Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Computer Vision Analysis of Broiler Carcass and Viscera

Jørgensen, Anders

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

COMPUTER VISION ANALYSIS OF BROILER CARCASS AND VISCERA

BY
ANDERS JØRGENSEN

DISSERTATION SUBMITTED 2018

AALBORG UNIVERSITY
DENMARK

Computer Vision Analysis of Broiler Carcass and Viscera

PhD Dissertation
Anders Jørgensen

Dissertation submitted June 13th, 2018

Curriculum Vitae

Anders Jørgensen

Anders Jørgensen has a M.Sc in Informatics with specialization in Vision, Graphics and Interactive System from Aalborg University, 2011. As part of his master he studied a semester at University of Auckland, New Zealand. After his graduation he worked three and a half years as a Research Assistant on the project "Bedre brug af hallen" at the Department of Media Technology. His focus for the project was method development, practical implementations and data collection.

His main research interest are computer vision and machine learning with emphasis on real world applications. As part of this PhD he has been employed at IHFood which specialises in computer vision for inspection of organic material.

Curriculum Vitae

Abstract

Poultry processing plants want to increase their slaughter rates but the manual food safety inspection is holding it back. The veterinarians conducting the inspection are already hard pressed to keep up with inspection rates of more than 3 birds per second. To maintain a high quality in future products, we need an automatic computer vision system to handle the inspection.

This thesis focusses on the inspection of broiler viscera, an area that has received little attention from the research community in the past years. Both the viscera and the carcass are examined during the health inspection and are equally important for the condemnation decision. Through this work methods have been developed for segmentation and diagnosis of the viscera set. A method have been put forward for segmentation of the viscera using both RGB and RGB-D data. Diagnosis methods have been developed for classification of healthy/unhealthy viscera and for classification of the most common liver diseases.

Ground truth is obtained by having veterinarians grade a subset of the data. Three veterinarians graded the same samples and the variation in their labels indicate that the inspection task is indeed difficult. Small deviations in the veterinarians' threshold for when a disease is a disease can result in large variations even in an unhealthy/healthy categorisation.

Data collection has been a crucial part of this work. The non-rigid nature of the viscera give a high variation in both position and appearance of the organs. On top of this comes the many different diseases which individually have a large variation as well. A large dataset is therefore key if a developed inspection system should generalise well.

Abstract

Resumé

Kyllingeslagterier ønsker at øge deres slagtehastigheder, men er begrænset af den manuelle fødevarekontrol. Dyrlægerne, som står for kontrollen, er allerede hårdt presset for at følge med ved de nuværende slagtehastigheder på over 3 fugle i sekundet. Hvis vi i fremtiden vil sikre en høj fødevarekvalitet, er der behov for et automatisk computer vision system til at foretage inspektionen.

Denne afhandling fokuserer på inspektion af kyllingens indmad, et felt der har ikke har fået meget opmærksomhed fra forskermiljøet i de senere år. Både indmad og krop bliver inspiceret efter slagtning og hver del er lige vigtig for den samlede veterinære bedømmelse. Gennem dette arbejde er der udviklet metoder til både segmentering og diagnosticering af indmadssæt. En metode er præsenteret for segmentering af indmad i både RGB og RGB-D data. Diagnosticeringsmetoder er udviklet både i forhold til en syg/rask inddeling og til klassifikation af de mest almindelige leversygdomme.

Ground truth er fundet ved at lade dyrlæger klassificere en del af vores data. Tre dyrlæger har klassificeret de samme samples og variationen i deres klassificeringer indikerer hvor svært det er at diagnosticere kyllingeindmad. Små afvigelse i hver enkelt dyrlæges grænse for hvornår en sygdom er til stede, kan give store variationer for selv en syg/rask inddeling.

Indsamling af data udgør en betydelig del af denne afhandling. Den løse opbygning af indmaden giver en stor variation i både placering og udseende for organerne. Oven i det kommer de mange sygdomme der hver især også indeholder en stor variation i deres udseende. Et stort datasæt er derfor nødvendigt hvis et automatisk inspektionssystem skal generalisere, så det kan bruges flere slagterier.

Resumé

Contents

Contents

Contents

Contents

Contents

Thesis Details

Thesis Title:	Computer Vision Analysis of Broiler Carcass and Viscera
PhD Student:	Anders Jørgensen
Supervisors:	Professor Thomas B. Moeslund, Aalborg University
	PhD Jens Fagertun, IHFood

The main body of this thesis consists of the following papers and technical reports. The technical reports are written for internal use at IHFood and desribe the unpublished work done during the PhD.

Papers

B *Detecting Gallbladders in Chicken Livers using Spectral Analysis*
Anders Jørgensen, Eigil Mølvig Jensen, Thomas B. Moeslund. British Machine Vision Conference, Machine Vision of Animals and their Behaviour (MVAB), Swansea, UK, 2015

C *Organ Segmentation in Poultry Viscera Using RGB-D*
Mark Philip Philipsen, Jacob Velling Dueholm, Anders Jørgensen, Sergio Escalera, Thomas B. Moeslund. Sensors, 18-1, 2018

E *Reaching behind Specular Highlights by Registration of Two Images of Broiler Viscera*
Anders Jørgensen, Jens Fagertun, Thomas B. Moeslund. Submitted to Journal of Applied Poultry Research

F *Diagnosis of Broiler Livers by Classifying Image Patches*
Anders Jørgensen, Jens Fagertun, Thomas B. Moeslund. Scandinavian Conference on Image Analysis (SCIA), Tromsø, 2017

H *Classify broiler viscera using an iterative approach on noisy labelled training data*
Anders Jørgensen, Jens Fagertun, Thomas B. Moeslund. Submitted to British Machine Vision Conference (BMVC) 2018

I *Weight Estimation of Broilers in Images using 3D Prior Knowledge*
Anders Jørgensen, Jacob Velling Dueholm, Jens Fagertun, Thomas B.
Moeslund. Submitted to Journal of Food Science

Technical reports

A *Preanalysis for Health Inspection of Viscera using Computer Vision*
Anders Jørgensen

D *Three Camera Setup for Image Capture of Moving Broiler Viscera*
Anders Jørgensen

G *Graded Data and Rerun of Results from Paper F*
Anders Jørgensen

Preface

My industrial PhD is a collaboration between the section of Media Technology at Aalborg University, Aalborg and IHFood, who specialises in computer vision inspection of organic material. This thesis is submitted as a collection of papers in partial fulfilment of a PhD study and it consists of both published and submitted papers, and technical reports meant for internal use at IHFood.

The work focusses on inspection of broiler viscera and automatic collection of images for a large dataset. The food inspection field is very competitive, so some information have unfortunately been redacted from the technical reports. The first part introduces the work, a state of the art of the research field and a conclusion. The papers and reports written during the PhD are appended in part two.

During my work I talked to multiple veterinarians who were so kind to help me understand what they were looking for during the inspection. I apologise beforehand for my crude simplification of the diseases and self-made terms used in this thesis.

I would like to thank my supervisor Thomas B. Moeslund, for his good ideas and opinions and for opening my eyes for computer vision during my bachelor. Great thanks to IHFood for hiring me as an industrial PhD and also to my supervisors at IHFood, first Eigil Mølvig Jensen who welcomed me to the company and later Jens Fagertun who always challenged my methods, with the purpose of improvement. A thank you to Mads Dabros as well, for his mentoring in Machine Learning, and the rest of the colleagues at IHFood for lending an ear when I needed advice. To my colleagues in the VAP group at AAU I would like to thank them for the discussions, off and on topic, and for saving me a seat and table when my stays in Aalborg became less and less frequent.

Anders Jørgensen
Aalborg University, June 13, 2018

Preface

Part I

Overview

Chapter 1

Introduction

Technology and science have always been a part of agriculture. The ancient Egyptians developed a convoluted irrigation system, utilising the flooding of the Nile, to become some of the first to employ farming on a large scale [1]. Mills for grinding corn were first driven by water then later by wind making it accessible in regions without access to a river [2]. With the invention of steam engines came the first tractors which would revolutionise farming and render horses obsolete, especially when the steam engines were replaced by internal combustion engines in the beginning of the 20th century [3]. Late 20th century came automation e.g in the shape of robotic milking systems [4] and today we see the rise of computer vision technology for sorting tasks, foreign body detection and quality control in food production plants [5].

The need for fast and reliable quality inspection is more relevant than ever. The earth's population of 7.5 billion people is increasing and doesn't seem to stabilise before reaching 11 billion in 2100 [6]. Feeding 3.5 billion extra people will require an efficient food production. At the same time we see that the gross domestic product (GDP) per capita is increasing as well. The GDP per capita in China, adjusted for purchasing power, is 15 times higher in 2016 than it was in 1990 [7]. With increased wealth follows an increased demand for high quality food.

The Danish export of poultry has increased from 1,900 tonnes in 1990 to 42,200 tonnes in 2016 [8]. Food makes up 18 % of the Danish export [9] and therefore quality and food safety inspection is important to avoid food scandals that can hurt the export. China saw this in 2008 when the melamine milk scandal broke [10]. Melamine, mostly used to produce the plastic melamine resin, was used to create a false level of protein content in watered down milk and infant formula. Six babies died and nearly 53,000 children were hospitalised [11] and countries were quick to ban import of all dairy products from China [12].

Europe had its own scandal in 2013 when it was uncovered by the Food Safety Authority of Ireland that multiple beef burgers contained both porcine and equine DNA [13]. A specific beef burger product from the Silvercrest plant in Co Monaghan contained 29 % horse meat. Although the undeclared meat posed no immediate health hazards, the scandal emphasised the importance of quality control during multiple stages of the supply chain.

As the world's population grows, so does the demand for meat. In 1961 an average person in the world consumed 23.09 kg meat per year, in 2013 that number has gone up to 43.22 kg meat per year. Poultry makes out the largest portion of the increase, from 2.88 kg to 14.99 kg over the same time span [14].

Per capita meat consumption by type, kilograms per year, World
Average per capita meat consumption broken down by specific meat types, measured in kilograms per person per year. Data is based on per capita food supply at the consumer level, but does not account for food waste at the consumer level.
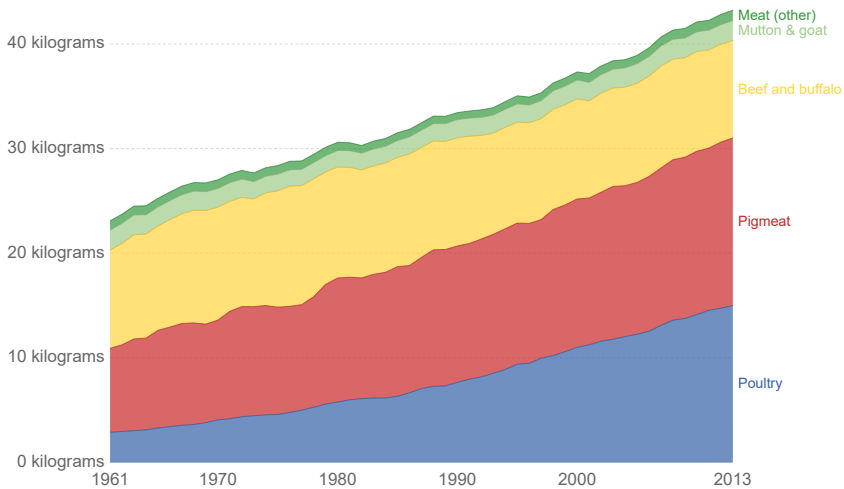
Source: UN Food and Agricultural Organization (FAO) OurWorldInData.org/meat-and-seafood-production-consumption/ • CC BY-SA

Fig. 1.1: Meat consumption per capita in the world. Poultry has seen the largest increase over the years.

In a time where meat production is criticised for being bad for the environment researchers have looked at the environmental impact of different types of meat. Although admittedly difficult to calculate, the green house gas emission of chicken is only around 1.6 $kgCO_2/kg$ compared to 22.6 $kgCO_2/kg$ for beef [15]. They found that there were large variations between countries, but it is clear that poultry is going to be a vital component in supplying the growing population with a source of protein. To ensure a viable quality inspection, in an industry that continuously pushes the production

speed, automation is needed. This thesis focusses on the aspect of automatic computer vision inspection of broilers with regard to both food safety and quality issues.

Chapter 1. Introduction

Chapter 2

Background

2.1   Food safety inspection of broilers

Slaughter and evisceration rates are typically between 12,000 and 13,500 birds per hour in Danish poultry processing plants [16]. This is the rate of a single processing line. Everything is highly automated to achieve such speeds, yet the health inspection is still performed by humans and it is becoming a bottleneck.

The veterinary authority in Denmark is The Danish Veterinary and Food Administration and the food inspection is handled by The Meat Control. Every slaughtered animal meant for consumption must go through an ante-mortem and post-mortem inspection. There is no minimum inspection time specified for poultry in neither EU or national regulations, however "*The speed of the slaughter line and the number of inspection staff present are to be such as to allow for proper inspection*" [17, 18]. With a slaughter rate of 12,000 birds per hour the average inspection time would be 333 ms for one person.

The post mortem inspection of poultry is being reviewed at EU level in the European Food Safety Authority (EFSA) [19]. The current regulations call for the post-mortem inspection to pay particular attention "*. . . to the detection of zoonotic diseases and diseases on OIE list*"[1] [17]. However, the critical zoonotic diseases, typically Salmonella and Campylobacter, are not visible in a post-mortem inspection. Salmonella is revealed by laboratory tests performed on samples collected in the rearing houses and in Denmark the birds are never brought to the slaughter house if a single sample is positive.

The real focus of the post-mortem inspection is on diseases that leave visible marks. For the carcass it is diseases like ascites, skin eczema, underweight and external factors like poorly bled birds and faecal contamination [18]. For

---

[1]OIE: The World Organisation for Animal Health

the viscera it is typically inflammation in the liver, heart or peritoneum but also faecal contamination and burst gallbladders which can be the result of a bad evisceration.

Another reason for doing a post-mortem inspection is from an animal welfare perspective. Defeathered birds can reveal scars, broken wings and bruises received either during transportation or in the rearing house. These cases can be reported back to the farmer, who can improve his or her practice. Slaughter houses also penalise the farmer for every unhealthy bird, so there is an economic incentive for animal welfare as well.

The post-mortem inspection happens at two stages. The first stage is the inspection of the carcass before evisceration and the second stage is the inspection of the viscera. The broilers and the viscera come by at a rate around 3.5 per second. Figure 2.1 shows the carcasses and viscera on the conveyor.

(a) Carcasses on the conveyor after evisceration.

(b) Viscera on the conveyor before inspection.

Fig. 2.1: Carcasses and viscera are transported on separate but synced conveyors after evisceration.

The inspection of the carcass is primarily done visually but the inspectors have a knife available, typically used when they suspect inflammation beneath the skin. They can take a broiler of the line for closer inspection and put it back up if it is deemed healthy.

The viscera are inspected by two persons standing next to each other along the conveyor. The viscera are positioned in a clamp that holds them so that the heart, lungs, liver and gizzard are mostly visible. As the evisceration is fully automatic there is a certain order to the organs' positions, however this cannot be guaranteed.

The carcasses travel on a separate conveyor along side the viscera. If an inspector spots a disease both the viscera and the carcass are removed from the line and temporarily stored together. The viscera are then checked again more thoroughly by a third inspector and if the viscera are deemed sick both the carcass and viscera are discarded, otherwise the carcass goes back onto

the line. The viscera are always discarded once removed from the line. The inspectors must also make sure that the evisceration has been done properly and otherwise remove the carcass from the line.
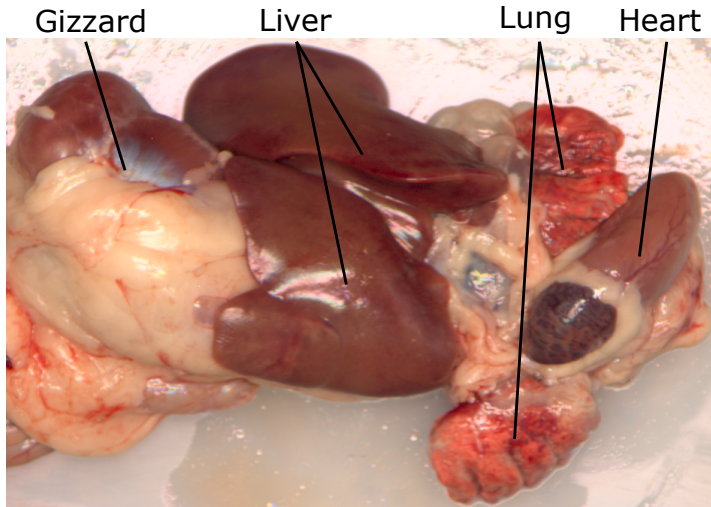
Fig. 2.2: Vital parts of the broiler viscera. Heart and liver are the most important organs for the visual health inspection as these are most frequently affected by diseases.

Computer vision could aid the inspectors at both stages. It could detect and remove unhealthy viscera from the conveyor and still leave the secondary inspection to a veterinarian, as it is today. Another solution would be a system with an extra conveyor. Of the many broilers that are slaughtered only about 0.7 % are unhealthy, so the vast majority will be healthy [18]. Samples that are clearly healthy would continue on the same conveyor and clearly unhealthy samples will be discarded automatically. The remaining samples would be side-tracked onto the extra conveyor for manual inspection. This way the veterinarians can focus on the borderline cases and they will have more time to do the inspection.

2.2 Related Research

As food production speeds are going up, so does the need for automation inspection. The availability of hyperspectral imaging, increasing frame rates and spatial resolution, has broadened the possibilities within food inspection. From detection of bruises on apples [20] to predicting the tenderness of red beef [21].

Inspection of broilers are no exception and it already happens at multiple stages in the production. Broilers are monitored in the rearing houses by the

farmers to keep track of the feed consumption to make sure they are following the ideal growth plan. This is a measure for the entire flock and then divided onto the number of birds. Additional systems have been developed to monitor weights on individual birds either mechanically or via computer vision systems [22, 23]. Even though only a small percentage of a flock is measured, it still gives a more varied image of the weight distribution.

Most research within this field has focused on the post mortem inspection of the carcasses. Many have investigated the problem of detecting skin tumours, which leaves a circular scar but often without immediate discolouration. However using hyperspectral imaging many have been successful in detecting skin tumours [24–28]. Systems developed by [25, 26] uses PCA to reduce the large number of wavelengths into a small number of principal components. Both collected their samples in a laboratory environment using a hyper spectral line scan camera. The two works reported classification rates on pixel level of 96 % and 97.18 %, respectively. [24, 28] employed feature selection methods which can be used to reduce the number of recorded spectral bands and thereby speed up recording time. With classification at tumour level they achieved a true positive rate of 90.6 % and 80.0 %, respectively, but [24] also reports over twice the number of false positives compared to [28].
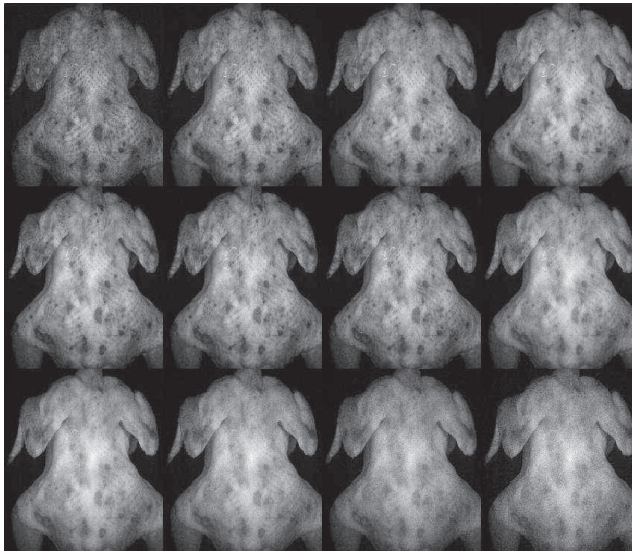
Fig. 2.3: Carcass with tumor captured at 12 different wavelengths. Source: [27]

Researchers at USDA Agricultural Research Service have long been working with universities to develop a system for classifying wholesome and unwholesome birds in a production environment. By unwholesome they refer

to systemically diseased chickens which are detectable by their smaller size or reddish skin. In [29] a dual camera system was used to capture images at 570 nm and 700 nm of every bird. The images were divided into 107 blocks and the average intensity of each block was fed to a neural network. The system achieved 94 % and 87 % accuracy for wholesome and unwholesome chicken, respectively, working in-line at a speed of 91 birds per minute.

Starting with an offline multispectral setup, [30] investigated sets of three and four wavelengths captured with an area scan CCD camera and achieved accuracies of 95.7 % and 97.7 %. Through several iterations, now using a line-scan camera, [31–33] developed a system capable of running at 140 birds per minute identifying 99.8 % of the wholesome birds and 97.1 % of the unwholesome. The system used fuzzy logic to score each pixel from the breast area and the average of all scores to classify the broiler.

A lot of the development in quality control is going on in the industry. The companies Marel, Meyn, Foodmate and Baader Linco all have products for quality grading of broiler carcasses, but the details are very limited. Typically, they use a front or back camera or both. Marel reports that they detect broken wings, red and blue bruises, faecal stains, remaining feathers and skin damage [34]. Meyn looks for similar defects but also have a foot pad inspection system for detecting burns in the footpad [35]. The feet are typically exported to Asia, but the burns are also a measure for broiler welfare. A high percentage of foot pad burns could indicate poor or inadequate litter in the rearing house. Foodmate specifies similar features as Marel and Meyn but also includes missing parts, one leg hanging, hock burns and bile stains [36]. Hock burns are like foot pad burns while bile stains typically occur due to a bad evisceration. Baader Linco detects defects similar to those by Foodmate [37]. Besides the usual front and back camera, they have the option to install two side cameras, which can increase the detection rates beneath the wing. None of the companies specifies any detection rates.

During the evisceration there is a risk that the broiler carcasses become contaminated with faeces, ingesta, bile, etc. These can carry bacteria or foul taste and infected carcasses must either be removed from the line or washed. [38] developed a system for detecting faeces and ingesta on the surface of broiler carcasses at a rate of 140 birds per minute. They used a 3-CCD camera with optical trim filters at 515.4 nm, 566.4 nm and 631 nm and achieved an accuracy of 96 % in detection rate but also reported "*many false positives*". [39] developed a fully automatic system for detecting contaminants on the carcass surface. The computer processing the two images, captured at 500 nm and 710 nm, also control a spray system to wash and disinfect the carcasses. The true positive rate of the system is 91.6 %.

Wooden breast is a relatively new phenomenon within poultry production. The abnormality affects the breast muscles where areas become pale and hard [40]. [41] developed a system for rapid on-line detection of wooden

Fig. 2.4: Image of a commercial quality inspection system. Source: [34]

breast using NIR spectroscopy. Measuring the protein content, they achieved a 100 % accuracy in detecting wooden breast from normal fillets on a test set of 79 samples. The work was done on chicken fillets after cut-up and with the skin removed.

Inspection of viscera has received less attention from both research institutions and companies. The viscera are transported alongside the carcass after evisceration until it has passed the manual inspection, after which the edible organs are harvested from the viscera set. Quality inspection can happen after harvest where the organs can be presented individually but food safety inspection must be done while the viscera travels along side the corresponding carcass. Here it can be a challenge as the organs are unorganised and can cover each other.

[42] developed a method to detect spleen enlargement in turkeys as this abnormality is related to virus diseases. The spleen and the liver show similar colours in regular RGB images but they found that the spleen was easy segmentable in UV lighting and achieved a detection rate of 95 %. [43] photographed both heart and liver in an effort to classify the four classes: airsacculitis, cadaver, normal, and septicaemia. Using white light, they found that the liver couldn't be used to separate normal and septicaemia and the heart couldn't be used to separate normal and cadaver. By combining the models for both organs, they were able to classify the birds into the four classes with an accuracy of 82.5 %. Similar discoveries were made by [44] who used NIR to capture images of chicken hearts to detect the diseases: airsacculitis, ascites, normal, cadaver, and septicaemia. Cadaver proved must difficult again and scored 82% true positives in the five-class problem, whereas the other diseases were detectable with rates between 92 % and 100 % of the 125 test samples. In work by [45] they used visible and near infrared spectroscopy to

separate chicken livers into the classes septicemic and normal. Samples were cut from the center of the liver and a neural network was trained to classify the 100 test samples. Classification rates of 98 % and 94 % was achieved for normal and septicemic livers, respectively.

These works all used harvested organs where the hearts and livers were removed from the viscera set and the conveyor. Most research is from around 2000 and no newer research seems to be available and no research seems to focus on inspection of poultry viscera in-line. Work has however been done for the viscera of pigs. [46] developed a method to segment the offal into heart, liver, lungs, diaphragm, and upper portion. After evisceration the pig's viscera are presented similar to broiler viscera. Using Auto-context they report F-scores between 0.778 and 0.957 for the five parts of the viscera.

2.3 Summary

Many proposed methods for detection and classification of diseases in broilers focus on a single disease e.g. tumours or septicaemia. But condemnations of broilers can be due to many causes. [47] mentions at least 12 causes in a Danish report about aiding tools for the food safety inspection. [48] mentions 16 causes in their work on prevalence rates in British reared broilers. In order to replace or provide meaningful aid for the veterinarians an automatic inspection system needs to detect most of these diseases.

The viscera have received little focus from researchers compared to the carcass, but the inspection is equally important from a food safety perspective. This inspection must take place while the viscera are still hanging on the conveyor as the traceability is lost when the organs are harvested. This increases the technical challenge as the organs can hide each other and multiple cameras are needed to cover the set.

In 2011, EFSA released a report where it was suggested that palpation and incision techniques should be omitted during the post-mortem inspection of pigs [49]. The risk of cross-contamination was found to be greater than the risk of missing a disease that was not immediately visible. In cases of reasonable suspicion, the pig should be removed from the slaughter line and treated separately. Moving towards visual inspection only, increases the competitiveness of a computer vision system.

Research within the topic of food inspection is mainly driven by the industry who wants to produce better products but cheaper and faster. Even though broiler processing plants already operate at impressive speeds most equipment is manually adjusted and far from the digitalisation heralded by Industry 4.0. An automated inspection system would move the inspection knowledge directly into a database which can be monitored for sudden changes in trends. A rise in the number of missing hearts from the viscera

sets could indicate that the evisceration equipment needs adjustment. There is a large potential in detecting errors and malfunctions early in order to keep the conveyors running at maximum capacity.

Chapter 3

Scope of the thesis

The work presented in this thesis originates from the hypothesis that "Computer vision can be used to automate food safety inspection of broiler viscera". This implies automatic diagnosis of diseases in images captured of the viscera as they move along the conveyor. To replace the manual inspection, it must detect and recognise multiple diseases just like the veterinarians performing the inspection task today.

The viscera must be diagnosed in-line which means that the organs cannot be neatly arranged for the camera. A part of this thesis has therefore focussed on segmentation of the viscera set, as a preliminary step before the diagnosis.

During the PhD, an additional research question arose, namely whether 3D features can be used to increase the performance of weight estimation of broiler carcasses in 2D images. The application could replace expensive in-line physical scales.

These topics; diagnosis, segmentation, and weight, all require data for which there exist no publicly available dataset. Data collection and annotation have therefore been a large part of this thesis.

All work are described in articles or internal reports and are divided into these four topics:

- Data collection and annotation: Appendix A, D, E

- Organ segmentation: Appendix B, C

- Computer vision based diagnosis: Appendix F, H, G

- Computer vision based weight estimation: Appendix I

Figure 3.1 shows the relations between the works presented in the thesis. The ellipses represent datasets and the the boxes contains references to the individual works.

Fig. 3.1: The structure of the work in the thesis. Ellipses represent datasets, boxes represents papers and reports.

The content of the four topics will be described briefly in the following sections. The details for each work are available in part II.

3.1 Data collection and annotation

Data is vital when applying any machine learning algorithm to a new domain and in computer vision this data comes in the form of images. With organic material a large dataset is key, as it is near impossible to capture the entire variance of a population in a finite dataset, which is important if the developed method should generalise well. Data acquisition should therefore be automatic in order to capture thousands of images. Another problem with data collection is when the prevalence of some classes is very low. As only 0.7 % of the broilers are unhealthy, an automatically collected dataset will be very imbalanced towards healthy birds. This must be dealt with either through the data collection or by the classification method.

[50] designed a dual camera system capable of capturing images at 540 nm and 700 nm using interference filters. Their setup was mounted on a cart and used a proximity sensor to detect shackles when recording images online at a poultry processing line. The two grey scale cameras were capable of recording images at the required rate of 91 bpm at a resolution of 768x484. [51] used a EMCCD camera coupled with a line scan spectrograph to acquire multi spectral images of broilers at a rate of 180 bpm. The images had a height of 502 pixels and recorded at three predefined spectral bands achieving between 95 an 120 scan lines per bird.
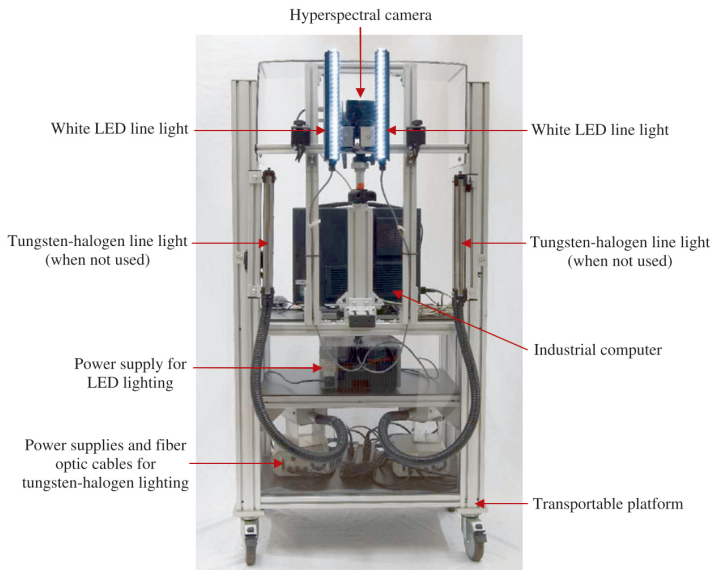
Fig. 3.2: Image of a setup for recording images of broiler carcasses online. Source: [51]

Our first step was to identify important wavelengths that could best de-

scribe the diseases on the liver and the heart. This work is described in appendix A. These images were recorded manually in a multi spectral setup capable of capturing images in the range of 430 nm to 720 nm in 10 nm steps. Figure 3.3 shows an example of a captured viscera set.

Fig. 3.3: 30 images of a viscera set captured at wavelengths between 430 nm and 720 nm.

This setup used 30 seconds to capture a full sweep, but helped us identify three specific wavelengths to use in our automatic recording setup. This setup, described in appendix D, was designed to be installed in-line at Danpo A/S in Aars and automatically record images of every passing viscera. Using three cameras and three light sources, five images were captured for each viscera, as two of the cameras captured images with alternating light sources. The cameras have a CMOS sensor with a resolution of 4096x3000. A vital part of this setup is the automatic labelling of the images. This is accomplished with a sensor that detects if a viscera set has been removed from the conveyor. Images of removed viscera were labelled unhealthy and saved while most of the healthy viscera images were deleted to obtain a more balanced dataset. The system operates at a production rates of 210 bpm and is triggered by a inductive sensor so that it automatically adjusts to the conveyor speed. Figure 3.4 shows the three cameras installed at Danpo A/S.

The recording system was constructed in collaboration with Linco Food Systems, a company that designs and builds poultry processing systems. They provided the metal constructions and the trigger signals, while camera, light and software was designed and built by the author. The system was installed in October 2016 at Danpo A/S. Recording can easily be turned on and off through a remote connection and it has since recorded over 60.000 viscera sets.

The use of multiple cameras was to investigate how many cameras that

Fig. 3.4: Recording system designed as part of this work.

was needed to cover the entire viscera set. The purpose of the alternating light sources is to construct a method that produces images without specular highlights without the use of polarising filters. This work is described in appendix E. The wet and glossy surfaces of the viscera produce some unpredictable specular highlights that are hard to avoid by light design alone. The developed method detects specular highlights and recover the area from the alternate image with another light source, using a colour normalisation technique shown in figure 3.5.
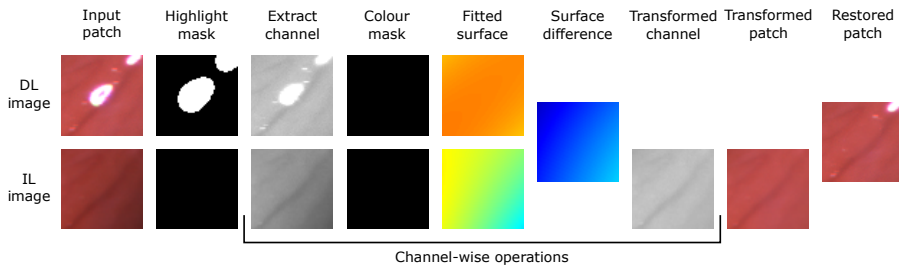
Fig. 3.5: Specular highlights are detected and pixels from the alternate image are colour corrected and imported into the first image. Source: Paper E.

Because the viscera travels along the conveyor between capturing the two images, the method requires a precise registration to map the the transformation between the two images. This proved difficult due to the low amount of landmarks on the liver surfaces, but it was handled by constraining the transformation between the two images. The results were visually pleasing

and a qualitative test on synthetic specular highlights showed good results as well. Figure 3.6 shows the result of a spectral highlight removal on the heart.

(a) Direct Light (DL).   (b) Restoration result.   (c) Indirect Light(IL).
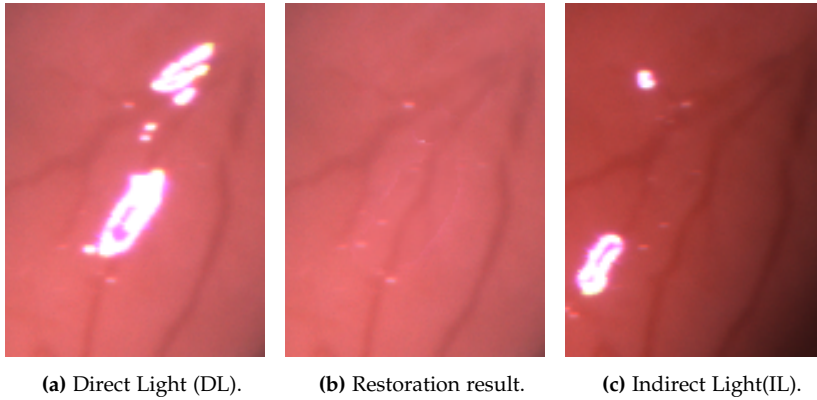
Fig. 3.6: Restoration on the heart. The result image (b) has the frontal light of the DL image, but without the specular highlights. Source: Paper E.

A large part of the PhD was used on performing the initial data collection with the multispectral setup and designing the automatic three camera recording system. Having access to large amounts of data is crucial to build models that generalise well, especially with the rise of deep neural networks. And when the prevalence of diseased viscera is much lower than healthy viscera, it is simply infeasible to capture a balanced dataset manually. Having a tap of data installed directly in a slaughter house is a major plus for further development of an automatic health inspection system.

3.2   Organ segmentation

A good classification method often relies on a good segmentation, as it constrains the problem to know were in the image to look. But segmentation of viscera is not a simple task. The colour and size of the organs can vary greatly due to natural causes and diseases can change an organ's visual appearance completely. Most previous work done on inspection of poultry viscera is done on viscera that has been harvested from the viscera set. However, to use the findings for food safety measures the inspection must take place while the viscera can be traced back to the broiler carcass, in order to reject it.

The viscera have a weakly constrained order as it is removed from the carcass. The clamp, in which it is transported on the conveyor, typically connects behind the heart and the lungs. Beneath it hangs the liver, spleen and gallbladder and then comes the gizzard and intestines. Figure 3.7 shows

four examples of viscera as it hangs in the clamp after evisceration. The gallbladder and spleen are often hidden behind the liver and the location of the heart and the lungs of vary a lot.
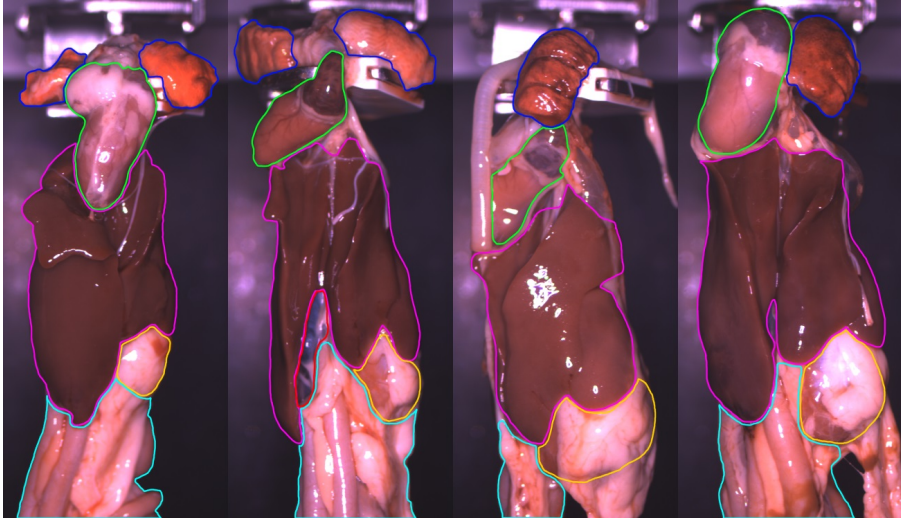
Fig. 3.7: Viscera sets as they are presented after evisceration. Blue = lungs, green = heart, purple = liver, red = gallbladder, yellow = gizzard and cyan = intestines.

During harvest of the edible organs (Liver, heart and gizzard) it can be a problem if the gallbladder is picked with the liver. The gallbladder contains bile which has a strong foul taste that can contaminate the livers if it is punctured. We developed a method to detect if the gallbladder was still attached to the liver after harvest, appendix B. Using a multispectral setup we captured images from 430 nm to 720 nm in 10 nm intervals which were used to separate the three classes: liver, gallbladder and fat. We found that the number of wavelengths could be reduced to two, namely 600 nm and 720 nm, without loss of performance.

In an effort to perform segmentation on an entire viscera set we recorded a separate dataset using an RGB-D sensor. As the colour of most of the organs is a shade of red the idea was to utilise the depth information to separate the organs. This work is described in detail in appendix C. 151 sets were recorded from four views resulting in 604 images. An example of a recorded viscera set can be seen in figure 3.8.

(a) RGB data. (b) Depth data. (c) Point cloud.
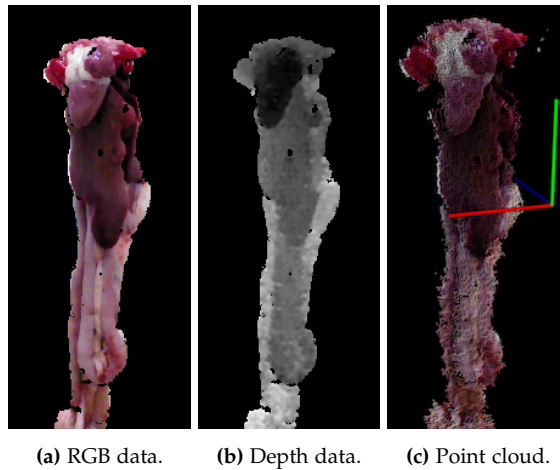
Fig. 3.8: Recorded RGB-D data combined in a point cloud. Source: Paper C.

After over-segmenting the images using VCCS [52], unary features are used in a Random Forest to label the supervoxels. Conditional Random Fields is then used to optimise the labels using edge features. The segmentation was done both with and without the depth information in order to investigate the impact of the extra dimension. Including the depth gave a performance boost of nearly two percentage points on average for the Jaccard index of the four classes heart, liver, lung and miscellaneous. An additional increase in performance was seen when adding unary features from a convolutional neural network. These features come from activation maps in network and totalled 4224 CNN feature maps. With CNN features, the segmentation achieved a total class average of 78.11 %. The results of the three different models can be seen in figure 3.1.

Method	Features	Misc.	Heart	Liver	Lung	Class avg.
RF+CRF	2D	90.66	57.69	80.59	**68.18**	74.28
RF+CRF	3D	91.28	63.02	82.38	67.43	76.03
RF+CRF	3D+CNN	**91.58**	**70.17**	**83.64**	67.05	**78.11**
ASA		96.32	88.65	88.63	82.49	89.63

Table 3.1: Results for the organ segmentation. Pixel-wise Jaccard index for the four classes. ASA (achievable segmentation accuracy). Source: Paper C.

A good segmentation of the viscera set simplifies the classification task when a diagnosis must be given. Classifiers can be trained per organ instead of one overall classifier for the entire viscera set. The results showed that both 3D and CNN features increased the performance of the segmentation.

For an in-line system the depth sensor might be impractical as it adds complexity to the hardware and the sensor might not be fast enough to record moving objects. CNN features can however be implemented in software and might be more suitable for a real-time application.

3.3 Computer vision based diagnosis

The manual food safety inspection is becoming a bottleneck at poultry processing plants who want to increase the production speed. The veterinarians who are responsible for the inspection are already working hard to keep up with the more than 200 birds, one person must inspect every minute. If production speeds should be increased any further some form of automatic inspection is needed.

Diagnosing the viscera is not an easy task even for veterinarians, especially at rates over three per second. Diseases develop over time and look different depending on the severity. Figure 3.9 shows a collage of viscera sets where all livers are diagnosed with cobblestone, a disease where the liver develops bulges and dark trenches. All viscera sets should be removed from the conveyor yet the examples in the bottom row of the image can be difficult to spot on the conveyor. The decision could also be affected by a previous decision, also known as the question order bias [53]. Veterinarians might also have different thresholds for different diseases. It is believed that a computer vision system could not only increase the detection rate but also improve the consistency of the classification.
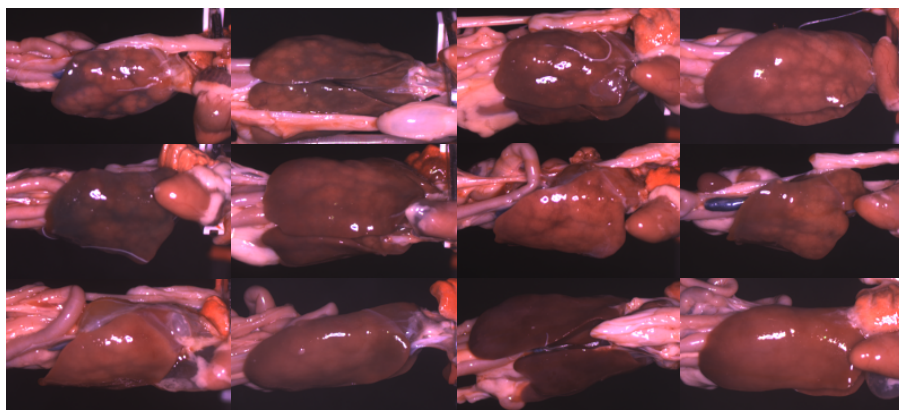
Fig. 3.9: Collage of viscera sets with livers all diagnosed with cobblestone. Source: Paper F.

It has not been possible to find work by others who have done diagnosis of entire viscera set. Others have focussed on the carcass or individual diagnosis

of harvested organs, as described in section 2.2. In appendix F we diagnose the liver in the viscera set using the RGB version of the segmentation method developed in appendix C. Using 1,476 images, graded by a veterinarian, the livers were first classified patch-wise by a convolutional neural network. The network was trained on patches, a region of 75x75 pixels, manually extracted for the five classes: healthy, cobblestone, perihepatitis, necrotic hepatitis and miscellaneous. Examples of extracted patches can be seen in table 3.2.
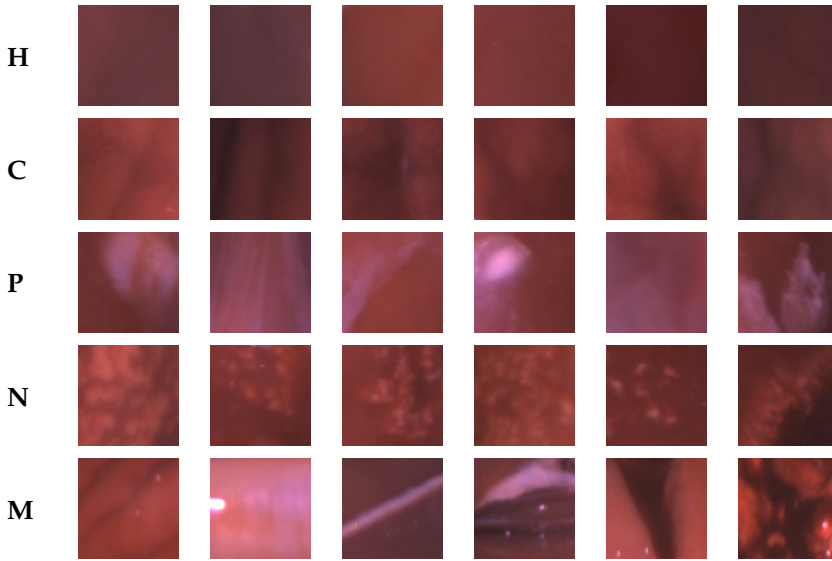
Table 3.2: Patch examples from the five classes used to train the convolutional neural network. Source: Paper F.

The convolutional neural network was used to evaluate patches in an input image, in a correlation-like fashion and in steps of five pixels. This resulted in five probability maps, one for each class in the network. Eight simple features was extracted from each probability maps and fed to a Random Forest Classifier which performed the diagnosis of the liver. The average of 10 five fold cross validations was used to present the classification results. The confusion matrix revealed true positive rates between 71.0 % and 80.3 % for the four classes.

The results were later revisited in appendix G. In an investigation about the difficulty of grading the viscera, we had three veterinarians grade the same data. Of the 3,984 graded images they agreed on 1081 healthy images and 1073 sick images, totalling 54.1 % of the images. Individually they had 97.0 %, 87.4 % and 99.2 % in these two classes, with the rest described as non-gradeable. This indicates that the grading is difficult and can be very

subjective. Bear in mind that the veterinarians only have one third of a second to make the decision in-line, whereas here they had the time they needed. By removing the samples where the veterinarians did not agree, from the dataset used in appendix F, the true positive rates increased to between 74.6 % and 94.5 %. The overall classification error dropped from 22.4 % to 14.4 %.

In a more holistic approach to detect whether a set of viscera should be removed from the conveyor or not, we used the images graded by the veterinarians for validation and test, while training on the images labelled automatically by the recording system described in appendix D. The labels given automatically are not always correct, so a method was used to relabel the mislabelled samples. In an iterative approach, described in appendix H, we used a convolutional neural network to find wrongly predicted training images so that these could be labelled manually. Figure 3.10 shows the number of relabelled training samples per iteration. P_D describes the average certainty for the wrong prediction and it can be seen that it drops for iteration five and six. This indicates that the classifier is no longer certain and it was observed that the samples got harder to label as well.
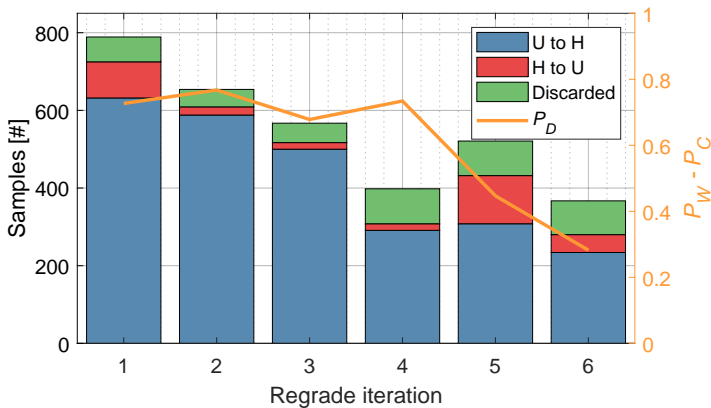
Fig. 3.10: Relabelled training samples for each iteration. P_D describes the certainty of the wrong prediction. U = Unhealthy and H = Healthy. Source: Paper H.

After the sixth iterations, the convolutional neural network was trained one last time and achieved an overall test classification error of 14.0 %. 6000 images were labelled during the six iterations which means only 17.68 % of the training data was manually labelled.

Labelled data is required for supervised learning and deep neural networks need lots of it. But hiring external experts to label your data is expensive and time consuming. We demonstrated that by using noisy automatically labelled data for training one can still obtain good classification results for a difficult problem. The relabelling was done by the author after receiving

little training from a veterinarian.

We have developed two methods for performing diagnosis, one which focused on classifying liver diseases and one for determining the overall healthiness of the viscera set. In a practical setup these two could be combined so the overall classification is used to determine whether the viscera should be removed from the conveyor as a food safety measure. A second method could then inspect the viscera set for statistical purposes by determining the disease or diseases. The last step is then no longer time critical as the decision has already been made to remove it from the conveyor.

3.4 Computer vision based weight estimation

The weight of each broiler is used in the processing plants to optimise the production and increase yield. This is done by adjusting the cut-up equipment to the weight and thereby getting the optimal cuts when parting the carcass into legs, wings and breast. But mechanical weight are expensive and tend to decrease in accuracy over time as the components gets worn. Camera based methods can only estimate the weight from the size of the bird and will therefore inherently be less accurate, but being cheaper they can be installed multiple places along the processing line. The images can also be used to estimate the size and weight of individual body parts, providing more details than a mechanical weight.

Camera based weight estimation is used in many areas of the industry and the principle is very simple. Extract features and find a model that best describes the correlation between the features and the weight. Essentially it comes down to finding appropriate features. [54] developed a system to estimate the weight of live broiler from a top view image captured in a rearing house. They found a strong correlation between the pixel area of the bird and the weight and managed to estimate the weight of the birds from day 7 to day 42. [55] developed a similar system but for weight estimation of pigs. Using only the pixel area of the pigs they achieved an accuracy of 96.2 % for individuals. Both of these systems monitor the animals as they grow, meaning that they can have multiple measurements over time. In production environments the weight must often be estimated from a single image, perhaps using multiple views. [56] used a single camera and two mirrors to get three views of the produce on the conveyor. Working at 30 items per second the method demonstrated an overall estimation error of 4.4 % corresponding to maximum two gram for the largest items.

IHFood have collected a large dataset of broiler images from one of their ClassifEYE systems. Single view weight estimation is limited by the fact the we cannot see what happens in the axis away from the camera or on the back side of the object. We opted to circumvent this by adding prior knowledge

in form of a statistical 3D model based on 45 3D scans of broilers collected at Danpo A/S in Aars. Examples can be seen in figure 3.11. The statistical model was reduced to seven components with PCA and then fitted to eight landmarks found on the broiler in the 2D image. The fitting was constrained to 2D as the orientation of the bird in the image is known. From the fitted statistical model 3D features were extracted and used in combination with the 2D features. The added 3D features gave a small yet significant performance improvement and resulted in a mean average percentage error of 3.47 %. 102,412 images were used for training and 34,060 images were used for test.
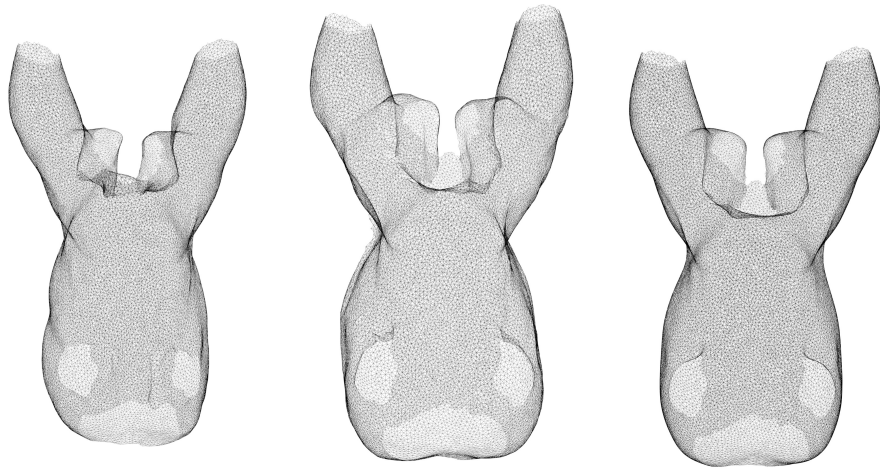
(a) 3D model of 1160 g bird.　　**(b)** 3D model of 1770 g bird.　　**(c)** Mean model of all birds.

Fig. 3.11: 45 birds have been 3D scanned and combined into a mean model shown in figure c. All images are the same scale.

The statistical model was created from just 45 3D scans and these were obtained from broilers from a Danish slaughter house, whereas the 2D images came from a British slaughter house. With the 3D features we saw a small performance increase, but the method should be tested with a larger number of 3D scans and birds from the same slaughter house to find the true potential of the method. The method was still fast enough to operate real-time, so the practicality falls on how well you can optimise the 3D scanning process when building the statistical model.

Chapter 4

Contributions

This chapter will summarise the contributions made through this thesis. It is believed by the author that this is the first large study on the topic of automatic food safety inspection of broiler viscera. The main contributions are:

Identified important wavelengths for liver and heart diseases. The works in appendix A and B use images recorded with a multi spectral setup to determine the wavelengths that can best distinguish diseases on the liver and heart, and separate the gall bladder from liver and fat.

Large dataset collected in-line. The report in appendix D describes the creation of an in-line image capture system used to create a large dataset of viscera as they are currently presented for the veterinarians at the slaughter house. Working on images directly from the production line makes the transition from research to application much shorter. Appendix E presents a method for obtaining an image without specular highlights and real data restored underneath.

Illustrated the subjectivity in todays inspection. The veterinarians responsible for the health inspection today is doing a very good job at a very difficult task. But the report in appendix G shows that the grading can be very subjective simply because the inspection task is so complex and there is no easily definable thresholds. The difference between the veterinarians' grades must be considered when evaluating a computer vision based system.

Segmentation method for broiler viscera. A method for segmentation of broiler viscera is presented in appendix C. A good segmentation algorithm

can be vital for a correct classification of the diseases present in the viscera set.

Methods for diagnosing broiler viscera. Diagnosis of the viscera has been approached in two different ways. Appendix F presents a method in which the liver is first found with a segmentation method and then classified patch-wise by a CNN. The classifier reports the disease affecting the liver. The work presented in appendix H takes a more holistic approach and takes the entire viscera set as input. A classifier then determines whether the viscera must be removed from the conveyor or not.

Improved weight estimation with 3D prior knowledge. Weight estimation from images can be used in many areas of the production. By adding prior knowledge from a statistical 3D model we improved the performance for weight estimation of broiler carcasses, while still capturing the images with a normal RGB camera. This work is presented in appendix I.

4.1 Work published by the author

Part of the PhD

- A. Jørgensen, J. V. Dueholm, J. Fagertun, T. B. Moeslund, *Weight Estimation of Broilers in Images using 3D Prior Knowledge*, Submitted to Journal of Food Science

- A. Jørgensen, J. Fagertun, T. B. Moeslund, *Classify broiler viscera using an iterative approach on noisy labelled training data*, Submitted to British Machine Vision Conference (BMVC), Newcastle upon Tyne, 2018

- A. Jørgensen, J. Fagertun, T. B. Moeslund, *Reaching behind Specular Highlights by Registration of Two Images of Broiler Viscera*, Submitted to Journal of Applied Poultry Research

- A. Jørgensen, E. M. Jensen, T. B. Moeslund, *Detecting Gallbladders in Chicken Livers using Spectral Analysis*, Machine Vision of Animals and their Behaviour, BMVC Swansea, 2015

- M. P. Philipsen, J. V. Dueholm, A. Jørgensen, S. Escalera, T. B. Moeslund, *Organ Segmentation in Poultry Viscera Using RGB-D*, Sensors, 18-1, 2018

- A. Jørgensen, J. Fagertun, T. B. Moeslund, *Diagnosis of Broiler Livers by Classifying Image Patches*, Scandinavian Conference on Image Analysis (SCIA), Tromsø, 2017

- A. Jørgensen, *Preanalysis for Health Inspection of Viscera using Computer Vision*, Internal report for IHFood

- A. Jørgensen, *Three Camera Setup for Image Capture of Moving Broiler Viscera*, Internal report for IHFood

- A. Jørgensen, *Graded Data and Rerun of Results from Paper F*, Internal report for IHFood

Not part of the PhD

- R. Gade, A. Jørgensen, M. M. Jensen, T. Alldieck, M. Abou-Zleikha, M. G. Christensen, T. B. Moeslund, M. K. Poulsen, R. G. Larsen, J. Franch, *Automatic analysis of activities in sports arenas using thermal cameras*, 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), Workshop on Human Tracking and Behaviour Analysis, Naples, 2016

- O. B. Jensen, T. B. Moeslund, R. Gade, A. Jørgensen, C. Lassen, M. Frølund, I. S. G. Lange, D. L. Murzea, M. A. Meyer, N. T. Andersen, *Termiske kameraer i fremtidens trafikplanlægning*, Trafik & Veje, 93-5, 2016

- R. Gade, T. B. Moeslund, S. Z. Nielsen, H. Skov-Petersen, H. J. Andersen, K. Basselbjerg, H. T. Dam, O. B. Jensen, A. Jørgensen, H. Lahrmann, T. K. O. Madsen, E. B. Skouboe, B. Ø. Povey, *Thermal Imaging Systems for Real-Time Applications in Smart Cities*, International Journal of Computer Applications in Technology, 53-4, 2016

- R. Gade, A. Jørgensen, T. B. Moeslund, *Long-Term Occupancy Analysis using Graph-Based Optimisation in Thermal Imagery*, IEEE conference on Computer Vision and Pattern Recognition, Portland, 2013

- R. Gade, A. Jørgensen, T. B. Moeslund, R. K. Jensen, *Automatic Occupancy Analysis of Sports Arenas*, European Association for Sports Management (EASM) conference, Aalborg, 2012

- R. Gade, A. Jørgensen, T. B. Moeslund, *Occupancy Analysis of Sports Arenas Using Thermal Imaging*, International Conference on Computer Vision Theory and Applications, Rome 2012

Chapter 4. Contributions

Chapter 5

Conclusion

The work presented in this thesis all concern the inspection and analysis of broilers and broiler viscera. Especially the viscera have been in focus as these have received little attention in the last 15 years and present a difficult challenge due to the way they are presented at the slaughter houses.

Two methods have been proposed for classification of the viscera. In an effort to remove the unhealthy viscera sets from the conveyor a method was presented which could classify the unhealthy from the healthy viscera sets. This was done by training a CNN on weakly categorised data and iteratively regrading mislabelled samples manually. This way under 20 % of the training samples needed to be manually graded. A second method was presented which could classify the liver into healthy and the three most frequent diseases. By classifying the liver in patches, the method generated intensity plots for each of the four classes and a Random Forest was used to give the overall classification for the liver.

Using an RGB-D sensor in an offline setup a method was proposed to segment the viscera into liver, lung, heart and other. A challenging task as the organs often are occluded by each other or by fat. The use of 3D information provided by the depth sensor improved the performance which was improved yet again with the use of CNN features.

Weight estimation of broilers from images is already used in the industry. In an effort to improve the accuracy we proposed an approach of adding prior knowledge from a statistical 3D model. This model was generated from 45 3D scans of broilers collected at a Danish poultry processing plant. Using Active Shape Modelling it was possible to extract 3D features for each image, giving a small improvement in accuracy.

The data used for the classification methods stem from a camera setup built for recording images of viscera in-line. The setup included three sets of camera and light, encased in stainless steel to cope with the high humidity

in the production environment. Due to the low prevalence of diseases in the broilers the system was designed to have an inbuilt sorting mechanism to avoid cluttering the hard drive with unwanted images of healthy viscera sets. The system now runs automatically and generates thousands of images per week.

Grading viscera is not an easy task, even for veterinarians. Each disease exists on a continuous scale of severity and size, and multiple diseases can be present at the same time. By having three veterinarians grade the same data we investigated the difficulties in diagnosing the viscera and the results showed a large variation in the grades. This also indicates that a 100 % detection rate is unobtainable even for humans and can therefore not be expected of a computer vision system either. Such a system should however be able to perform the detection more consistently.

Through the work presented in this thesis we have studied the main components required to answer the hypothesis that "Computer vision can be used to automate food safety inspection of broiler viscera". We have built a system for recording images in-line at full conveyor speed and obtained good performance both in the binary case of healthy/unhealthy and in classifying the most frequent diseases. It is the author's believe that this technique could be commercialised within a few years.

5.1 Future work

There is a huge potential in developing a commercial system that can perform fast and reliable food safety inspection of broilers and their viscera. But the field is changing, and new legislation could change the objective of the inspection and the reason for condemnation. In the future, carcasses might not be discarded due to liver diseases if there is no indication that the disease has affected the meat. But even in that case it is still important to detect and removed the diseased broilers and viscera from a quality perspective.

For any system the processing time will be important. One of the purposes of an automatic inspection is to increase the slaughter rates towards 15,000 birds per hour. That translates to 240 ms of processing time per bird and as multiple views are required to inspect the entire viscera set, time could become a challenge. But the wall time before an action should be made at the conveyor will likely be longer, so given enough parallel computing power it should be possible.

The veterinarians have a fundamental knowledge that is difficult to put into a computer vision system. As developers of classification methods, we want to simplify the decision making into "which one of these classes is it?". But a viscera might be slightly affected by one disease and more severely by another disease and when these two diseases are combined it is some third

disease. Trying to put this complex medical knowledge and know-how into simple classes is difficult and requires a great deal of attention.

To operate as an automatic food safety inspection system, it needs to be approved by the appropriate authorities. A partial system that could produce a guaranteed clean stream, as for example, 80 % of the birds would likely be easier to get approved and it would still be an aid for the veterinarians who could instead concentrate on the difficult decisions. Processing speeds are not going to slow if we are to keep up with the increased demand for chicken meat. With new legislation and a demand for a more uniform inspection the future is bright for the use of computer vision for food safety inspection.

Chapter 5. Conclusion

References

[1] H. Kees, T. James, and I. Morrow, *Ancient Egypt: A Cultural Topography*, ser. A Phoenix book. University of Chicago Press, 1977. [Online]. Available: https://books.google.dk/books?id=iLuMAAAAIAAJ

[2] D. T. Swift-Hook, Ed., *History of Wind Power*, 2012, pp. 41–72.

[3] H. Bonnett, *Discovering traction engines*. Shire Publications, 1975.

[4] W. Rossing and P. Hogewerf, "State of the art of automatic milking systems," *Computers and Electronics in Agriculture*, vol. 17, no. 1, pp. 1–17, apr 1997.

[5] D.-W. Sun, *Computer Vision Technology in the Food and Beverage Industries*. Woodhead Pub Ltd, 2012. [Online]. Available: https://www.safaribooksonline.com/library/view/computer-vision-technology/9780857090362

[6] United Nations, Department of Economic and Social Affairs, "World Population Prospects: The 2017 Revision," Tech. Rep., 2017.

[7] The World Bank, "GDP per capita, PPP (current international $)," 2018. [Online]. Available: https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?end=2016{&}locations=CN{&}start=1990{&}view=chart

[8] Danmarks Statistik, "Slagtninger og produktion af fjerkræ efter enhed, ANI6," mar 2018. [Online]. Available: http://www.statistikbanken.dk/ANI6

[9] Landbrug og Fødevarer, "Udenrigshandel 2012 - 2016," Tech. Rep., 2016. [Online]. Available: https://www.lf.dk/~/media/lf/tal-og-analyser/fodevareklyngens-udenrigshandel/2016/udenrigshandel-2012-2016--final-.pdf?la=da

[10] X. Pei, A. Tandon, A. Alldrick, L. Giorgi, W. Huang, and R. Yang, "The China melamine milk scandal and its implications for food safety regulation," *Food Policy*, vol. 36, no. 3, pp. 412–420, jun 2011.

[11] The Associated Press, "Nearly 53,000 Chinese children sick from milk," 2008. [Online]. Available: http://www.nbcnews.com/id/26827110/{#}. WroP{_}z-sYuU

[12] Reuters, "China milk banned in Asia, Africa, Europe Union," sep 2008. [Online]. Available: https://in.reuters.com/article/us-china-milk-recalls/factbox-china-milk-banned-in-asia-africa-europe-union-idINTRE48P3Z120080926

[13] Department of Agriculture, Food and the Marine, "Equine DNA and Mislabelling of Processed Beef Investigation," mar 2013. [Online]. Available: https://www.fsai.ie/uploadedFiles/Enforcement_and_Audit/Horse_Meat/Equine-DNA-DAFM-March-2013.pdf

[14] H. Ritchie and M. Roser, "Meat and Seafood Production & Consumption," 2018. [Online]. Available: https://ourworldindata.org/meat-and-seafood-production-consumption

[15] J. Lesschen, M. van den Berg, H. Westhoek, H. Witzke, and O. Oenema, "Greenhouse gas emission profiles of European livestock sectors," *Animal Feed Science and Technology*, vol. 166-167, no. 167, pp. 16–28, jun 2011.

[16] Marel Poultry, "The world of Poultry Processing," 2018. [Online]. Available: https://marel.com/files/pdf/world-of-stork-poultry-en.pdf

[17] European Parliament, "Corrigendum to Regulation (EC) No 854/2004," 2004. [Online]. Available: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32004R0854

[18] U. Löhren, "Overview on current practices of poultry slaughtering and poultry meat inspection," *EFSA Supporting Publications*, vol. 9, no. 6, pp. 1–58, jun 2012. [Online]. Available: http://doi.wiley.com/10.2903/sp.efsa.2012.EN-298

[19] European Food Safety Authority (EFSA), "Scientific Opinion on the public health hazards to be covered by inspection of meat (poultry)," *EFSA Journal*, vol. 10, no. 6, pp. 1–179, jun 2012. [Online]. Available: http://doi.wiley.com/10.2903/j.efsa.2012.2741

[20] G. ElMasry, N. Wang, C. Vigneault, J. Qiao, and A. ElSayed, "Early detection of apple bruises on different background colors using hyperspectral imaging," *LWT - Food Science and Technology*, vol. 41, no. 2, pp. 337–345, mar 2008.

[21] G. ElMasry, D.-W. Sun, and P. Allen, "Near-infrared hyperspectral imaging for predicting colour, pH and tenderness of fresh beef," *Journal of Food Engineering*, vol. 110, no. 1, pp. 127–140, may 2012.

[22] B. Dutchman, "Swing 20," 2018. [Online]. Available: https://cdn.bigdutchman.com/fileadmin/content/egg-poultry/products/en/egg-production-poultry-growing-weighing-systems-Big-Dutchman-en.pdf

[23] A. K. Mortensen, P. Lisouski, P. Ahrendt, A. Krogh Mortensen, P. Lisouski, and P. Ahrendt, "Weight prediction of broiler chickens using 3D computer vision," *Computers and Electronics in Agriculture*, vol. 123, pp. 319–326, 2016.

[24] Z. Du, M. K. Jeong, and S. G. Kong, "Band selection of hyperspectral images for automatic detection of poultry skin tumors," *IEEE Transactions on Automation Science and Engineering*, vol. 4, no. 3, pp. 332–339, 2007.

[25] J. Fletcher and S. Kong, "Principal component analysis for poultry tumor inspection using hyperspectral fluorescence imaging," *Proceedings of the International Joint Conference on Neural Networks, 2003.*, vol. 1, pp. 149–153, 2003.

[26] C. Xu, I. Kim, and M. S. Kim, "Poultry Skin Tumor Detection in Hyperspectral Reflectance Images by Combining Classifiers," pp. 1289–1296, 2007.

[27] I. Kim, M. S. Kim, Y. R. Chen, and S. G. Kong, "Detection of skin tumors on chicken carcasses using hyperspectral flourescence imaging," vol. 47, no. 5, pp. 1785–1792, 2004.

[28] S. Nakariyakul and D. P. Casasent, "Fast feature selection algorithm for poultry skin tumor detection in hyperspectral data," *Journal of Food Engineering*, vol. 94, no. 3, pp. 358–365, 2009.

[29] K. Chao, Y.-R. Chen, W. R. Hruschka, and F. B. Gwozdz, "On-line inspection of poultry carcasses by a dual-camera system," *Journal of Food Engineering*, vol. 51, no. 3, pp. 185–192, feb 2002.

[30] C. C. Yang, K. Chao, Y. R. Chen, and H. L. Early, "Systemically diseased chicken identification using multispectral images and region of interest analysis," *Computers and Electronics in Agriculture*, vol. 49, pp. 255–271, 2005.

[31] K. Chao, C. C. Yang, and M. S. Kim, "Spectral line-scan imaging system for high-speed non-destructive wholesomeness inspection of broilers," *Trends in Food Science and Technology*, vol. 21, no. 3, pp. 129–137, 2010.

[32] C.-C. Yang, K. Chao, M. S. Kim, D. E. Chan, H. L. Early, and M. Bell, "Machine vision system for on-line wholesomeness inspection of poultry carcasses." *Poultry science*, vol. 89, pp. 1252–1264, 2010.

[33] K. Chao, C. C. Yang, M. S. Kim, and D. E. Chan, "High throughput spectral imaging system for wholesomeness inspection of chicken," *Applied Engineering in Agriculture*, vol. 24, no. 4, pp. 475–485, 2008.

[34] Marel, "IRIS GDR," 2018. [Online]. Available: https://marel.com/poultry-processing/systems-and-equipment/broilers/grading-weighing-and-distribution/automatic-quality-grading/iris-gdr/432?prdct=1{&}parent=432

[35] Meyn, "Meyn Footpad Inspection System," 2018. [Online]. Available: https://www.meyn.com/media/1884/prle{_}footpad-inspection-system{_}rgb{_}03.pdf

[36] Foodmate, "INVISION CAMERA GRADING SYSTEM," 2018. [Online]. Available: https://www.foodmate.nl/uploads/pdf/pdfeng/InvisionCameraGradingSystem.pdf

[37] Baader, "ClassifEYE," 2018. [Online]. Available: https://www.baader.com/en/products/poultry_processing/distribution/classifeye/index.html

[38] B. Park, K. C. Lawrence, W. R. Windham, and M. P. Snead, "Real-time image processing for rapid contaminant detection on broiler carcasses," in *Proceedings of SPIE - The International Society for Optical Engineering*, Y.-R. Chen and S.-I. Tu, Eds., vol. 5587, no. November 2004, nov 2004, p. 101.

[39] K. J. C. Rui Kang, Kai Yang, Xia Xia Zhang, Wei Wu, "Development of Online Detection and Processing System for Contaminants on Chicken Carcass Surface," *Applied Engineering in Agriculture*, vol. 32, no. 1, pp. 133–139, feb 2016.

[40] H.-K. Sihvo, K. Immonen, and E. Puolanne, "Myodegeneration With Fibrosis and Regeneration in the Pectoralis Major Muscle of Broilers," *Veterinary Pathology*, vol. 51, no. 3, pp. 619–623, may 2014.

[41] J. P. Wold, E. Veiseth-Kent, V. Høst, and A. Løvland, "Rapid on-line detection and grading of wooden breast myopathy in chicken fillets by near-infrared spectroscopy," *PLoS ONE*, vol. 12, no. 3, p. e0173384, mar 2017.

[42] Y. Tao, J. Shao, K. Skeeles, and Y. R. Chen, "Detection of splenomegaly in poultry carcasses by UV and color imaging," *Transactions of the Asae*, vol. 43, no. 2, pp. 469–474, 2000.

[43] K. Chao, Y. R. Chen, H. Early, and B. Park, "Color image classification systems for poultry viscera inspection," *Applied Engineering in Agriculture*, vol. 15, no. November, pp. 363–369, 1999.

References

[44] K. Chao, Y. . R. Chen, W. R. Hruschka, and B. Park, "CHICKEN HEART DISEASE CHARACTERIZATION BY MULTI–SPECTRAL IMAGING," *Applied Engineering in Agriculture American Society of Agricultural Engineers*, vol. 99, no. 171, pp. 99–106, 2001.

[45] C. Hsieh, Y. R. Chen, B. P. Dey, and D. E. Chan, "SEPARATING SEPTICEMIC AND NORMAL CHICKEN LIVERS BY VISIBLE/NEAR–INFRARED SPECTROSCOPY AND BACK–PROPAGATION NEURAL NETWORKS," *Transactions of the ASAE*, vol. 45, no. 2, pp. 459–469, 2002.

[46] T. Amaral, I. Kyriazakis, S. J. McKenna, and T. Plotz, "Segmentation of organs in pig offal using auto-context," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, apr 2016, pp. 1324–1328.

[47] C. Claudi-Magnussen and H. Daugaard Larsen, "Automatiserede hjælpeværktøjer til kødkontrol på kyllingeslagterierne," Danish Meat Research Institute, Tech. Rep., 2011.

[48] C. E. Part, P. Edwards, S. Hajat, and L. M. Collins, "Prevalence rates of health and welfare conditions in broiler chickens change with weather in a temperate climate," *Royal Society Open Science*, vol. 3, no. 9, p. 160197, sep 2016.

[49] European Food Safety Authority, "Scientific Opinion on the public health hazards to be covered by inspection of meat (swine)," *EFSA Journal*, vol. 9, no. 10, p. 2351, oct 2011. [Online]. Available: http://doi.wiley.com/10.2903/j.efsa.2011.2351

[50] K. Chao, B. Park, Y. R. Chen, W. R. Hruschka, and F. W. Wheaton, "Design of a dual-camera system for poultry carcasses inspection," vol. 16, pp. 581–587, 09 2000.

[51] S. C. Yoon, B. Park, K. C. Lawrence, W. R. Windham, and G. W. Heitschmidt, "Line-scan hyperspectral imaging system for real-time inspection of poultry carcasses with fecal material and ingesta," *Computers and Electronics in Agriculture*, vol. 79, no. 2, pp. 159–168, 2011.

[52] J. Papon, A. Abramov, M. Schoeler, and F. Worgotter, "Voxel Cloud Connectivity Segmentation - Supervoxels for Point Clouds," in *2013 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, jun 2013, pp. 2027–2034.

[53] A. Blankenship, "Psychological Difficulties in Measuring Consumer Preference," *Journal of Marketing*, vol. 6, no. 4, p. 66, apr 1942. [Online]. Available: http://www.jstor.org/stable/1246085

[54] M. B. R. Mollah, M. A. Hasan, M. A. Salam, and M. A. Ali, "Digital image analysis to estimate the live weight of broiler," *Computers and Electronics in Agriculture*, vol. 72, no. 1, pp. 48–52, jun 2010.

[55] M. Kashiha, C. Bahr, S. Ott, C. P. Moons, T. A. Niewold, F. O. Ödberg, and D. Berckmans, "Automatic weight estimation of individual pigs using image analysis," *Computers and Electronics in Agriculture*, vol. 107, pp. 38–44, sep 2014.

[56] D. G. Bailey, K. A. Mercer, C. Plaw, R. Ball, and H. Barraclough, "HIGH SPEED WEIGHT ESTIMATION BY IMAGE ANALYSIS," *Proceedings of the 2004 New Zealand National Conference on Non Destructive Testing*, 2004.

Part II

Appendix

Paper A

Preanalysis for Health Inspection of Viscera using Computer Vision

Anders Jørgensen

1 Introduction

This report will focus on finding the best applicable light for automatic data collection. The approach is to extract features at multiple wavelengths and select the most important wavelengths based on the performance in a classification. The recorded viscera sets have been graded by veterinarians.

2 Organs in the viscera set

Figure A.1 shows a viscera set with the organs marked. The liver (marked with blue) generally has a dark brown/red color and a uniform surface with respect to both color and texture. It is almost split in two, both parts shaped roughly like a chicken breast fillet. It is fattest where it is attached to the other organs and gets thinner towards the edges. The heart (marked with red) has a greyish red color and is shaped like an ellipsoid but fatter in one end. 1/4 down there is a ring of fat and above that the heart has a dark purple tint. Blood veins often runs along the sides of the heart but the details can be difficult to see, as the heart is covered by a membrane called pericardium.

The lungs (marked with green) are placed next to the heart and usually have a clear red color. The brightness of the red color depends on the amount of blood left in the lungs and can therefore vary between the birds. Its surface is wrinkled and have some deep trenches shaped by the chicken's ribs. The Gizzard (marked with yellow) is placed under the liver and is shaped like a ball. It can be difficult to see the gizzard as it is often covered by the liver or by fat.
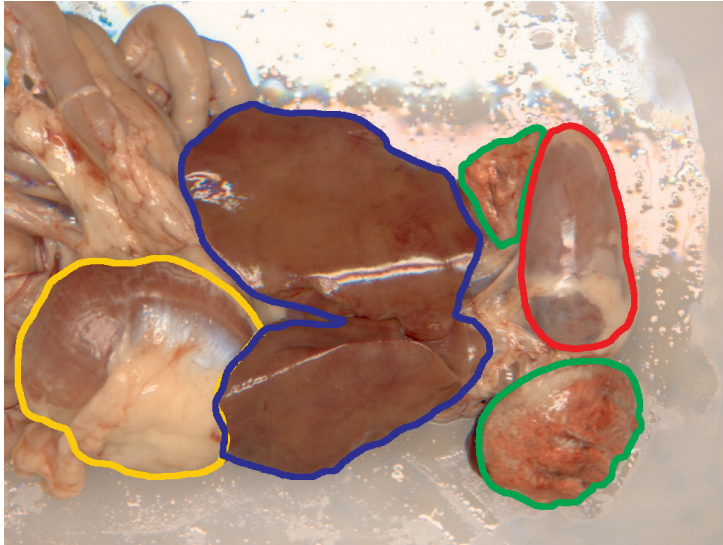
Fig. A.1: The large organs in the chickens. The liver is marked with blue, heart with red, lungs with green, and the gizzard with yellow.

3   Data set

The data set has been recorded at Danpo A/S in the period from September 24. to October 9th. 331 sets of viscera have been recorded. Each set have been recorded at 30 different wavelengths, from 430 nm to 720 nm with 10 nm intervals.

3.1   Image recording

The images are captured with a monochrome Basler Scout camera couple with a VariSpec Liquid Crystal Tuneable Filter. The filter has a bandwidth of 10 nm and a range from 400 nm to 720 nm. Due to very low light intensities at the short wavelengths, no images are captured below 430 nm. The resolution of the camera is 1024x768 pixels. A sketch of the setup can be in figure A.2a.

The filter is mounted in front of the lens and due to the physical size of the filter it requires a narrow field of view to avoid vignetting in the images. The distance between the camera and the viscera is therefore around one meter. This is however not a bad thing as the lamps get quite hot and could cause the surfaces on the viscera to dry out during the 30 seconds it takes to record a full sweep of images from 430 nm to 720 nm.
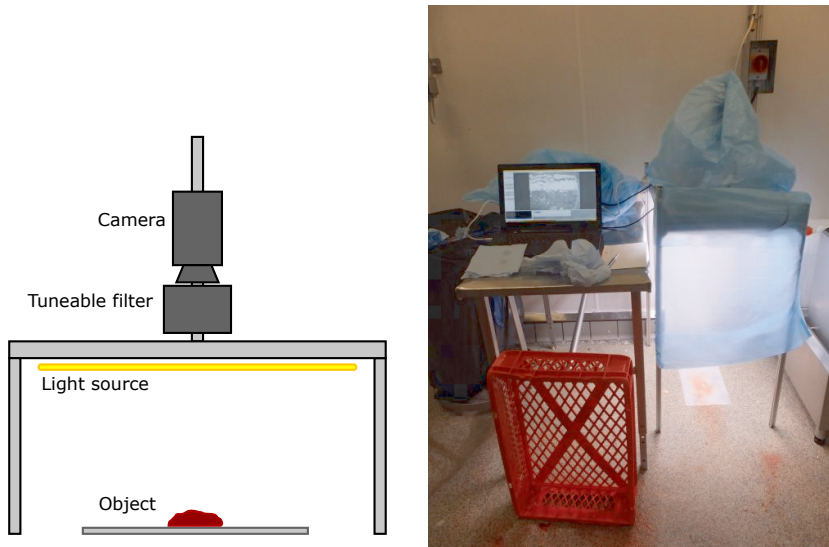
(a) Sketch of the spectral setup. The physical size of the tuneable filter limits the camera's field of view.

(b) The equipment set up at Danpo A/S. Most of it is wrapped in plastic to avoid water or blood sprays from the nearby processing equipment.

Fig. A.2: The spectral setup. Figure A.2a shows a sketch of the setup and figure A.2b shows the setup at Danpo A/S.

All images are recorded on 7 different days over a 14-day period. The room was cleaned daily which meant that the equipment had to be taken down at the end of each recording and a new light calibration made at each setup. The calibration was made to ensure a uniform light intensity across the entire spectrum. This was controlled by changing the camera's exposure time for each wavelength. Figure A.3 shows all exposure times from all calibrations. The camera is calibrated to the first gray patch (Neutral 8) on a X-Rite ColorChecker Classic [1]. In figure A.3 it can be seen that the calibration was fairly similar during the recording period. The high spikes in the light intensity is due to the light source being flourescent tubes.
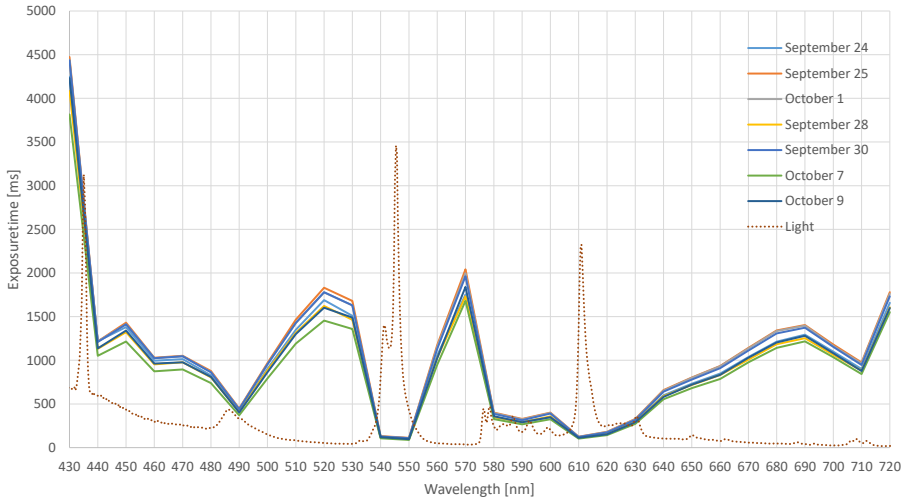
Fig. A.3: Comparison of exposure times across the recording days at Danpo A/S. The intensity of the lamps are depicted with the dotted line.

3.2 Images in the data set

The data set contains images of 331 sets of viscera. The viscera was categorized by veterinarians during the recording and divided into the five categories seen in table A.1. The data set contains all the sick viscera sets re-

Diagnose	Count
Healthy	75
Peritonitis	103
Necrotic hepatitis	83
Pericarditis	45
Liver changes	50

Table A.1: Share of diagnoses in the data set.
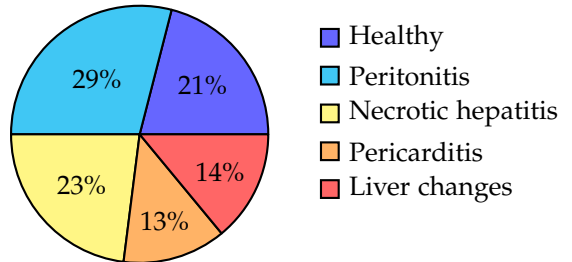
Table A.2: Share of diagnoses in the data set.

moved from the line during the recording hours. Some sets have more than one diagnosis so the sum of all counts in table A.1 is larger than 331.

An example of a sweep of a viscera set can be seen in figure A.4.

Fig. A.4: Example of an image sweep for a viscera set. Starting at 430*nm* in the top left corner to 720*nm* in the bottom right corner.

4 Viscera diagnoses

For all recorded viscera sets, only two organs have causes the viscera to be discarded, namely the liver and the heart. To simplify the analysis, we want the heart and the liver graded separately. The liver can be affected by necrotic hepatitis, peritonitis or liver changes. The heart can be affected be pericarditis or peritonitis. As there is no clear visible difference between pericarditis and peritonitis, when just looking at the heart, these are put in the same category. The difference is that peritonitis affects the entire peritoneum and not just the pericardium. The new grading terms are made clear in table A.3.

Diagnoses		
Viscera	**Liver**	**Heart**
Healthy	Healthy	Healthy
Peritonitis	White	Inflammation
Liver changes	Changes	-
Necrotic hepatitis	Inflammation	-
Pericarditis	-	Inflammation

Table A.3: Class names for entire viscera sets and for liver and heart individually.

4.1 Examples of each diagnose

No colour images were captured of the viscera. The colour images presented in this report is therefore generated from the 30 captured monochrome images. The purpose is to generate images that are pleasing for humans to look at, because they match that of the real world.

Each pixel in each image is multiplied by three factors for each wavelength that specifies how much this wavelength contributes to the red, green and blue colour channels. The factors are obtained from this website, `http://www.cvrl.org/cmfs.htm`, by the Colour and Vision Research laboratory at the Univesity of London.

Healthy viscera

Figure A.5 shows two examples of healthy viscera. These are taken from the conveyor after the inspection site and are therefore indirectly approved as healthy, as they were not taken of the conveyor at the inspection site. The colour of the liver can vary between the healthy sets, but each liver generally has a uniform colour over the entire surface.
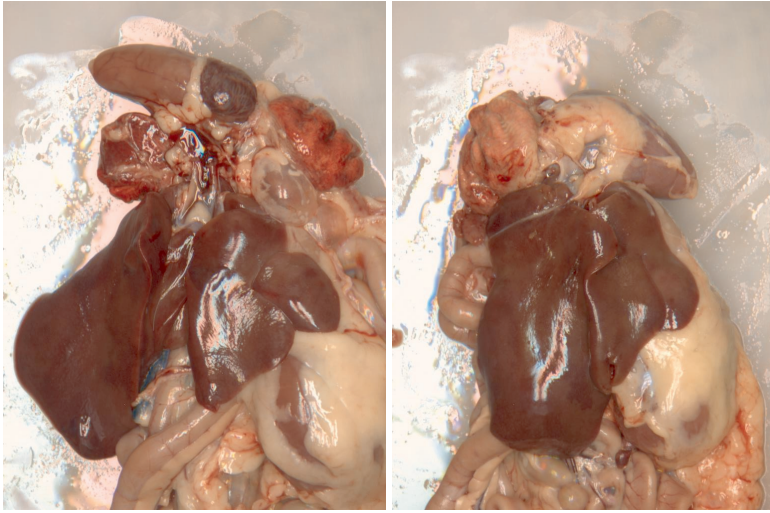
Fig. A.5: Examples of healthy viscera sets. See figure A.13 for more examples.

Necrotic hepatitis (Liver inflammation)

Liver inflammation can be recognized by the dots or specks on the liver. There is a large variation in colour, size and shape of the dots. In addition to the dots, the liver is sometimes discoloured "behind the dots", often green or gray. The dots can also group together and form discoloured and rough areas on the liver. See figure A.6 for examples.
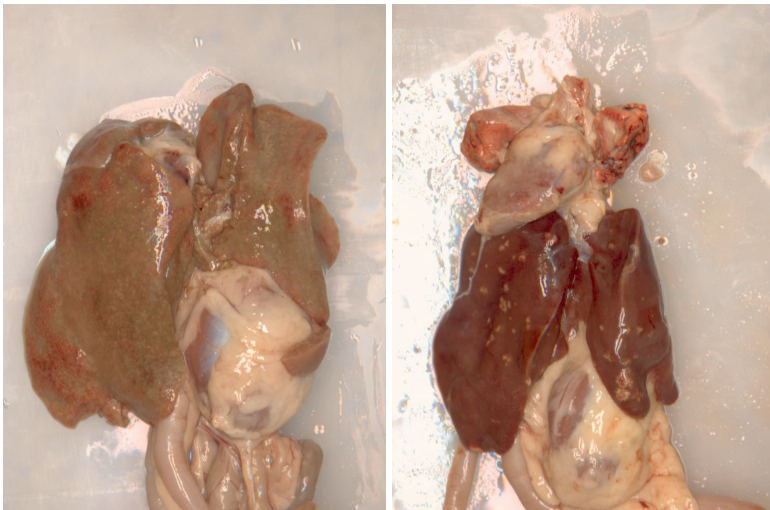
Fig. A.6: Examples of liver inflammation. See figure A.14 for more examples.

Liver changes

Liver changes are here used as a common term for changes in colour, shape or texture of the liver. It is often correlated with ascites which is detected and removed at the inspection before the evisceration. In this data set they are often seen as dents on the liver. The dents can be difficult to see in the images if not for the dark trenches formed between the dents. Often the liver will also be slightly discoloured. See examples in figure A.7.
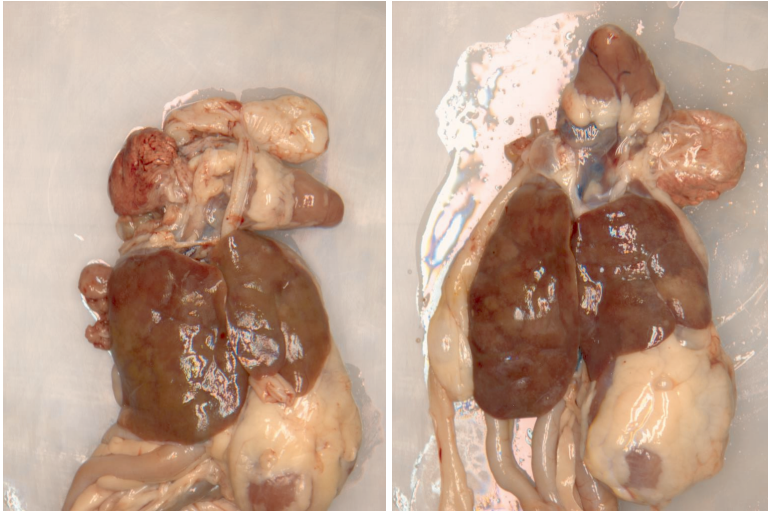
Fig. A.7: Examples of liver changes. See figure A.15 for more examples.

Peritonitis (white liver)

Peritonitis is seen when the peritoneum gets attached to the liver. In this report it is also called white liver, due to the white colour of the peritoneum. The colour can however vary from almost transparent, where it only gives the liver a white tint, to yellow where it looks like melted cheese. The infectious state can also cause the liver to get attached to other organs, for instance the gizzard. It can be very difficult to see this in the images as there are no visible clues and the liver often lies on top of the gizzard anyway. Peritonitis often affects both the heart and the liver at the same time. If the heart is affected this is categorized as heart inflammation, as described in table A.3. See examples of white liver in figure A.8.
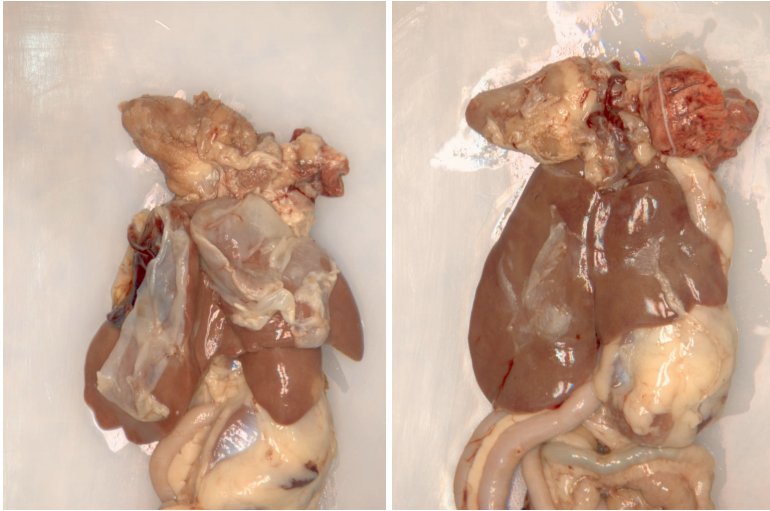
Fig. A.8: Examples of white liver. See figure A.16 for more examples.

Pericarditis/Peritonitis (heart inflammation)

This category describes inflammation of the heart in general. Visually there can be a large variation in the inflammations, see figure A.9. The membrane around the heart most often becomes white/yellow and opaque, but it can also become hard and show sediments of pus. In other cases the membrane gets filled with a yellowish or dark fluid.
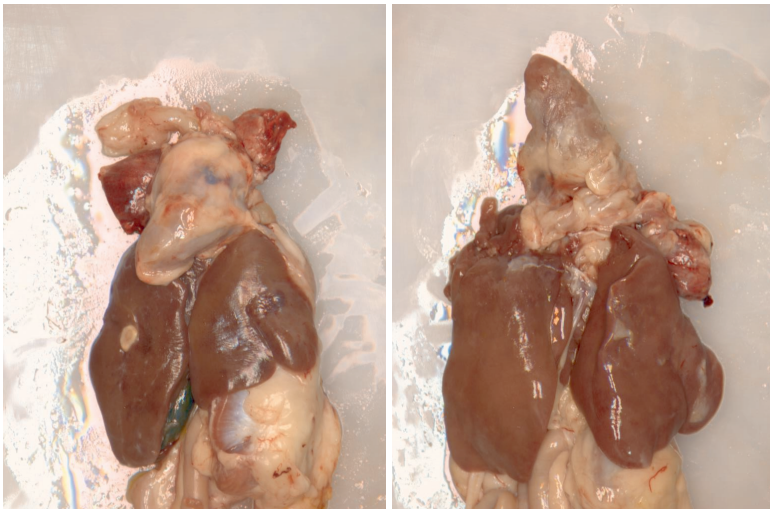
Fig. A.9: Examples of heart inflammation. See figure A.17 for more examples.

5 Spectral analysis

To investigate which wavelengths are important to classify the viscera sets, clear samples have been cropped from the images in the data set. By clear samples means regions that only belong to one of the classes in table A.3. These patches have a size of 25x25 pixels. This is small enough to easily pick areas that only represent one class and large enough to depict the variation in pixels that is present in e.g. necrotic hepatitis. The small size also means that multiple patches can be extracted from the same organ. Due to the size difference between liver and heart, more samples can be extracted from the liver than the heart. Figure A.10 shows an example of an extracted patch for each of the heart classes and figure A.11 one for each of the liver classes.

The idea behind extracting small patches is to avoid contaminating the classes with data that actually belongs to another class. Looking at the image to the right in figure A.6, it is clear that even though this liver has necrotic hepatitis it still contains areas that are not affected or at least looks healthy. If you extract all pixels from this liver many of them will belong to the healthy class even though the liver is classified as necrotic hepatitis. The patches furthermore ease the data processing as all samples now have the same size, whereas the organs vary greatly in size and shape.

(a) The 30 channels of a healthy heart patch.

(b) The 30 channels of an inflamed heart patch.

Fig. A.10: Examples of two 25x25 heart patches. Starting from 430 nm top left to 720 nm bottom right.
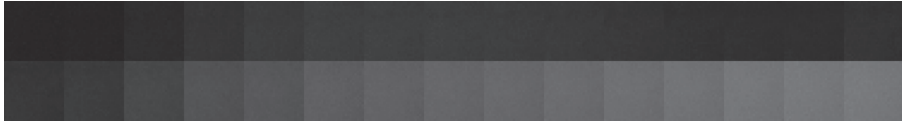
(a) The 30 channels of a healthy liver patch.
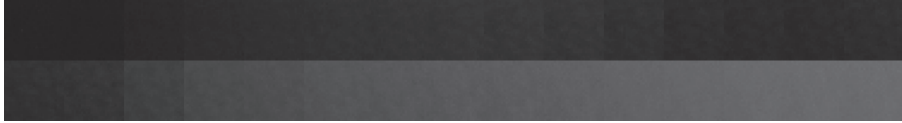
(b) The 30 channels of a changed liver patch.

(c) The 30 channels of a white liver patch.

(d) The 30 channels of a inflammed liver patch.

Fig. A.11: Examples of four 25x25 liver patches. Starting from 430 nm top left to 720 nm bottom right.

5.1 Extracted patches

A total of 1594 patches have been extracted, divided as described in table A.4.

Diagnosis	Count	Abbr.
Healthy liver	395	LG
White liver	451	LH
Liver inflammation	322	LB
Liver changes	122	LF
Health heart	152	HG
Heart inflammation	152	HB

Table A.4: Extracted patches per diagnosis
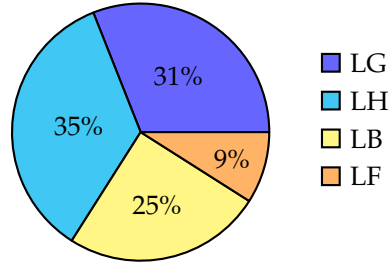
Table A.5: Share of patches extracted for the liver diagnoses.

There is generally more liver patches than heart patches as the liver is larger than the heart. The one exception is liver changes as this class were less represented in the original data set than the other liver classes and because liver changes often are found together with other diseases which makes it difficult to extract clear patches.

5.2 Feature extraction algorithms

The patches contain both intensity and textural information. To utilize this it was chosen to extract some simple features from each patch. The five algorithms for extracting the features are described in table A.6. The algorithms are applied individually to all 30 wavelengths per patch which means a total of 150 features per patch.

Abbreviation	Description
Int	Average intensity for all pixels in the image.
Ent	The entropy calculated from all pixels in the image. A measure of the information in the patch.
Var	The variance of all pixels in the image. A measure of the variance in the image.
Sob	A 3x3 Sobel kernel is applied to the image. The average intensity of the absolute pixel values is then found. A high value should indicate many edges in the image.
SobVar	A 3x3 Sobel kernel is applied to the image. Afterwards the variance is calculated. A measure for the variation between edge/not edge in the image.

Table A.6: The five feature extraction algorithms. Image here refers to the intensity image for one wavelength in a patch.

6   Graphical representation of features

This section has been removed to not disclose confidential information.

7   Selection of wavelengths and features

All wavelengths mentioned in this section have been replaced by a randomly generated set of letters to not disclose confidential information.

To make it more practical to record images at a high frame rate it is necessary to reduce the number of wavelengths. This is done by removing wavelengths that do not contribute positively to the classification.

The classifier is trained with the Multi-Class AdaBoost algorithm [2]. This method calculates the information gain for each split and sums the gains for the entire model after training, giving an overall performance score for each feature. For representation purposes are the highest score set to 1 and the following feature scores are given relative to the first.

This can be used to remove features with a low score while monitoring the overall classification results. A feature can have a low score for two reasons, if it contains no useful information or if it contains information already explained by another feature.

Table A.7 shows the 10 most important features for the liver classifier.

Top 10	Feature and wavelength	Score
1	*Var CC*	1
2	*Int NF*	0.6749
3	*SobVar HK*	0.5219
4	*Sob AU*	0.5045
5	*SobVar EM*	0.3321
6	*Ent VV*	0.2644
7	*SobVar VV*	0.2364
8	*Int VV*	0.2262
9	*Sob CC*	0.2020
10	*SobVar CC*	0.2003

Table A.7: Top 10 features for the liver classifier and their relative scores. The features are named by the extaction algorithm and the wavelength.

The classifier doesn't know about the correlation between neighboring wavelengths and will always choose the feature that splits the most observations. However, as indicated in section 6 will neighboring wavelengths be highly correlated and can therefore replace each other without a noticeable loss of classification accuracy. This means that the features *Var CC*, *Int NF*, *Sob AU*, *Sob CC*, and *SobVar CC*, which are all among the top 10 best features, can be extracted from a single wavelength, e.g. *NF*.

The classifier achieves an accuracy of 95.3% for the four liver classes using all 150 features. Table A.8 shows the accuracy when the classifier is trained with 10 different collections of features. All reported results are the average of five five-fold stratified cross validation runs.

Selected features	Accuracy
All features at all wavelengths	95.3 %
Top 10	95.4 %
Top 3	93.0 %
Top 1	67.2 %
All features at *NF* nm	84.5 %
All features at *HP* nm	85.6 %
All features at *KB* nm	76.0 %
All features at *NF* nm and *HP* nm	90.7 %
All features at *HP* nm and *KB* nm	93.0 %
All features at *NF* nm and *KB* nm	95.1 %
All features at *NF* nm, *HP* nm and *KB* nm	95.5 %

Table A.8: Accuracy for the liver classifier trained with different sets of features. Accuracy describes the total True Positive Rate for all classes.

The 10 best features achieves an accuracy similar to all features at all wavelengths, yet from table A.7 it can be seen that these features comes from just 6 wavelengths. It makes sense to reduces the number of wavelengths even further, as these are expensive in both hardware requirements and recording time, whereas the features are relatively cheap to extract once you have the image in memory. From table A.8 can it be concluded that the wavelengths *NF* nm and *KB* nm can provide a similar accuracy as all features at all wavelengths, 95.1 % vs 95.3 %. By added a third wavelength, *HP*, giving a total of 15 features, the classifier achieves an even higher accuracy of 95.5 %.

The wavelengths are selected based on the assumption that they can be recorded simultaneously in one frame by an RGB camera, as visualized in figure A.12. They all give a high response in each colour channel with a minimum of crosstalk.
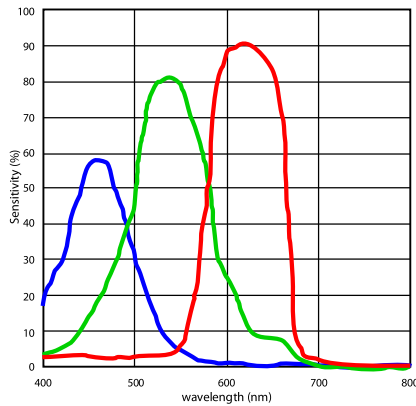
Fig. A.12: Spectral response curves for a typical RGB camera with an infrared cutoff filter. The wavelengths *NF* nm, *HP* nm and *KB* nm give a high response in each of the three colour channels with a small amount of crosstalk.

8   Conclusion

This study shows, that when looking at liver diseases, the number of wavelengths can be drastically reduced without compromising classification accuracy. The results in table A.8 shows that all features at the wavelengths *NF* nm and *KB* nm can achieve the same level of accuracy as all wavelengths with all features. Adding *HP* increases the accuracy even further and results in an RGB image with a response in all three colour channels. This is valuable if the images are to be inspected manually either by veterinarins or developers.

The classification in this study have been done on 25x25 pixel samples, to keep the classes pure and to focus on the task of finding useable wavelengths. Liver changes are difficult to represent with a small patch as the structural

changes affect the entire liver on a larger scale than liver inflammation. A more holistic approach for detecting liver changes must be investigated for an automatic classification system.

The graphs in section 6 show that the spectral response of the diseases are highly correlated from one wavelength to the next. This is an advantage when it comes to designing a specialized recording system, as it leaves some leeway when selecting camera and LEDs. This also means that a classification system is less likely to drop significantly in performance if the LED's peak wavelength shift due to aging.

9 Example images

Healthy viscera
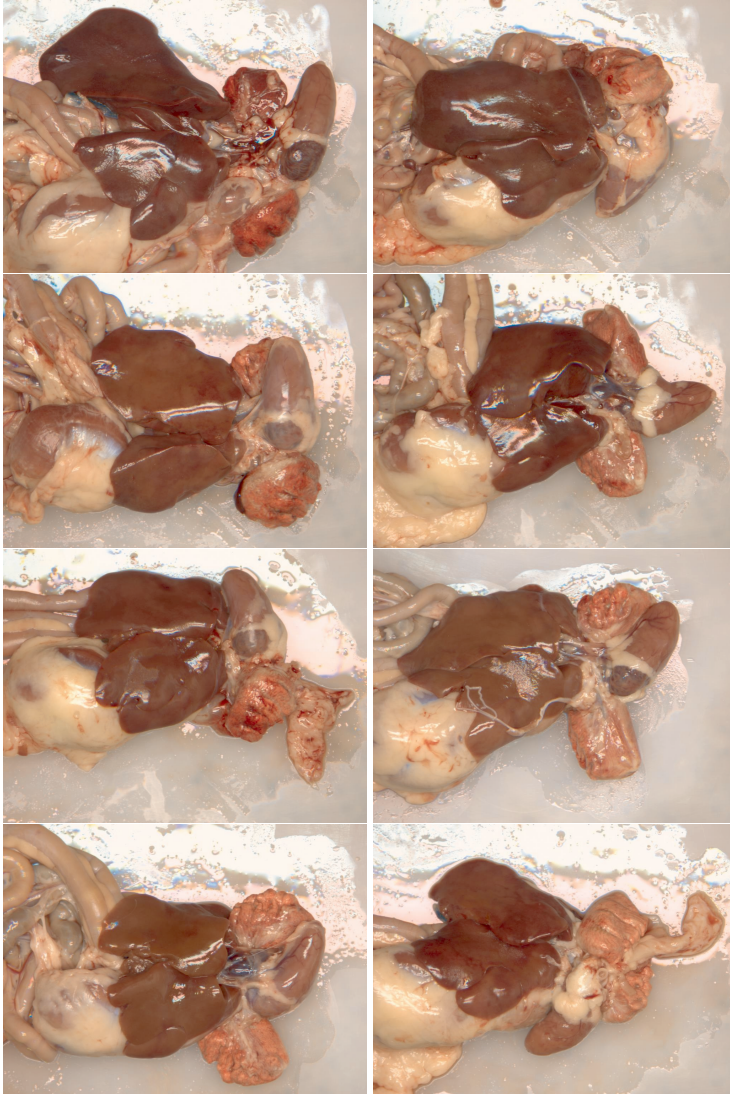
Fig. A.13: Examples of of healthy viscera.

Liver inflammation
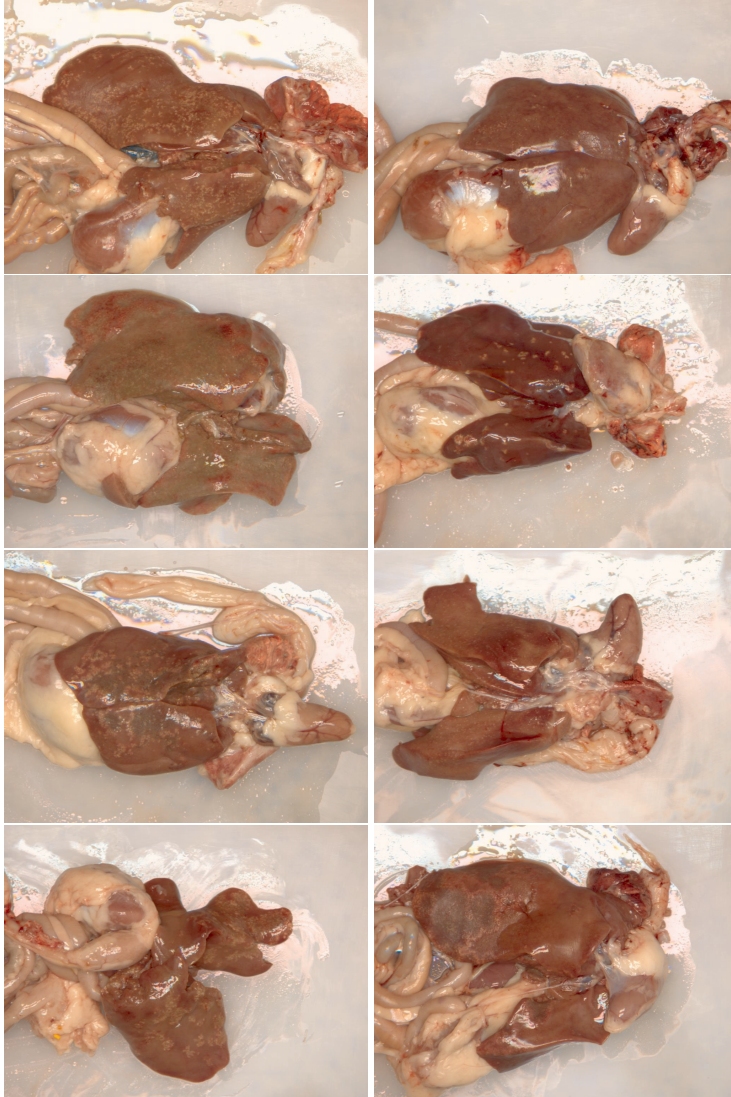
Fig. A.14: Examples of liver inflammation.
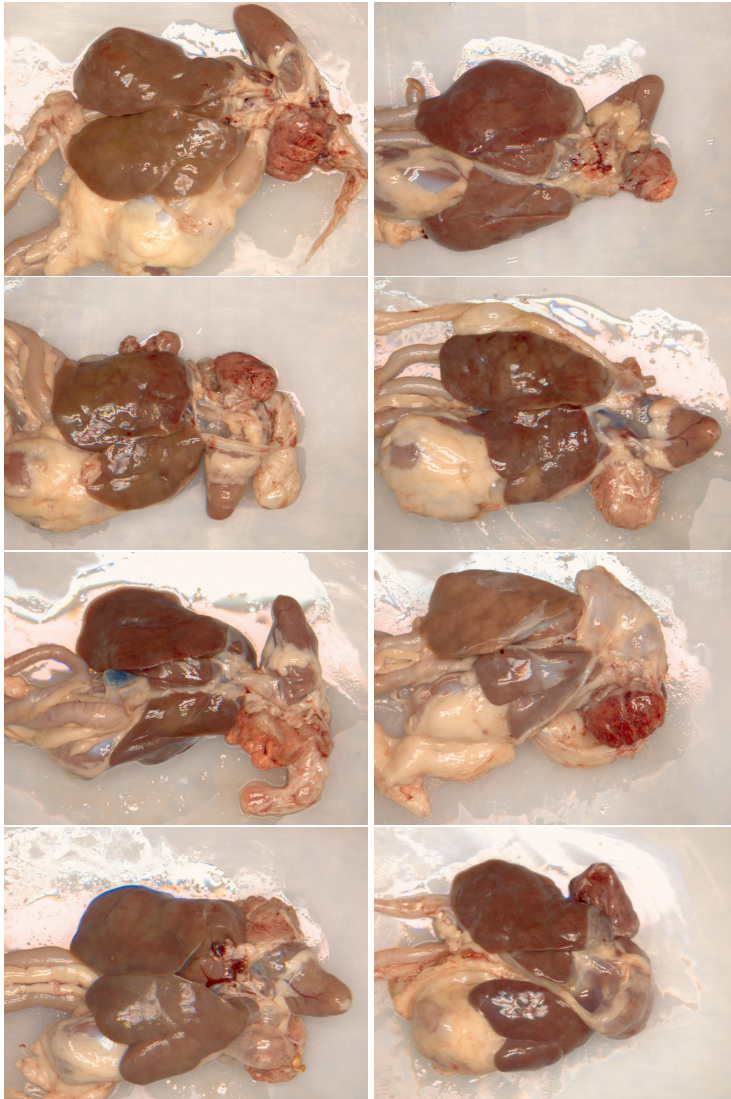
Liver changes

Fig. A.15: Examples of liver changes.

White liver (Peritonitis)
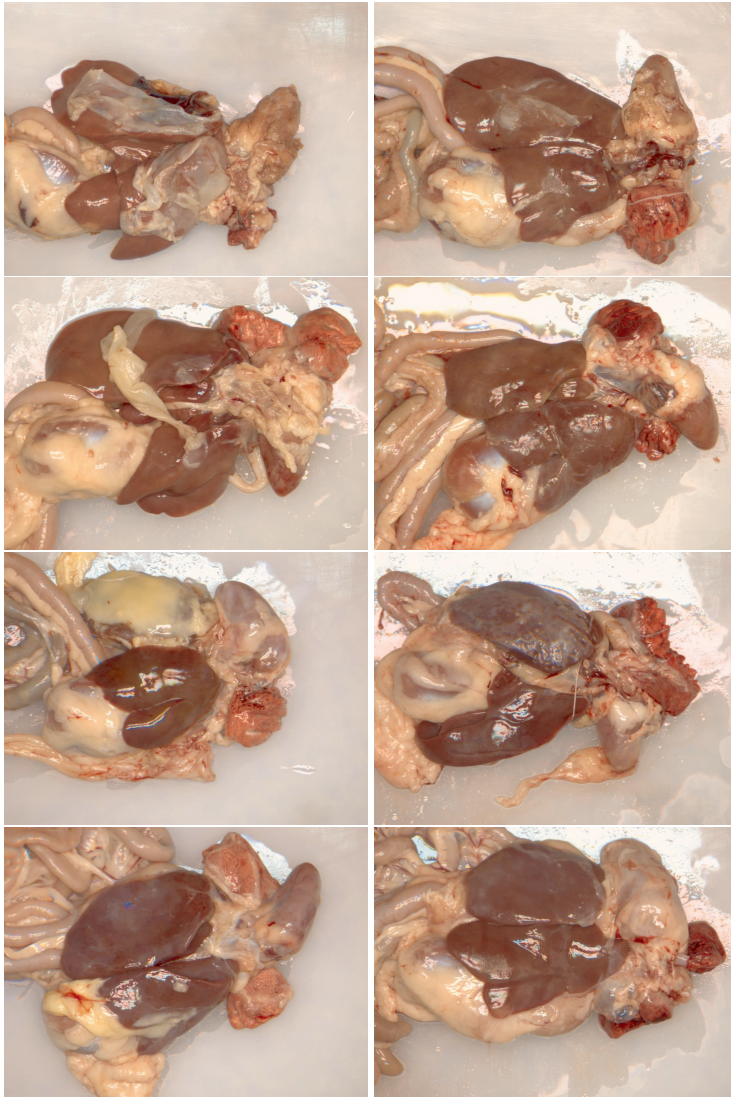
Fig. A.16: Examples of white liver.

Heart inflammation
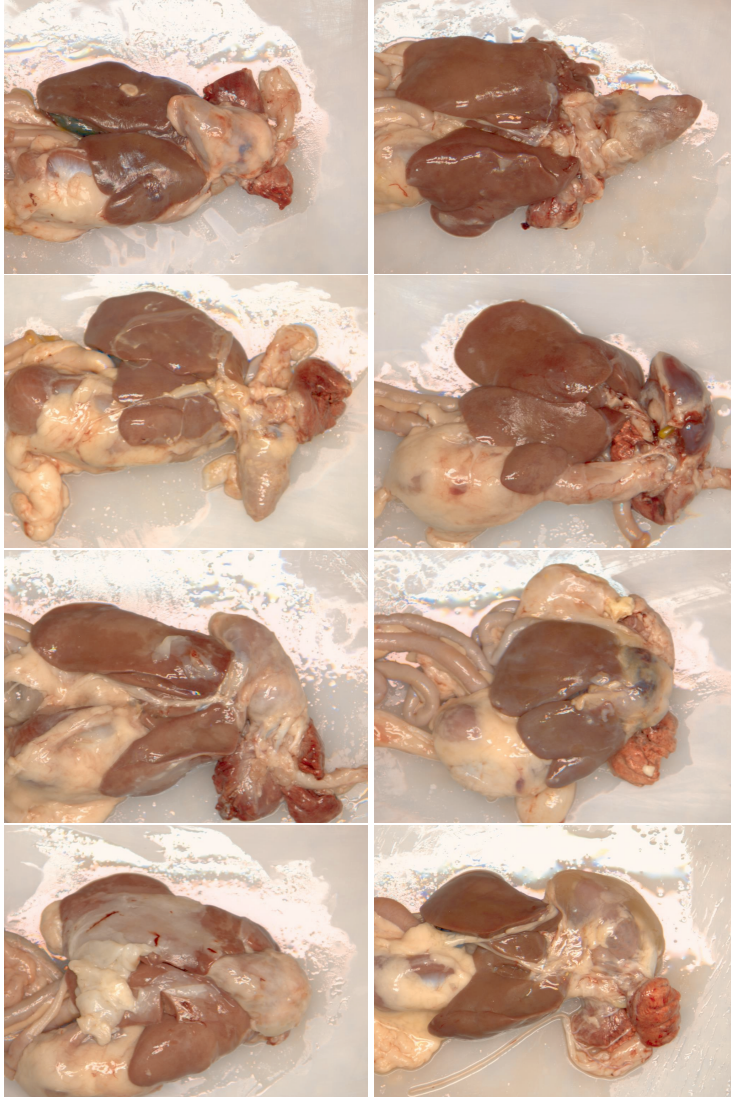
Fig. A.17: Examples of heart inflammation.

References

[1] X-Rite, "Colorchecker classic," http://xritephoto.com/colorchecker-classic, December 2015.

[2] T. Hastie, S. Rosset, J. Zhu, and H. Zou, "Multi-class AdaBoost," *Statistics and Its Interface*, vol. 2, no. 3, pp. 349–360, 2009. [Online]. Available: https://web.stanford.edu/~hastie/Papers/SII-2-3-A8-Zhu.pdf

Paper B

Detecting Gallbladders in Chicken Livers using
Spectral Analysis

Anders Jørgensen, Eigil Mølvig Jensen, and Thomas B.
Moeslund

Abstract

This paper presents a method for detecting gallbladders attached to chicken livers using spectral imaging. Gallbladders can contaminate good livers, making them unfit for human consumption. A data set consisting of chicken livers with and without gallbladders, has been captured using 33 wavelengths within the visible spectrum. This work shows how to reduce the high number of wavelengths while maintaining a high accuracy. A classification tree has be trained to evaluate if a gallbladder is present and whether it is suitable for automatic removal, which could increase profits for the processing plants. As a preliminary study this shows good results with a classification accuracy of 91.7%.

1 Introduction

The consumption of chicken meat is increasing in almost every corner of the world. From 2000 to 2011 there has been a 31 % increase in the top chicken producing countries, according to the International Poultry Council [1]. To maintain a high quality while increasing the production speed, poultry processing plants relies on computer vision for a larger numbers of inspection tasks. Already existing works can detect skin tumours [2, 3] and faecal matter [4, 5] while others sort chickens by wholesome/unwholesome [6, 7]. Similar methods also exist in commercial systems by Linco and Meyn [8, 9].

But quality inspection of the chicken's organs have received little attention in research. After evisceration, the liver, heart and gizzard are sorted out from the remaining guts. These can be used for sausages, pâté or sold as is. The quality decides whether they are sold for human consumption or are to be used in fodder for animals. And this has a high affect on the price of these organs. Today this quality assessment relies heavily on manual inspection.

Due to the chicken's anatomy, the gallbladder is sometimes by mistake extracted with the liver during evisceration. This means that the liver is unfit for human consumption and thereby only worth 25 % that of a clean liver. It is therefore important to separate livers with and without gallbladders correctly. This paper presents a preliminary study proposing a method for detecting gallbladders among livers using spectral images within the visible spectrum. It further investigates whether detected gallbladders are fit for automatic removal.

2 Related work

Hyperspectral imaging in relation to food quality and safety receives a growing attention these years, as described by Huang et. al in [10]. Both in superfi-

cial inspection of beef-marbling [11], estimating bacteria on the surface [12] to internal quality like tenderness. Using the NIR spectrum, Elmasry et. al [13] are capable of estimating the pH-value and tenderness of fresh beef. Work by Kamruzzaman et. al [14] show that it was also possible to estimate water holding capacity for lamb meat using images captured within the NIR spectrum.

Little research has been done regarding chicken viscera. Tao et. al [15] use ultra violet light to detect splenomegaly in turkey carcasses. The spleen appears almost completely dark in images captured at a central wavelength at 365 nm, making it recognizable from the liver.

Spectroscopy has also been used to detect septicemia in chicken livers [16]. Using a neural network they are able to detect 94 % of the septox livers in their sample.

3 Image acquisition

For this study, a data set consisting of 60 image sets was captured off-line at a chicken processing plant. 30 image sets contains just a liver, and 30 image sets contains a liver with a gallbladder attached. The images have been captured using a monochrome Basler Scout camera combined with a VariSpec Liquid Crystal Tunable Filter. The filter has a bandwidth of $10nm$ and a range from $400nm$ to $720nm$, both inclusive. It was chosen to capture images in steps of $10nm$, which results in 33 grey scale images in each image set. The image resolution is $1024x768$. Examples from the data set can be seen in figure B.1.
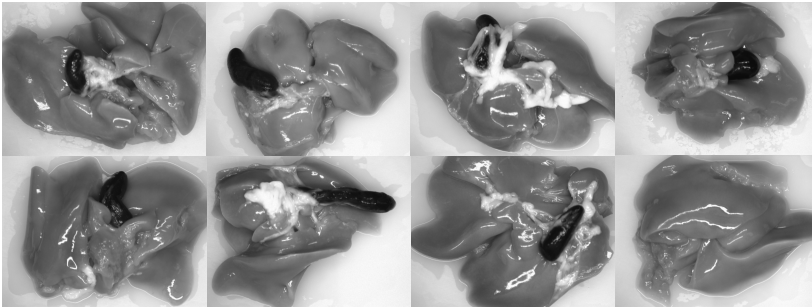
Fig. B.1: Examples of liver images, captured at $650nm$, from the data set. The black organ is the gallbladder and the white matter is fat.

Before capturing the data set, the lighting was measured using an Ocean Optics spectrometer. This was done at the white region of an X-rite Colour Checker Classic [17]. The result can be seen in figure B.2.
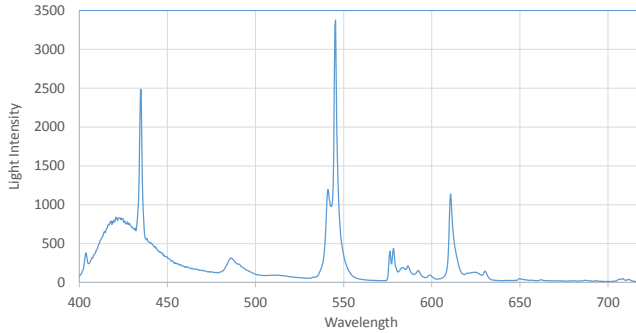
Fig. B.2: The spectral response of a white surface under the image capture setup.

To ensure an equal brightness level across the spectrum, the camera's exposure time must be adjusted for each wavelength. High light intensity means a low exposure time. This was done through a calibration routine where the average pixel value for the white region of the colour checker was equalised across all wavelengths.

4 Methods

4.1 Dimensionality reduction

The 33 grey scale images are most likely highly correlated, hence many of them can be removed without loss of information. For an on-line system, capturing 33 images at production speeds near 3 birds per second, would also be impractical and expensive. A better solution would be to select the important wavelengths and capture images at those.

For this, Hierarchical Dimensionality Reduction has been used, as this method finds the most uncorrelated subset of already existing features [18]. Pixels extracted from annotated images, containing the three classes *"liver"*, *"fat"* and *"gall"*, were concatenated in a $n \times 33$ matrix, where 33 is the number of wavelengths. n is the total of 85,352 gall pixels, 321,193 liver pixels and 47,515 fat pixels.

First a classification tree was trained using 10 fold cross validation to find the accuracy when using all wavelengths. The tree was set to contain a maximum of 100 nodes. Then, using the correlation matrix, the two most correlated wavelengths are selected. The Shannon entropy is calculated for both wavelengths and the one with the lowest entropy is then removed from the data set. The decision tree is then trained again. These steps are repeated until the accuracy of the classification tree starts to fall. The results can be seen in figure B.3.
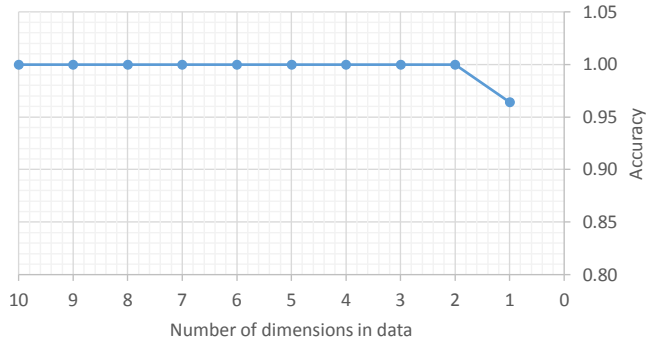
Fig. B.3: The accuracy of the decision tree as the dimensions are reduced. From 33 to 11 wavelengths are not depicted in the graph.

The accuracy is ≈ 0.999 until one wavelength remains, where it drops to 0.964. It is therefore chosen to use two wavelengths, to maintain a high accuracy. The two remaining wavelengths before the accuracy drops are $600nm$ and $720nm$. The resulting intensity space can be seen in figure B.4. This shows that the three classes can be separated using these two wavelengths.
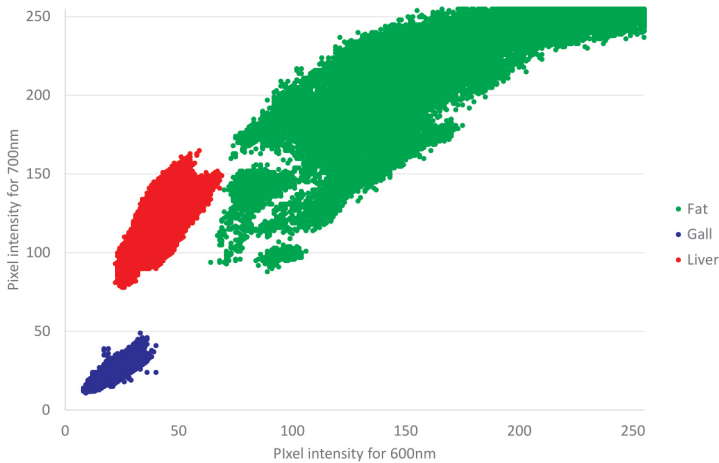
Fig. B.4: The intensity space using images captured at wavelength $600nm$ and $720nm$.

4.2 Segmenting the images

Using the decision tree trained during the dimensionality reduction, segmenting the gallbladder is straight forward. Every image is smoothed with a $5x5$ median filter to removed pixel outliers and then segmented by running each pixel through the tree where it is classified as either *"liver"*, *"fat"* or

"gall". Examples of segmented images can be seen in figure B.5. The white background is being classified as "fat".
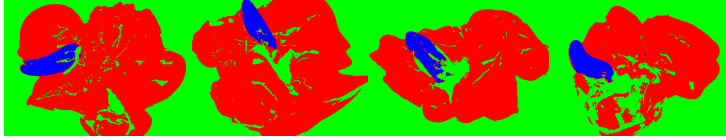
Fig. B.5: 4 examples of segmented images. Blue is gall, red is liver, and green is fat and background.

4.3 Evaluate the images

The data set should be classified into three groups. Images without a gallbladder("No Gall"), images with a gallbladder that can be removed("Good Gall") and images were the gallbladder can not be removed("Bad Gall"). In this study a gallbladder is considered removable if the entire contour is visible. A robot will in these situations be able to remove the gallbladder without puncturing it or damaging the liver.

Two features are extracted for this purpose. The first feature is the area of the gall in pixels. Blobs are created from the blue pixels in the segmented images. The area of the blobs is then summed, given the total area of the gallbladder in pixels. Holes inside the blobs are discarded, as these are either fat or reflections. Fat is not a problem for removal, as long as it is not connected to the liver.

The second feature is the ratio of the gall area divided by the area of the convex hull of all the gall blobs. In some segmented images the gall pixels are split into smaller blobs, often due to strings of fat on top of the gallbladder. All blobs in an image are therefore joined with the convex hull, to estimate the total area of the gallbladder. This can be seen in figure B.6.
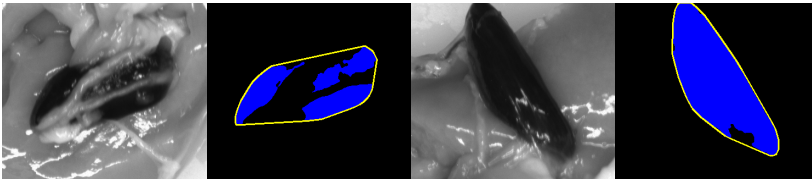
Fig. B.6: Finding the gallbladder in the images. The blue area describes the gall and the yellow line shows the convex hull around the gall areas.

5 Results

A classification tree has been trained using 10 fold cross validation, this time with a maximum of 4 nodes. The resulting confusion matrix can be seen in figure B.7. The overall accuracy of the system is 91.7%. 89.5% of "Bad Gall" were classified correctly, 72.7% of "Good Gall" were correct, and 100% "No Gall" were correctly classified.

		Predicted Class				TPR/FNR
		Bad Gall	Good Gall	No Gall		
True Class	Bad Gall	17 / 89.5%	2 / 10.5%	0 / 0.0%		89.5% / 10.5%
	Good Gall	3 / 27.3%	8 / 72.7%	0 / 0.0%		72.7% / 27.3%
	No Gall	0 / 0.0%	0 / 0.0%	30 / 100.0%		100.0% / 0.0%

Fig. B.7: Confusion matrix showing the accuracy of the trained classifier.

6 Discussion

Of the small data set zero "No Gall" images were misclassified. So clean livers pass though the system as they should. Three "Good Gall" images were misclassified as "Bad Gall". This would be expensive for a company as good livers are being sold as animal feed. Two "Bad Gall" images are predicted as "Good Gall". This error could also prove expensive if the gallbladder is punctured by the removal robot and ends up contaminating a whole batch of good livers.

(a) (b)
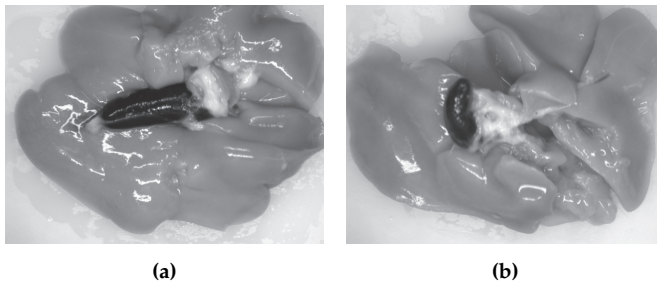
Fig. B.8: (**a**) This gallbladder is misclassified as "Good Gall". (**b**) This gallbladder is mis- classified as "Bad Gall"

Two error examples can be seen in figure B.8. The gallbladder shown in figure B.8(**a**) is wrongly classified as "Good Gall". The gall is almost com-

pletely free, but the "top" is covered with some liver and fat. The gallbladder shown in figure B.8(**b**) is simply much smaller than the average and therefore misclassified as "Bad Gall".

The objects were segmented using images captured at $600nm$ and $720nm$. These are both in the upper end of the visible spectrum, especially $720nm$ which also is the maximum of the used VariSpec filter. It would be interesting to extend the data set to include images captured at larger wavelengths, to investigate the information hidden outside the visible spectrum.

Overall, this preliminary study has shown that it is possible to detect the gallbladder on livers and, to some extend, tell whether they are removable or not. The gallbladder is easily separated from both the liver and fat using just one wavelength, yet two wavelengths are required for separating all three classes. Having two wavelengths also reduces the dependency of lighting conditions, especially if the segmentation is moved to colour space instead of intensity space. In general, the results looks promising but it should be tested on a larger data set before drawing conclusions.

References

[1] I. P. Council, "Chicken meat consumption for top producing countries," http://www.internationalpoultrycouncil.org/documents/ TopChickenConsumption.pdf, 2011.

[2] S. Nakariyakul and D. P. Casasent, "Fast feature selection algorithm for poultry skin tumor detection in hyperspectral data," *Journal of Food Engineering*, vol. 94, no. 3-4, pp. 358–365, 2009.

[3] Z. Du, M. K. Jeong, and S. G. Kong, "Band selection of hyperspectral images for automatic detection of poultry skin tumors," *IEEE Transactions on Automation Science and Engineering*, vol. 4, no. 3, pp. 332–339, 2007.

[4] B. Park, S. C. Yoon, W. R. Windham, K. C. Lawrence, M. S. Kim, and K. Chao, "Line-scan hyperspectral imaging for real-time in-line poultry fecal detection," *Sensing and Instrumentation for Food Quality and Safety*, vol. 5, pp. 25–32, 2011.

[5] S. C. Yoon, B. Park, K. C. Lawrence, W. R. Windham, and G. W. Heitschmidt, "Line-scan hyperspectral imaging system for real-time inspection of poultry carcasses with fecal material and ingesta," *Computers and Electronics in Agriculture*, vol. 79, no. 2, pp. 159–168, 2011.

[6] K. Chao, C. C. Yang, and M. S. Kim, "Line-scan spectral imaging system for online poultry carcass inspection," *Journal of Food Process Engineering*, vol. 34, no. 2011, pp. 125–143, 2011.

[7] C.-C. Yang, K. Chao, M. S. Kim, D. E. Chan, H. L. Early, and M. Bell, "Machine vision system for on-line wholesomeness inspection of poultry carcasses." *Poultry science*, vol. 89, pp. 1252–1264, 2010.

[8] Linco, "Classifeye," http://www.lincofood.com/en/products/poultry_processing/transfer_measurement/classifeye/index.html, 2015.

[9] Meyn, "Quality grading system," https://www.meyn.com/en-GB/69/quality-grading-system.html, 2015.

[10] H. Huang, L. Liu, and M. O. Ngadi, "Recent developments in hyperspectral imaging for assessment of food quality and safety." *Sensors (Basel, Switzerland)*, vol. 14, no. 4, pp. 7248–76, 2014.

[11] Y. Li, J. Shan, Y. Peng, and X. Gao, "Nondestructive assessment of beef-marbling grade using hyperspectral imaging technology," *ICAE 2011 Proceedings: 2011 International Conference on New Technology of Agricultural Engineering*, pp. 779–783, 2011.

[12] E. Z. Panagou, O. Papadopoulou, J. M. Carstensen, and G. J. E. Nychas, "Potential of multispectral imaging technology for rapid and non-destructive determination of the microbiological quality of beef filets during aerobic storage," *International Journal of Food Microbiology*, vol. 174, pp. 1–11, 2014.

[13] G. Elmasry, D. W. Sun, and P. Allen, "Near-infrared hyperspectral imaging for predicting colour, pH and tenderness of fresh beef," *Journal of Food Engineering*, vol. 110, no. 1, pp. 127–140, 2012.

[14] M. Kamruzzaman, G. ElMasry, D. W. Sun, and P. Allen, "Prediction of some quality attributes of lamb meat using near-infrared hyperspectral imaging and multivariate analysis," *Analytica Chimica Acta*, vol. 714, pp. 57–67, 2012.

[15] Y. Tao, J. Shao, K. Skeeles, and Y. R. Chen, "Detection of splenomegaly in poultry carcasses by UV and color imaging," *Transactions of the Asae*, vol. 43, no. 2, pp. 469–474, 2000.

[16] B. P. Dey, Y. R. Chen, C. Hsieh, and D. E. Chan, "Detection of septicemia in chicken livers by spectroscopy,," *Poultry Science*, vol. 82, no. 2, pp. 199–206, Feb. 2003.

[17] X. rite Photo, "Colorchecker classic," http://xritephoto.com/ph_product_overview.aspx?ID=1192&action=overview, 2015.

[18] D. G. S. Richard O. Duda, Peter E. Hart, *Pattern Classification*. Wiley-Interscience, 2001.

Paper C

Organ Segmentation in Poultry Viscera Using RGB-D

Mark P. Philipsen, Jacob V. Dueholm, Anders Jørgensen, Sergio Escalera, and Thomas B. Moeslund

Abstract

We present a pattern recognition framework for semantic segmentation of visual structures, i.e. multi-class labeling at pixel-level, and apply it to the task of segmenting organs in the eviscerated viscera from slaughtered poultry in RGB-D images. This is a step towards replacing the current strenuous manual inspection at poultry processing plants. Features are extracted from feature maps such as activation maps from a Convolutional Neural Network (CNN). A Random Forest classifier assigns class probabilities, which are further refined by utilizing context in a Conditional Random Field. The presented method is compatible with both 2D and 3D features, which allows us to explore the value of the benefit of adding 3D and CNN derived features. The data set consists of 604 RGB-D images showing 151 unique sets of eviscerated viscera from four different perspectives. A mean Jaccard index of 78.11% is achieved across the four classes of organs by using features derived from 2D, 3D and a CNN, compared to 74.28% using only basic 2D image features.

1 Introduction

Poultry is both the most popular and the fastest rising source of meat in the USA [1]. The number of slaughtered chickens sold to US consumers was 8.6 billion in 2010, with the US broiler industry valued at $45 billion [2]. The health inspection at poultry processing plants involves visual inspection of the viscera, i.e. the internal organs of the chicken, after it has been extracted from the carcass. The inspection is necessary in order to ensure that the evisceration has been successfully performed and to examine whether the chicken shows any signs of disease. Hearts and livers are sold for human consumption. Therefore, it is important that these organs are extracted undamaged and show no signs of disease. Visual abnormalities such as inflammation makes organs such as live and heart, if not a health risk, then unappetizing to look at. Incomplete removal of viscera is a quality issue in chickens that are sold whole.

The current process for inspecting the viscera involves strenuous manual labour, which limits the throughput of the processing plant. The operators must inspect approximately three chickens per second. In addition to inspecting the viscera and corresponding carcasses, the birds that fail the inspection must be manually removed from the production line. Manual inspection is inherently slow, expensive and inconsistent; all of which represent a challenge when attempting to comply with the increasing demands from consumers.

Machine vision has previously been explored for inspection and grading of poultry. A great introduction to the problem and existing work on vision-based poultry inspection is given in [3], where a general purpose multispec-

tral classification system for high-speed online inspection of chicken carcasses is presented. In [4], a hyper-spectral imaging system for detection of external feces and ingesta is proposed. External inspection has been commercially available for some time, e.g. with the ClassifEYE [5] grading system for detecting damaged and impure carcasses. Inspection of the viscera, on the other hand, has received less attention. The most recent example of work in this area is [6], where entire poultry viscera are classified into four categories, i.e. normal, airsacculitis, cadaver, and septicemia, using a neuro-fussy model of RGB and spatial features. This model is able to classify the viscera correctly 82.5% of the time. In this paper, we take a different approach as the viscera is first segmented into the organs of interest. This segmentation will then serve as the basis for the subsequent work that deals with detection of disease and quality control. As an example, in [7], the segmentation approach described in our previous work [8], is used as the basis for classifying patches of liver as healthy or according to different diseases. The organs that are of interest during inspection are: heart, liver, and lungs. The rest of the viscera consists of the intestines, minor organs and connective tissue; all of which are to be classified together. Viscera are non-rigid bodies without straight lines or sharp edges and obey only a weak spatial arrangement. The nature of these objects are thus very different from the data that most object segmentation research deals with.

With this work, we exploit the recent advances in computer vision and pattern recognition by applying a modern visual perception framework to an unusual data set in a problem domain that has seen limited progress in the last decade. We make use of a convolutional neural network (CNN) and 3D derived features and demonstrate their contribution to the overall performance of the semantic segmentation of viscera. This paper is an extension of the work in our original conference paper [8]. The extended work includes four times the amount of data, CNN features and the addition of a conditional random field (CRF) on top of the random forest (RF) classifier. These improvements, in combination, lead to an increase in the mean pixel-wise Jaccard index from the 61.5% achieved in [8] to the 78.11% achieved here.

Contributions

This work can be summarized as: (1) Using CNN activation maps as super-pixel features. (2) Quantifying the value of CNN- and 3D-derived features for semantic segmentation. (3) Proposing a method for visual localization of organs enabling automated quality control.

2 Related work

Quality control of organic material is based largely on hyper spectral imaging (HSI). HSI makes it possible to capture nuances in colors that are normally not visible with RGB cameras [9–13]. In [14], splenomegaly is detected on poultry carcasses using ultra violet (UV) and color imaging. The use of UV aids in separating the spleen from the liver; something which proves difficult in RGB images. In [15], it is concluded that near-infrared imaging can be used to access quality measures like tenderness and color of fresh beef. [16] investigated 33 wavelengths in the visible spectrum and discovered two wavelengths, namely $600nm$ and $720nm$, that are optimal for detecting the gallbladder attached to the chicken liver.

More complex approaches with segmentation of multiple organs are less explored. An example is found in [17], where RGB images of pig offal are segmented into five classes (heart, liver, lungs, diaphragm, and an upper portion). This is achieved with a modified auto-context algorithm with an updating atlas, showing a small yet consistent improvement over the regular auto-context algorithm. A comparable problem domain is medical machine vision, e.g. the analysis of magnetic resonance brain scans. Given the relatively rigid structure of the brain, atlas-based approaches have proven successful in deforming expert-annotated samples to the target image. Multiple atlases can be used to increase the accuracy at the expense of higher computational cost [18, 19]. [20] compares an atlas-based approach to a tree-based approach using both computed tomography and magnetic resonance scans. A multivariate regression forest is used to predict the 3D bounding volumes of multiple organs based on appearance, spatial context and confidence prediction. The tree-based model was found to be more efficient than the atlas-based approaches, noting that tree-based methods implicitly contain a prior assumption of spatial information. The tree-based models are preferred in this application due to the weak spatial structure. Generalized models with little to no prior assumptions have also recently been proposed for various medical segmentation tasks suitable for a wide range of modalities [21]. Based on standard 3D features, they show improvements on three medical volume data sets using an iterative learning scheme of stacked contextual classifiers, where each stacked classifier adds complexity.

Most existing research in semantic scenes segmentation is focused on scenes with man-made objects. Contrary to the organs in our viscera dataset, man-made objects comprise, to a large extend, of straight lines and clean edges. Two widely used datasets with these types of objects are NYU-D 1 [22] and NYU-D 2 [23]. They both contain RGB and depth (RGB-D) images of indoor scenes. [24] is one example of recent work that addresses the NYU-D V2. They apply a framework, where features are extracted from superpixels

and classified using an RF. The label predictions are then refined using a CRF, which applies a pairwise smoothness term and learn contextual relationships between the different classes. A similar approach is found in [25]; instead of operating directly on the image channels of the RGB-D image, a point cloud is the basis for the oversegmentation and feature extraction. A similar approach is seen in [26] where a pixel-wise RF with random offset features and a CRF are expanded by using a stacked random forest. The stacked classifier learns when the previous classifier is mistaken and attempts to correct the errors.

Deep learning architectures for semantic pixel-wise segmentation have, as in many other areas, raised the bar significantly from the former approaches that relied on handcrafted features and classic classifiers. The VGG CNN architecture [27] has proven very versatile. Even though it was originally intended to address the image classification problem, it now forms the basis of many of the deep learning approaches that attempt to solve dense semantic segmentation data sets. [28] shows that the VGG architecture outperforms AlexNet [29] and GoogLeNet [30] when using its weights in their fully convolutional network (FCN). This is the first FCN to be rearchitected from pre-trained classification networks, such as AlexNet [29], VGG [27] and GoogLeNet [30]. The rearchitecting process is primarily done by discarding the final classification layer, converting the fully connected layers to convolutional layers and appending 1x1 convolutions in order to predict class scores at each coarse output location. By fine-tuning the entire pretained set of weights they gain a 30% better performance compared to only fine-tuning the output classifier. [31] proposes a new FCN architecture for semantic pixel-wise segmentation, where the low resolution feature maps produced by the encoding convolutional layers of the VGG model are upsampled to input image resolution by passing max-pooling indices from the encoding part of the network to the corresponding decoding layers. Finally, their architecture produces probability maps for each of the classes. The performance of their method is evaluated on the CamVid road scene segmentation dataset, the PASCAL VOC 2012 dataset and the SUN RGB-D indoor scene understanding dataset [32]. They compare their approach to several other deep learning and non-deep learning approaches, and they mostly achieve state-of-the-art performance along with the lowest computational cost. Again, in [33], the pre-trained weights of the VGG network are used in another adaption of a CNN to semantic image segmentation. They overcome the problems with the loss of spatial information associated with the use of CNNs, by replacing downsampling operators in the final max-pooling layers and upsampling the filters in subsequent convolutional layers.The score maps are then upsampled to input resolution using bilinear interpolation and passed to a fully connected CRF, which captures fine edge details and long-range dependencies. In [34], the authors propose to use the activations from across multiple layers

of a CNN to describe pixels and perform fine-grained semantic segmentation. The early layers contain low level information with a high localization precision, while the later layers capture high level semantic information that is much less sensitive to pose and placement. The combination of both types of features results in pixel-wise feature vectors that allow for high localization precision and great semantic classification.

3 Chicken viscera dataset

A data set consisting of 151 sets of chicken viscera was collected for our previous work [8]. Each set is captured from four different perspectives as shown in Figure C.1, resulting in a total of 604 RGB-D images. The viscera were captured from a distance of $35cm$ and placed in a hanger similar to the ones used on the production line. The viscera were taken directly from the production line and placed in the hanger, while retaining the same orientation as on the line. Data is captured using the Intel RealSense F200 3D

(a)	(b)	(c)	(d)
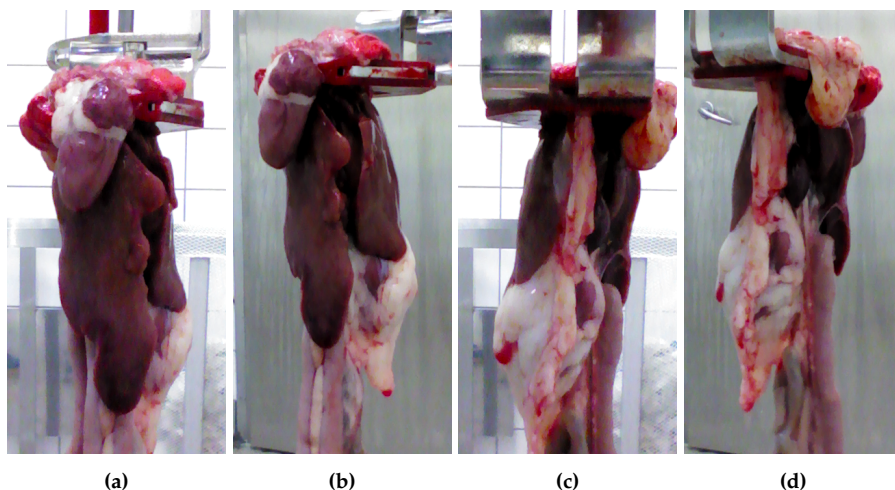
Fig. C.1: A single set of viscera as presented on the production line. Each set is captured from four perspectives. (a) Center Front. (b) Right Front. (c) Center Back. (d) Right Back.

camera. The RGB images are registered to the depth-maps, using the calibration provided by the Intel RealSense SDK, resulting in RGB-D images with a resolution of 480×640 pixels with an approximate region of interest of 150×350 pixels. Figure C.2 shows the different representations of the RGB-D input data, specifically the RGB image C.2a, the depth image C.2b and a point cloud C.2c.
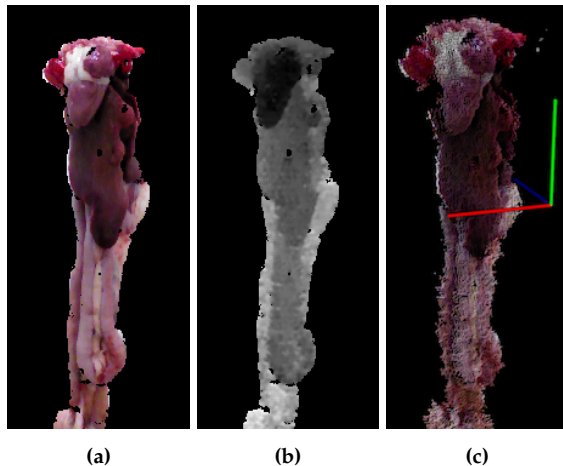
Fig. C.2: (a) RGB. (b) Depth. (c) Point cloud.

3.1 Ground-truth annotations

A ground-truth (GT) is needed to train and evaluate supervised learning algorithms. For this data set, the GT is obtained by manually annotating pixels belonging to the three classes of interest, namely heart, liver and lungs, and an additional fourth category encompassing miscellaneous organs and tissue. An ignore region is established around each organ, as labeling data is an ill-posed problem where the assignment of labels near borders is ambiguous. Figure C.3a shows an example of a viscera in RGB with the four classes lined out. Figure C.3b shows the corresponding grayscale pixel-level annotation and Figure C.3c shows the same annotation including a 2-pixel ignore boundary around each class. The ignore region furthermore contains pixels with no depth measurement, which can be seen as holes in the image. Pixels in the ignore region are excluded from both training and evaluation. The use of ignore regions is also seen in the popular pascal segmentation competition [35]. Ignore regions are introduced to avoid penalizing systems because of ambiguities in the labelling of pixels. At the same time, these are the areas that are the most challenging for semantic segmentation systems, thus ignoring them will unavoidably lead to improvements in the performance metrics. The specific width of the ignore regions used on this dataset, was the result of discussions with end-users. Because of the significant size difference for various organs and occlusion, the number of pixels for each class is skewed with 5% heart, 20% liver, 5% lung and 70% misc.

Unlike in our preliminary work [8], where only 151 images from a single perspective were used, this work makes use of all 604 images. Since the same viscera recur multiple times in the data set, although from different
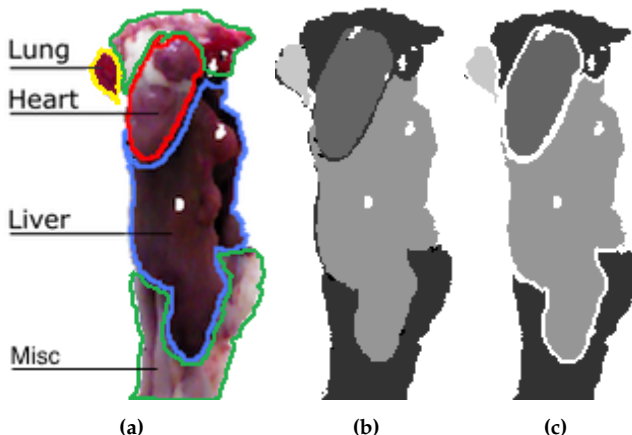
Fig. C.3: (a) RGB image of entrails with labeled organs. (b) Manual annotation in 2D. (c) Manual Annotation with ignore region used onwards in this work. The labels indicate with increasing intensity: misc., heart, liver and lung.

viewpoints, it is emphasized that images of the same viscera does not appear across training and test set.

4 Segmentation approach

In this paper we apply a framework for semantic segmentation inspired by the work of Müller [24] and Wolf [25]. The framework, as shown in Figure C.4, consists of oversegmenting the point cloud into supervoxels, from where features are extracted from feature maps, which are used to assign class probabilities in the RF classifier. A CRF optimizes the label assignments by taking the similarities between neighbors into account. The funding program supporting this work aims at demonstrating concepts in a collaboration between end-users, companies and one or more universities. Issues such as computational-, memory-, speed requirements are normally not included in such project. Instead a proof-of-concept is the target. In case of a successful project, the idea is that the company continues to mature the technology. We have therefore not been focusing on implementation issues and used standard C++ and python libraries whenever possible. Our current implementation is therefore not real-time. The used methodology is tailored to the specific type of dataset, in particular the very small dataset size and the organic objects. The presented method will not perform favorably when compared to the end-to-end deep learning based techniques that dominate large public RGB-D semantic segmentation benchmarks. Likewise, these state-of-the-art methods are difficult to get to perform well on such a specialized and lim-

ited dataset. The novelty of our framework lies in the use of CNN features and their application on an unusual data set, which leads to the omission of features commonly used addressing man-made objects.
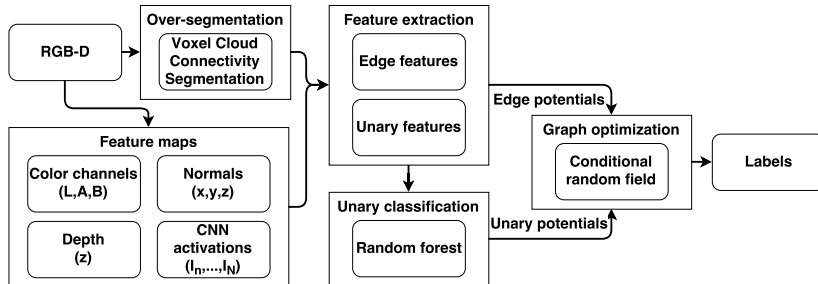

Fig. C.4: Overview of the segmentation framework.

Both the RF and the CRF are supervised classification methods that require training on labeled data. The dataset of 151 viscera sets are split into a RF training set of 91 sets, a CRF training set of 30 sets, while the remaining 30 sets are reserved for testing.

4.1 Oversegmentation

Oversegmentation is often used as a pre-processing step, where similar pixels are grouped into superpixels in the 2D case [36], or supervoxels in 3D [37]. Oversegmentation reduces the computational complexity at the later stages and can help reduce noise due to small variations, at the expense of a loss in precision depending on the oversegmentation's boundary adherence.

In this framework the Voxel Cloud Connectivity Segmentation (VCCS) [37] supervoxel segmentation algorithm is used, as the importance of each feature type that is used for oversegmentation, namely color, spatial, and geometric, can be adjusted using weights. This is convenient for exploring the value of features derived from 3D. VCCS produces the supervoxels by seeding the point cloud spatially evenly and an iterative clustering algorithm groups voxels within $\sqrt{3}R_{seed}$ of each seed in a 39-dimensional space based on spatial, color and geometric similarity.

4.2 Feature maps

Feature maps are obtained in preparation for the feature extraction and include the color channels from the LAB space, depth, normal magnitude for each of the three dimensions, and activation maps from across different layers of a CNN. Examples of each type of feature map are shown in Figure C.5. The viscera is subject to a loose composition, where heart and lungs are generally

found in the top while the liver usually configures in the central or bottom part of the image. Therefore, the position of a superpixel is a valuable feature. The CNN activation maps are extracted from each filter of each convolutional layer in the VGG-16 architecture [27]. From all of the convolutions layers, a total of 4224 CNN feature maps are produced. The used network weights originate from the network that was trained as part of the VGG team's entry to the ILSVRC-2014 competition and are used directly without fine-tuning. The specifics of the network are described in [27]. The activation maps that are lower resolution than the input are simply resized using linear interpolation. This use of CNN feature maps is inspired by the work of [34], where it is shown that extracting information at various layers across the network helps preserving the localization information present in the lower layers. This is done by creating pixel-wise feature vectors across the layers.

(a)   (b)   (c)   (d)   (e)   (f)   (g)   (h)
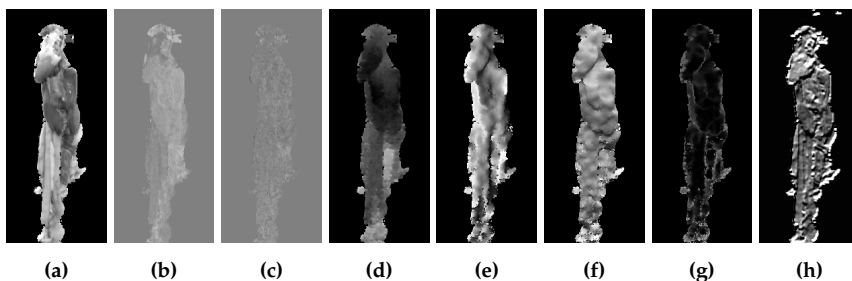
Fig. C.5: Feature map examples. (a) Channel L from LAB. (b) Channel A from LAB. (c) Channel B from LAB. (d) Depth map. (e) Normal magnitude in x direction. (f) Normal magnitude in y direction. (g) Normal magnitude in z direction. (h) CNN activation map.

4.3   Feature extraction

The features used for classification are extracted for each superpixel. The feature response for each feature map are averaged for all pixels in a superpixel. Two sets of features are extracted. Unary features are extracted for the RF, and secondly, pairwise features are found for the CRF.

Unary features are used in the RF to assign probabilistic labels to each superpixel. These probabilities serve as the unary component of the optimization performed later by the CRF. The features that are useful here are good at discriminating between the classes. The unary features used in the RF are listed in Table C.1.

Edge features are used to measure the similarity of neighboring superpixels and serve as the pairwise component of the CRF. The features that are used here must be good at describing the similarity of two superpixels. The edge features for the CRF are listed in Table C.1.

Table C.1: Unary and edge features.

Unary Features	Type	2D	3D	3D+CNN
LAB	Color	3	3	3
Center point	Spatial	2	3	3
CNN activation	Texture etc.	0	0	4224

Edge Features	Type	2D	3D	3D+CNN
LAB	Color	3	3	3
Center point	Spatial	2	3	3
Normal vector	Geometric	0	3	3

4.4 Unary classification

The unary potentials are used as an initial label estimate for the superpixels. Because the classification of the unary features is based entirely on local features, and since the neighborhood is taken into account at a later stage, these labels are assigned with a probability instead of one-hot encoding. Thereby, the later optimization can take into account the uncertainty of the initial classification. In this work the RF classifier [38] is used for the initial classification. The RF consists of an ensemble of label distributions in the leaf nodes trees. The label distributions are created, during training, from labeled training features that reach a particular leaf node when traversing through the RF.

The RF is implemented using the scikit-learn library [39]. The dataset is imbalanced due to the differences in organ size and occurrence. In order to compensate for this, prior probabilities that reflect the skewed distribution are assigned to each class. The number of trees in the forest is determined by examining the convergence of the out-of-bag error. The remaining parameters of the RF are optimized through a cross-validated grid search on 91 sets of viscera, reserved for training the RF. The class probabilities produced by the trained RF model are visualized with the example in Figure C.6. It is clear that the unary potentials give a rough location and segmentation for each organ. The model is certain for a few superpixels, especially for the miscellaneous part found in the bottom half of the image, as seen in Figure C.6d. The heart is less obvious with several faulty responses found on the liver. Using these local unary potentials alone and based on a hard-decision scheme where the most likely outcome is chosen across all classes, as seen in Figure C.7c, results in large parts of the liver being misclassified as miscellaneous.
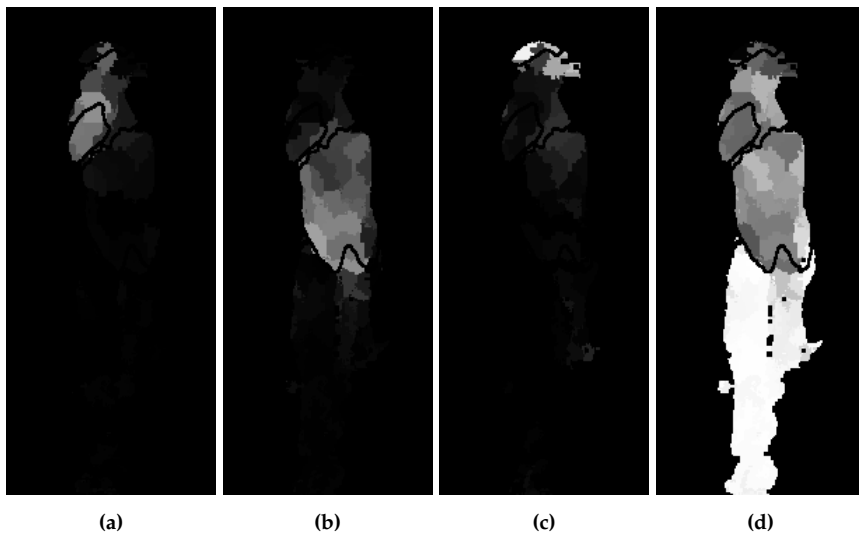
Fig. C.6: Random forest probabilities for each superpixel belonging to each of the four classes. (a) Heart. (b) liver. (c) lung. (d) misc. White values indicates higher certainty from the classifier.

4.5 Graph optimization

The random forest prediction based on local features can be further improved by taking neighborhoods into account. This is done using a CRF [40], where the learned edge potentials describe the likelihood of the different classes relationships between each other, according to the composition of the edge features between them.

The optimization that the CRF performs is based on the minimization of the energy function consisting of unary potentials and pairwise potentials. The unary potentials originates from the class label probabilities of the RF classifier. In this work the pairwise features consist of difference in position and color information, and normals in the 3D case, all learned from the separate training set. Note that each class is weighted according to the inverse class frequency. The CRF is implemented using the freely available PyStruct [41]. It is shown to have a smoothing effect as seen in Figure C.7, refining the RF classification by a large portion of especially the liver in this case. The CRF is able to correctly infer superpixels even over longer distances given certain unary potentials and discriminative pairwise potentials.

5 Evaluation

The value of adding 3D and CNN features is documented by progressively adding the different feature types to the semantic segmentation pipeline and
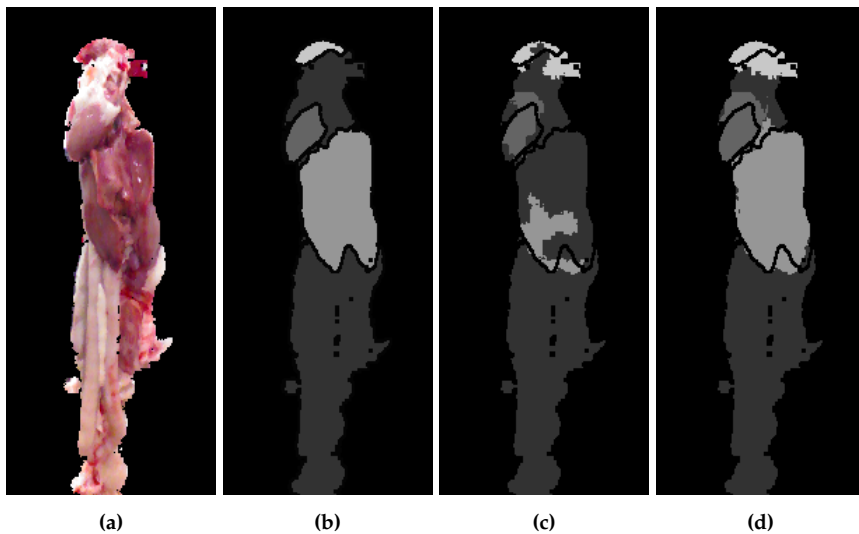
Fig. C.7: Final prediction where labels are, with increasing intensity: misc., heart, liver and lung. (a) input image (b) GT. (c) random forest prediction using 3D and CNN features, and (d) RF+CRF using 3D and CNN features.

evaluating the performance. The evaluation is based on 120 images displaying 30 unique viscera from different angles. Evaluation is performed at pixel-level. The regions with background and the 2-pixel wide boundary area around classes are excluded.

5.1 Quantitative analysis

The segmentation is scored by the Jaccard index shown in (C.1), measuring the similarity between the prediction and the manually labeled GT by counting true positives(tp), false positives(fp) and false negatives(fn). The evaluation is performed on pixel-level, where every pixel belonging to a given superpixel inherits the superpixel's label.

$$Jaccard\ index = \frac{tp}{tp + fp + fn} \tag{C.1}$$

The Jaccard index is found for each class and averaged, without weighting according to class occurrence, for a combined score. For the results based exclusively on 2D features, the features that are related to 3D are disabled. This is impacting the clustering into superpixels, as the geometric similarity is an important feature, as well as the RF classification from the lack of depth and the CRF from the lack of normal features. The additional information of both 3D and CNN is shown to improve the segmentation as seen in Table C.2,

Method	Features	Misc.	Heart	Liver	Lung	Class avg.
RF+CRF	2D	90.66	57.69	80.59	**68.18**	74.28
RF+CRF	3D	91.28	63.02	82.38	67.43	76.03
RF+CRF	3D+CNN	**91.58**	**70.17**	**83.64**	67.05	**78.11**
ASA		96.32	88.65	88.63	82.49	89.63

Table C.2: Pixel-wise Jaccard index for the four classes when evaluating using a 2-pixel wide ignore region.

especially on the more challenging heart class. The heart is small and much of it is covered by fat, which makes the color features much less effective. The lungs are also small and often occluded, therefore there is fewer pixels available for training and testing. Additionally, the lungs exhibit large variance in color, based on the amount of blood left in them.

The use of superpixels results in an upper limit on overall accuracy, which depends on the object boundary adherence of the oversegmentation. The impact of the oversegmentation is measured using an upper bound, namely the achievable segmentation accuracy (ASA) [42] defined in (C.2). Each superpixel/supervoxel S_k in $\mathcal{S}$ is labeled with the ground-truth label G_i with the largest intersection, and finally normalized where a score of 1 is a perfect fit. The maximal performance with the employed oversegmentation is show under ASA in Table C.2.

$$ASA_{\mathcal{G}}(\mathcal{S}) = \frac{\sum_k \max_i |S_k \cap G_i|}{\sum_i |G_i|} \tag{C.2}$$

The confusion matrices of Figure C.8 have been inspected to gain insight into the errors being made. The confusion matrices are normalized to account for the imbalanced dataset. The RF has a tendency to favor the miscellaneous class, with several misclassifications among all the other three classes. The CRF on the other hand, is a more balanced system with evenly distributed errors.

5.2 Analytic results

The examples shown in Figure C.9 are included to visualize the improvements shown in the quantitative analysis. Even though it is limited to a few percentage points, the improvement is visually evident, and can ultimately mean the difference between correctly segmenting e.g. a heart or not, as seen in the second row. The heart and lung classes are very small in terms of pixels. For this reason a small improvement in the precision of the unary classification may lead to a large jump in overall performance, especially if it leads to a small region of the heart to suddenly be found, allowing the CRF to
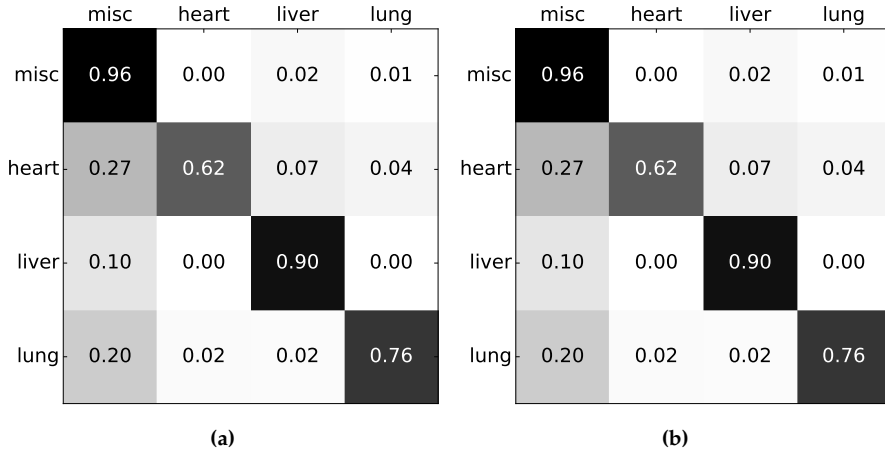
	misc	heart	liver	lung
misc	0.96	0.00	0.02	0.01
heart	0.27	0.62	0.07	0.04
liver	0.10	0.00	0.90	0.00
lung	0.20	0.02	0.02	0.76

(a)

	misc	heart	liver	lung
misc	0.96	0.00	0.02	0.01
heart	0.27	0.62	0.07	0.04
liver	0.10	0.00	0.90	0.00
lung	0.20	0.02	0.02	0.76

(b)

Fig. C.8: Normalized Confusion Matrices with the true label on the left and prediction on top. (a) RF with 3D+CNN features. (b) RF+CRF with 3D+CNN features.

get at better starting point for optimizing and correcting neighboring regions. The segmentation is found to be most challenged in the top where lung, heart and misc. are found in close proximity. On the other hand, segmentation of liver and misc. in the lower regions of the images is proving relatively robust. The RF seems to struggle in areas with high variance, resulting in almost arbitrary class predictions. This uncertainty is seen to traverse through the framework, complicating the task of the CRF. The CRF is highly dependent on the input from the RF, and thereby also the features used. Depending on the application, it is desirable to get somewhat continuous regions of the same label, i.e. the segmentation produces a single coherent region for the heart and another with liver etc. The CRF smoothing effect is found to overall improve the segmentation and produce more desirable regions.
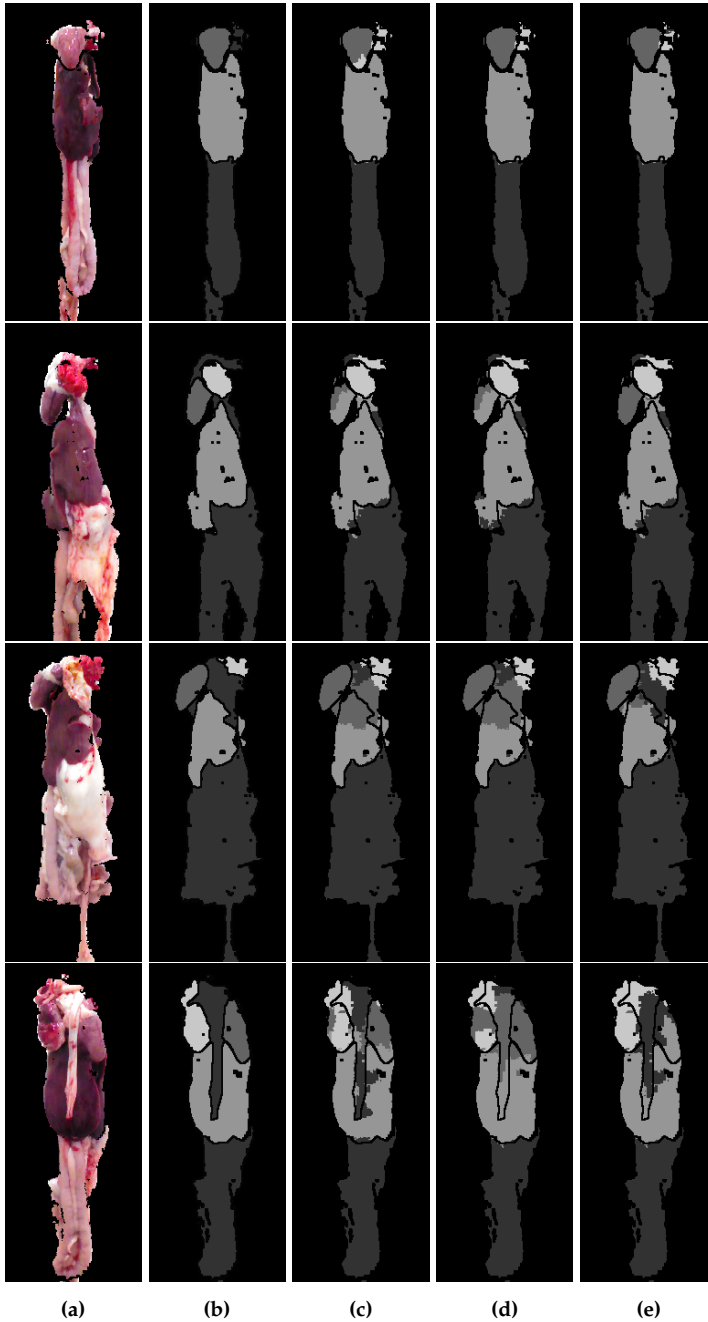
Fig. C.9: Examples for qualitative analysis. (a) RGB image. (b) ground-truth. (c) RF+CRF+2D. (d) RF+CRF+3D. (e) RF+CRF+3D+CNN.

We presented a framework for automatic visual inspection of poultry viscera, a domain with a weak-spatial arrangement without sharp edges and with organs often occluded by fat. We show how segmentation algorithms previously used on man-made objects are able to function on these deformable objects, scoring a mean Jaccard index of 74.28% across the four classes of organs by using basic 2D features. The addition of 3D features shows a potential improvement, with a score of 76.03% despite only small depth deviations of the viscera. Finally, adding CNN features achieves the mean Jaccard index 78.11%.

Author Contributions   Mark P. Philipsen conceived and designed the experiments with assistance from Jacob V. Dueholm. Anders Jørgensen assisted in the collection and preparation of data. Anders Jørgensen, Sergio Escalera provided feedback during the work and Thomas B. Moeslund supervised the entire process.

Conflicts of interest   The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript.

References

[1] D.-W. Sun, *Computer vision technology for food quality evaluation*. Academic Press, 2016.

[2] USDA, "Poultry statistics," https://www.ers.usda.gov/topics/animal-products/poultry-eggs/statistics-information.aspx, accessed: 2016-11-30.

[3] K. Chao, C. Yang, Y. Chen, M. Kim, and D. Chan, "Hyperspectral-multispectral line-scan imaging system for automated poultry carcass inspection applications for food safety," *Poultry science*, vol. 86, no. 11, pp. 2450–2460, 2007.

[4] B. Park, K. C. Lawrence, W. R. Windham, and D. P. Smith, "Multispectral imaging system for fecal and ingesta detection on poultry carcasses," *Journal of food process engineering*, vol. 27, no. 5, pp. 311–327, 2004.

[5] I. Food, "Classifeye - inspection, grading and quality assurance," http://www.ihfood.dk/poultry2111, accessed: 2016-12-02.

[6] K. Chao, Y.-R. Chen, H. Early, and B. Park, "Color image classification systems for poultry viscera inspection," in *Photonics East (ISAM, VVDC, IEMB)*. International Society for Optics and Photonics, 1999, pp. 110–120.

[7] A. Jørgensen, J. Fagertun, and T. B. Moeslund, "Diagnosis of broiler livers by classifying image patches," in *Scandinavian Conference on Image Analysis*. Springer, 2017, pp. 374–385.

[8] M. P. Philipsen, A. Jørgensen, S. Escalera, and T. B. Moeslund, "RGB-D Segmentation of Poultry Entrails," in *Articulated Motion and Deformable Objects: 9th International Conference*, vol. 9756, 2016, pp. 168–174.

[9] K. Chao, C. C. Yang, M. S. Kim, and D. E. Chan, "High throughput spectral imaging system for wholesomeness inspection of chicken," *Applied Engineering in Agriculture*, vol. 24, no. 4, pp. 475–485, 2008.

[10] B. P. Dey, Y. R. Chen, C. Hsieh, and D. E. Chan, "Detection of septicemia in chicken livers by spectroscopy,," *Poultry Science*, vol. 82, no. 2, pp. 199–206, feb 2003.

[11] H. Huang, L. Liu, and M. O. Ngadi, "Recent developments in hyperspectral imaging for assessment of food quality and safety." *Sensors (Basel, Switzerland)*, vol. 14, no. 4, pp. 7248–76, 2014.

[12] E. Z. Panagou, O. Papadopoulou, J. M. Carstensen, and G. J. E. Nychas, "Potential of multispectral imaging technology for rapid and non-destructive determination of the microbiological quality of beef filets during aerobic storage," *International Journal of Food Microbiology*, vol. 174, pp. 1–11, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.ijfoodmicro.2013.12.026

[13] C. H. Trinderup, A. L. Dahl, J. M. Carstensen, K. Jensen, and K. Conradsen, "Utilization of multispectral images for meat color measurements," in *Workshop on Farm Animal and Food Quality Imaging 2013*, 2013, pp. 42–48.

[14] Y. Tao, J. Shao, K. Skeeles, and Y. R. Chen, "Detection of splenomegaly in poultry carcasses by UV and color imaging," *Transactions of the Asae*, vol. 43, no. 2, pp. 469–474, 2000.

[15] G. Elmasry, D. W. Sun, and P. Allen, "Near-infrared hyperspectral imaging for predicting colour, pH and tenderness of fresh beef," *Journal of Food Engineering*, vol. 110, no. 1, pp. 127–140, 2012.

[16] A. Jørgensen, E. M. Jensen, and T. B. Moeslund, "Detecting gallbladders in chicken livers using spectral imaging," in *Proceedings of the Machine Vision of Animals and their Behaviour (MVAB)*. BMVA Press, September 2015, pp. 2.1–2.8. [Online]. Available: https://dx.doi.org/10.5244/C.29.MVAB.2

[17] T. Amaral, I. Kyriazakis, S. J. Mckenna, and T. Ploetz, "Weighted atlas auto-context with application to multiple organ segmentation," *Proc. WACV*, 2016.

[18] J. M. Lötjönen, R. Wolz, J. R. Koikkalainen, L. Thurfjell, G. Waldemar, H. Soininen, D. Rueckert, A. D. N. Initiative *et al.*, "Fast and robust multi-atlas segmentation of brain magnetic resonance images," *Neuroimage*, vol. 49, no. 3, pp. 2352–2365, 2010.

[19] H. Wang, J. W. Suh, S. R. Das, J. B. Pluta, C. Craige, and P. A. Yushkevich, "Multi-Atlas Segmentation with Joint Label Fusion," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 3, pp. 611–623, 2013.

[20] A. Criminisi, D. Robertson, O. Pauly, B. Glocker, E. Konukoglu, J. Shotton, D. Mateus, A. Martinez Möller, S. G. Nekolla, and N. Navab, *Anatomy Detection and Localization in 3D Medical Images*. Springer London, 2013, pp. 193–209.

[21] F. Sampedro, S. Escalera, and A. Puig, "Iterative multi-class multi-scale stacked sequential learning: Definition and application to medical volume segmentation," *Pattern Recognition Letters*, vol. 46, pp. 1 – 10, 2014.

[22] N. Silberman and R. Fergus, "Indoor scene segmentation using a structured light sensor," in *Proceedings of the International Conference on Computer Vision, Workshop on 3D Representation and Recognition*, 2011.

[23] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from rgbd images," in *European Conference on Computer Vision*. Springer, 2012, pp. 746–760.

[24] A. C. Müller and S. Behnke, "Learning depth-sensitive conditional random fields for semantic segmentation of rgb-d images," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, May 2014, pp. 6232–6237.

[25] D. Wolf, J. Prankl, and M. Vincze, "Fast semantic segmentation of 3d point clouds using a dense crf with learned parameters," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, May 2015, pp. 4867–4873.

[26] M. Thøgersen, S. Escalera, J. Gonzàlez, and T. B. Moeslund, "Segmentation of RGB-D indoor scenes by stacking random forests and conditional random fields," *Pattern Recognition Letters*, vol. 80, pp. 208 – 215, 2016.

[27] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[28] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.

[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[30] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.

[31] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *arXiv preprint arXiv:1511.00561*, 2015.

[32] G. Patterson and J. Hays, "SUN attribute database: Discovering, annotating, and recognizing scene attributes," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 2751–2758.

[33] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *arXiv preprint arXiv:1606.00915*, 2016.

[34] B. Hariharan, P. Arbeláez, R. Girshick, and J. Malik, "Hypercolumns for object segmentation and fine-grained localization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 447–456.

[35] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The PASCAL Visual Object Classes (VOC) Challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2010.

[36] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2281, 2012.

[37] J. Papon, A. Abramov, M. Schoeler, and F. Wörgötter, "Voxel Cloud Connectivity Segmentation - Supervoxels for Point Clouds," in *Computer Vision and Pattern Recognition (CVPR)*, June 2013, pp. 2027–2034.

[38] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[39] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[40] J. Lafferty, A. McCallum, and F. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data," in *Proceedings of the 18'th International Conference on Machine Learning, ICML*, vol. 1, 2001, pp. 282–289.

[41] A. C. Müller and S. Behnke, "PyStruct - Learning Structured Prediction in Python," *Journal of Machine Learning Research*, vol. 15, pp. 2055–2060, 2014.

[42] M. Y. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy Rate Superpixel Segmentation," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2097–2104, 2011.

Paper D

Three Camera Setup for Image Capture of Moving Broiler Viscera

Anders Jørgensen

The report has been used internally at IHFood

Abstract

This report describes the design of a capture system for gathering images to use in the development of a computer vision system for inspection of broiler viscera. The aim is to capture thousands of viscera images automatically. 3 cameras are used to gain more coverage of the viscera and ultimately to remove/handle specular reflections.

Besides collecting the images it is also vital that the viscera sets are graded and this report discusses how to get a balanced data set when less than 1 % of the broiler population is sick.

1   Introduction

The production of chicken meat has come a long way since the farmer chopped the head of an old hen behind the hen house. Today it is a big industry where trucks bring broilers to the slaughter houses where they are slaughtered at rates about 3.5 birds per second (bps). The slaughter houses want to increase their production speed, but health inspection is becoming a bottleneck as EU legislation requires a post mortem inspection of every slaughtered animal [1]. This check is also vital to maintain a high standard of food safety which is characteristic for Danish produce.

This report describes the development of a system designed to record multiple images of each viscera eviscerated from the slaughtered broilers. The images should be used in the development of a classification system that can aid the veterinarians during the health inspection. The recording system is developed in collaboration with Linco Food Systems A/S who handles the installation and the interface with the existing systems at Danpo A/S where the system is to be installed.

1.1   Post mortem inspection

The post mortem inspection today is done in two stages. The first inspection is done of the full body after the feathers have been plucked from the broiler. This is an inspection for overall signs of diseases and while mostly done visually the veterinarian is equipped with a knife for when it is necessary to check for symptoms beneath the skin. (However, experiences from post mortem inspection of pigs have led to a prohibition of touching the pigs as it increases the risk of diseases and bacteria spreading between the carcasses.)

The second stage is after evisceration where the viscera are presented alongside the carcass. The viscera are inspected for signs of diseases, in which case both the viscera and the carcass must be removed from the conveyor. Two inspectors are positioned side by side to make sure everything is detected. At this stage carcasses are also removed if they have been contami-

nated by gall during the evisceration. A picture taken from inspection stage two can be seen in figure D.1.

Fig. D.1: Picture taken as the viscera and carcasses enter the stage two inspection site.

During both stages the broiler come by at around 3.5 bps. That makes the average inspection time per broiler less than 300 ms, 600 ms per viscera set if the inspections only check every other set. But the inspectors at stage two must also have an eye on the carcasses, which again brings the inspection time per viscera down. A sketch of the inspection system can be seen in figure D.2.
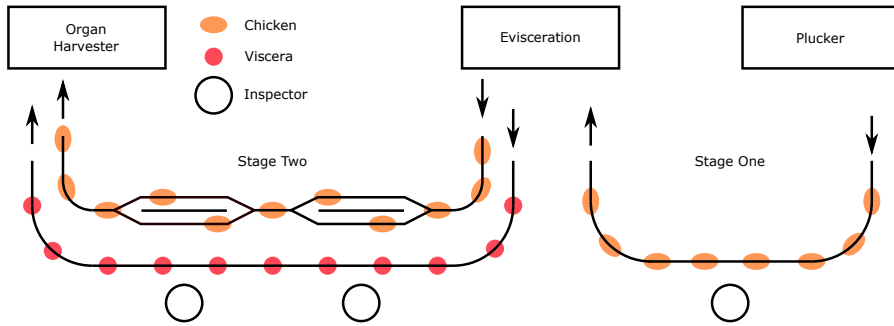
Fig. D.2: Sketch of the inspection area.

1.2 Broiler viscera

The viscera are inspected as diseases in the organs can affect the entire bird through the blood. The organs which are most often infected are the heart, liver and lungs, however during the two-week image recording for the pre-analysis, see paper A, only heart-, liver- and systemic diseases were discovered. Figure D.3 highlights the most vital organs in a viscera set.
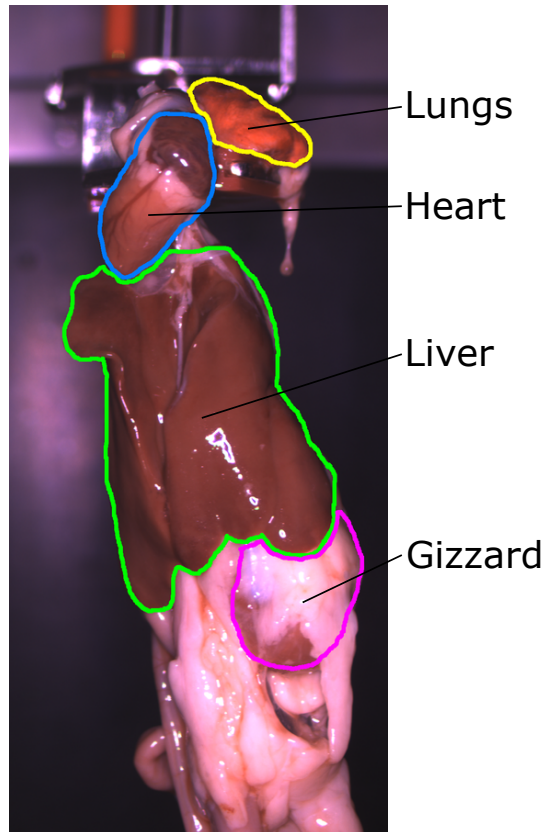
Fig. D.3: Viscera with vital organs marked. Green is liver, blue is heart, yellow is lung (only one of two is visible) and purple is gizzard. The part of the viscera "below" the image is the intestines.

1.2.1 Diseases

For a normal population of broilers, just about 1% of the flock is unhealthy. Roughly half of these are detected at stage one, according to the veterinarians, leaving 0.5% unhealthy broilers for stage two. This means that on average when slaughtering 12.600 broilers per hour, 63 unhealthy viscera sets are taken of the conveyor. The diseases are however often correlated within a flock of broilers from the same farm, meaning the number of unhealthy viscera will vary from flock to flock. For an overview of the most frequent diseases and their distribution see paper A.

1.3 Obtaining ground truth

An important aspect of data recording is the gathering of ground truth. The inspectors at stage two take the unhealthy viscera off the conveyor, so the images should be captured before that. In order to not interfere with the production, we cannot ask the veterinarians to grade the viscera as it is taken of the conveyor. It would also be near impossible to keep track of a viscera set and link the grading to an image captured by the camera a couple of seconds before. It is therefore found that the best solution is to have a veterinarian grade the images after they have been recorded. This means that the veterinarian will have to grade the viscera only based on images, but this is also what the computer vision software would do. It should later be investigated whether a grade from an image corresponds to the same grade of the real viscera.

Based on the normal broiler population, described in section 1.2.1, only 1 out of 200 viscera sets will be unhealthy. As the system should capture images before the grading, the resulting dataset will be very biased towards healthy sets. This is unwanted as the healthy images will fill up the hard drive and the veterinarians would spend their time grading too many healthy images. An automatic way to trim the dataset before grading is therefore needed. This is described further in section 2.2.

The inspectors at stage two have a mirror enabling them to look at the "backside" of the viscera set. The liver and heart most often face towards the inspectors but depending on how the clamp grabs the viscera during evisceration it happens that it is rotated around its own vertical axis. A vision system would need multiple cameras to inspect all around the viscera, but this is not deemed necessary for data collection. One can assume that the inspection task from the "backside" is equal to that of the "front". A purpose of this work is therefore also to analyse how many cameras are needed for a complete inspection of a viscera set.

2 Challenges and considerations

The viscera are transported via a conveyor belt where on they hang in a clamp. This is how they will be presented for a future vision system so this is how the images for the data set should be captured as well. The images should be captured automatically as it would be infeasible to do it manually. It is also important that the capture system does not interfere with the daily production, but is able to run autonomously.

This work is the continuation of a previous study about the spectral analysis of the major viscera diseases, paper A. Results of that study are used to determine the camera resolution and the spectral design of the light sources.

2.1 Presentation of viscera

Figure D.4 shows how the viscera sets are moved along the conveyor belt. The vital organs are gathered at the top near the clamp and the intestines hang below. The length on the intestines can vary a lot and sometimes they completely unroll and are dragged after the rest of the viscera as it can be seen to the right in figure D.4. As the diseases are mostly located in the heart, liver and lung, the cameras should focus on the area in the green box. This will also help maintaining a high spatial resolution on these organs.
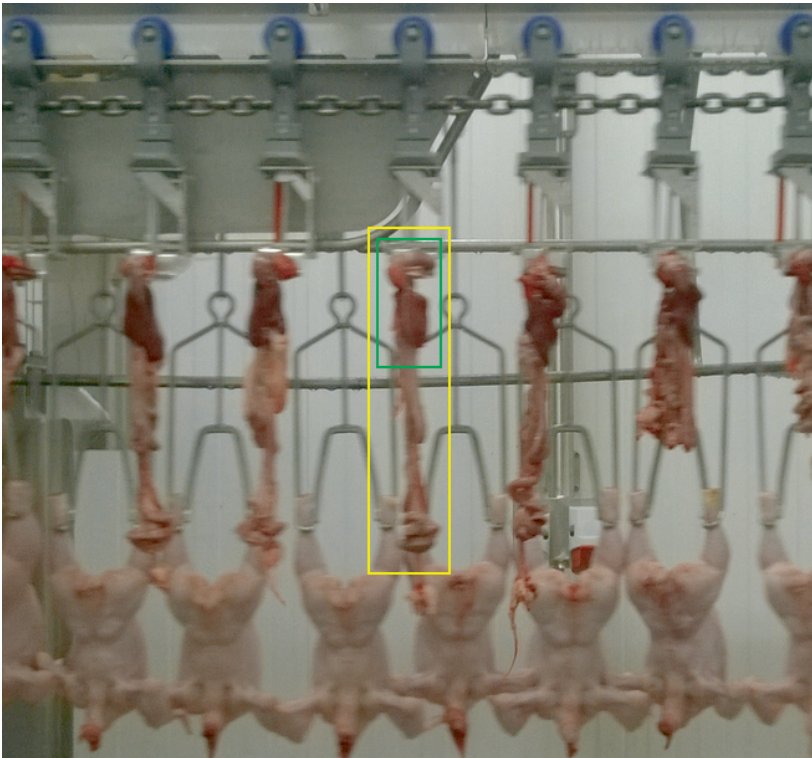

Fig. D.4: The viscera as it is presented on the conveyor in front of the carcasses. The green box denotes the heart, liver and lung and the yellow box marks the entire viscera set.

2.1.1 Conveyor speed

The capture system should preferably work at speeds up to 4 bps. This is faster than the normal processing speeds, but the chain speed is adjustable so that the speed can be increased if necessary. The viscera is positioned 15.24 cm (6") apart on the conveyor. This means that at full speed the chain moves 60.96 cm/s or 0.6096 m/s.

2.1.2 Setup location

The capture system is not positioned in the same location as the inspectors at stage two, due to limited space and to not interfere with their work. It is instead located in the evisceration room, just after the evisceration. This room has some very rough conditions with a high humidity and high temperatures. The room is also cleaned daily with chemicals and water pressure cleaners. The cameras and light sources should be able to withstand this kind of treatment. Figure D.5 shows the location where there is room for the camera setup.

Fig. D.5: Site were the cameras should be installed.

2.1.3 Room lighting

The evisceration room is lit with fluorescent lights. It will be possible to shield the viscera from some of the light, but the equipment should be easy to clean and the conveyor should be able to pass freely by the cameras. It can be seen in figure D.5 that the intestines extend a long way down from the clamp that holds the viscera. Because of the turns on the conveyor this makes the viscera swing back and forth as is moves along. It is important the viscera do not hit any covering as that would increase the risk of cross contamination. This limits the amount of ambient light you can block from the viscera.

2.2 Balancing the data set

To avoid capturing 199 healthy viscera sets for every sick, we need an automatic way to trim the incoming data. The sick viscera sets are removed by the inspectors which means that the clamp is now empty on the conveyor. This can be detected with an optical sensor. An overview of the processing line at Danpo can be seen in figure D.6. The images are captured right after the evisceration and the viscera continues to inspection stage two in figure D.2. The optical sensor is placed after the inspection and for each clamp it detects whether it is empty or not. If it is empty a message is sent to the vision system. Details are provided in section 3.7.
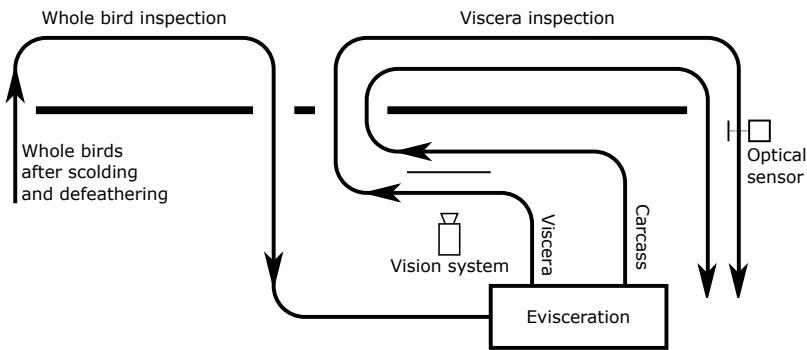

Fig. D.6: Sketch of the processing flow.

2.3 Design specifications

One way to deal with unwanted ambient light is to capture the images with a low exposure time. This will also limit blur when capturing moving objects but it requires a strong light source and introduces specular highlights on glossy surfaces. A common way to remove or reduce the effect of specular highlights is to use polarising filters. However, even with perfect polarising filters, this approach would require a light source four times stronger to get the same illumination in the image. In this setup we will therefore handle highlights by taking multiple images of the same viscera set with light sources at different positions. This should cause the specular highlights to be at different positions at the viscera set as well. The images can then be preprocessed into a single image without highlights or processed individually.

2.3.1 Light source considerations

In the pre-analysis, paper A, it was found that the best light for detecting liver diseases was at least two wavelengths, one between ██ nm and ██ nm

and one between ▮ nm and ▮ nm. This corresponds to a blue and a red colour. An LED, in the right package and with the right voltage specifications, should be picked with these two ranges.

The light source should be "pulseable", meaning that it can be switched on and off very quickly. It is possible to get more light from the LEDs as they can withstand a higher voltage when pulsed, than when turned on continuously. This way it also uses less power and dissipates less heat, as the light only flashes when the camera is exposing.

2.3.2 Camera considerations

The important camera specifications for this setup is resolution and frames per second (fps), which are often inversely proportional. The camera used in the pre-analysis, paper A, had a vertical resolution of 1024 pixels and covered roughly the area marked by green in figure D.4. The viscera photographed were neatly placed in the centre of the image, which can not be ensured for this setup. It is therefore estimated that the vertical resolution should be at least around 1500 pixels to account for the variation in viscera height.

A high frame rate is needed to avoid the viscera from moving too much between the two images. The viscera swings and wobbles as it moves along the conveyor and the closer in time the images are captured the smaller is the motion. Less local motion means it is easier to register the two images for specular highlight removal.

The images must be captured with both a red and a blue light source, as stated in section 2.3.1. This is typically done as two consecutive captures with a monochrome camera where the light pulses in turns. But because the two wavelengths are in each end of the visible spectrum it would be possible to capture both wavelengths at the same time with a colour camera and still have a minimum of cross talk between the colour channels. To minimize cross talk even more one could also use a prism camera with 2 or 3 sensors.

The viscera are transported at 60.96 cm/s which means it moves roughly 1 cm every 16 ms, which again corresponds to 60 fps. The camera should be able to burst at 60 fps, even though it only needs to capture 8 images per second (two images for every viscera set at 4 bps).

The camera should be trigged by an external source so it can be activated when a viscera set passes by the camera. This is standard for most industrial cameras.

3   System design

This chapter will describe the design of the hardware parts in the system. It evolves around three cameras and three light sources. An external trig deliv-

ered by Linco activates the recording routine that captures multiple images with changing light sources. A computer receives and saves the images.

3.1 Camera and light positions

Besides getting images of viscera the setup is also a test of how much one camera can cover. Two cameras are positioned next to each other at a roughly $30°$ angle, both looking horizontally at the viscera. A third camera is located slightly above the other cameras looking down onto the clamp where the heart and lungs are located.

The two cameras placed at the side of the viscera should share light sources, so that each camera captures two images, one with each light source turned on. The third camera only captures one image with its own light source. Figure D.7 shows a sketch of the setup. Each tube contains both a camera and a light source. The tubes should be placed at least 30 cm from the viscera so that they can be accessed by the cleaning personal. Moving the cameras closer would also require the use of a lens with a wide field of view and the distortions that follow.
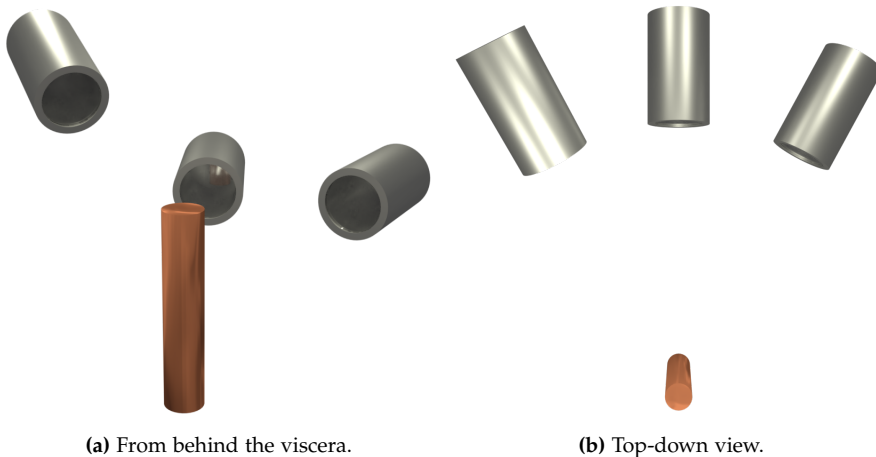

(a) From behind the viscera.　　　　**(b)** Top-down view.

Fig. D.7: Sketch of the setup. Each tube contains both a camera and a light source. The orange tube depicts the viscera set.

3.2 Casing for camera, lens and light

The casing for the camera and light must be able to keep the equipment dry in an environment with high humidity and daily cleaning by water pressure cleaners. IHFood have a tube design, used in other projects, that fits both camera, lens and light. The tube can be seen in figure D.8. It was decided to

use this casing as it is already tried and tested and fits the requirements for this system.

Fig. D.8: The vision tube fitted with camera and LED for another project. The LEDs are fitted on PCB around the lens.

The inner diameter of the tube is 84 mm and the inner length is 160 mm. This is therefore the maximum size of the camera and lens, including cables and connectors. The LEDs must be mounted on a piece of circular PCB mounted in the top of the tube. The PCB must have a hole in the middle where the lens can stick through. Figure D.9 shows a sketch of the inside of a tube.
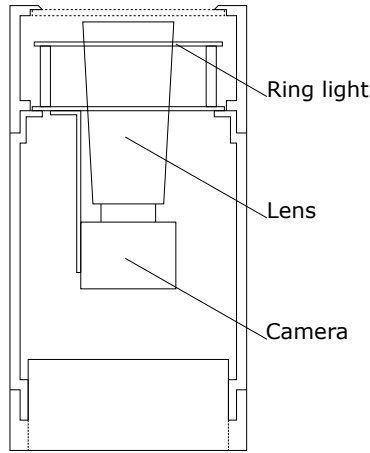
Fig. D.9: Sketch of the tube with camera, lens and ring light. Not exact measures.

3.2.1 Camera

Several cameras were investigated, but we ended up selecting the LW-AL-IMX253C-USB3 [2] from Imaging Solutions Group (ISG) for several reasons. It uses a 12MP IMX253 sensor from Sony that has a resolution of 4096x3000. The camera has a burst frame rate of 68 fps and is connected via USB3, so it requires no extraordinary hardware like CoaXPress. It uses C-Mount or CS-Mount lenses and it fits inside the tube. Important specifications can be seen in table D.1.

Camera Specifications	
Model Name	LW-AL-IMX253C-USB3, COLOR
Image Sensor Model	SONY Pregius IMX253
Image Sensor Format	1/1.2"
Max. Resolution	4096x3000, 12MP
Max. Frame Rate	68 fps
Pixel Size	3.45 μm x 3.45 μm
Shutter	Global
Interface Port	USB3
Dimensions (W x H x L)	52 mm x 46 mm x 52mm
Diagonal (Calculated)	64.43 mm
Lens Mount	C-Mount or CS-Mount

Table D.1: Camera specifications for LW-AL-IMX253C-USB3 [2].

It is a colour camera using a Bayer pattern for colour generation. One could therefore argue that the true image resolution is 2000x1500 if no in-

114

terpolation is used. If the camera is turned on the side it will still have a vertical resolution of 2000 pixels compared to the 1500 pixels specified in section 2.3.2.

These cameras come with an IR filter by default, but that was removed for this project as not to attenuate the light at the red wavelength.

3.2.2 Lens

The lens should be a C-mount or CS-mount to match the camera. The working distance should include 30-50 cm with a field of view that matches the green box in figure D.4. The physical size of the lens is also important as is should be able to fit in the tube with the camera.

It was decided to go with the Kowa LM25XC [3]. The specifications are listed in table D.2.

Lens Specifications	
Model Name	LM25XC
Main Sensor size	4/3"
Focal Length	25 mm
Iris range (F-stop)	f/2 - f/16
Focusing range (m)	0.15 m - ∞
Size (Ø x L)	45 mm x 89 mm
Mount	C-Mount
Horizontal Field of View	
2/3-inch Sensor	20.1 °
1-inch Sensor	28.9 °
4/3-inch Sensor	40.9 °

Table D.2: Lens specifications for Kowa LM25XC [3].

The focus range is well below 30 cm. The length is specified without the mount, which will be screwed into the camera's mount, so the total length of camera and lens will be 135 mm. Figure D.10 shows the camera with the lens mounted.

Fig. D.10: ISG camera with Kowa lens mounted.

The horizontal field of view (FoV) for the camera's 1/1.2-inch sensor can be roughly interpolated from the specified FoVs. Figure D.11 shows that the FoV should be around 24.4 for that sensor size.
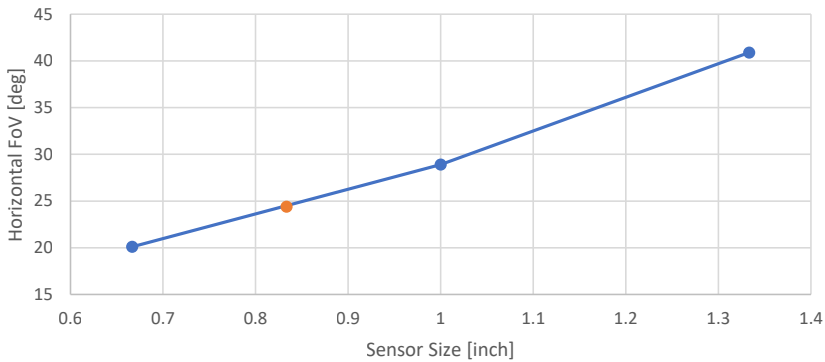

Fig. D.11: Estimating the horizontal FoV a sensor size of 1/1.2". Linear interpolation is used to connect the data points. The orange dot is placed at $x = 0.83\overline{3}$" yielding $y = 24.4°$.

The horizontal view width can be calculated as $2l_{wd}\sin(12.2°)$ for different work distances, l_{wd}. For 30 cm the view width is 12.7 cm and for 50 cm the view width is 21.1 cm. These numbers should be used as guidance, when positioning the cameras.

3.3   Light design

The light should fit inside the tube as described in section 3.2, on a PCB formed as a disc with a hole in the middle for the camera lens. The outer radius should be 46.5 mm with room for eight mounting holes along the perimeter. The hole in the centre should have a radius of 25 mm to make sure the lens will fit. As the wavelengths are already specified the task is to find LEDs that match these wavelengths and fit physically in the case. The ████████████████████ from Lumileds were the best option with the ██████████ (peak wavelength at ██ nm) and the ██████████ (peak wavelength at ██ nm) LEDs.

LED Specifications		
	████████	████████
Part Number	████████	████████
Peak Wavelength	██ nm	██ nm
Forward Voltage (Typical)	2.20 V	2.85 V
Maximum DC Forward Current	700 mA	1000 mA
Radiometric Power (Typical)	325 mW	575 mW
Size (W x H x L)[mm]	1.30 x 0.96 x 1.70	1.30 x 0.56 x 1.70

Table D.3: Specifications for the LEDs.

Each colour is fitted in a circle around the hole to equally distribute the light around the camera lens. It was decided to add a green LED, also from Lumileds, to make it possible to capture images with more life-like colours. This would be useful for the veterinarians when grading the images afterwards. It should be possible to switch off the green LEDs if they interfere with the other colour channels, so every set of LEDs can be switched off individually. Each ring consists of 30 LEDs. A sketch of the ring light can be seen in figure D.12.
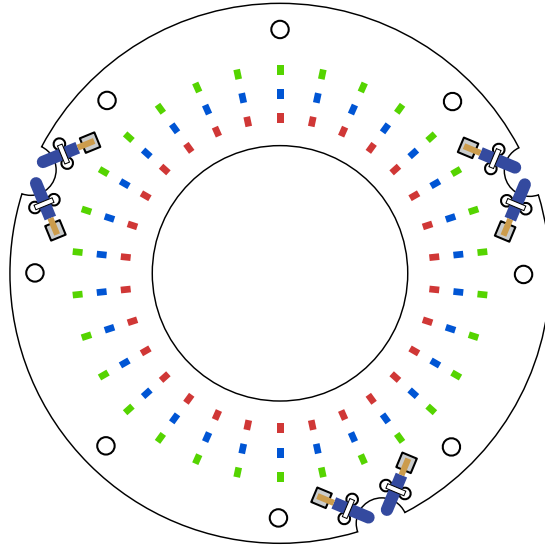
Fig. D.12: Sketch of the layout for the ring light PCB. The red, green and the blue diodes can be switched on and off individually.

The specifications for the green ████████████████ LED can be seen in table D.4.

LED Specifications	
	████
Part Number	████████
Peak Wavelength	██ nm
Forward Voltage (Typical)	2.85 V
Maximum DC Forward Current	1000 mA
Luminous Flux (Typical)*	100 lm
Size (W x H x L)[mm]	1.30 x 0.56 x 1.70

Table D.4: LED specifications for the Green LED. *Radiometric Power not specified for the green LED.

The LEDs are driven by Gardasoft PP420 light controllers [4]. These are configurable via a web interface and can control the pulse width by which the LEDs are strobed. Specifications are listed in table D.5.

Gardasoft PP420 Specifications	
User Interface	Ethernet
Output channels	Four independent constant current outputs with SafeSense^{TM}.
Output Current	From 0 mA to 10 A in steps of 2.5 mA. Up to 2 A per channel continuous or 10 A pulsed. Total 2 A maximum average output current from the controller
Trigger Inputs	4 opto-isolated digital inputs. Require 3 V to 24 V at 3 mA.
Pulse Width Timing	From 20 µs to 999 ms in steps of 20 µs. Timing repeatability 0.1 µs
Output Voltage	0 V to 47 V
Supply Voltage	Regulated 12 V to 48 V. The supply voltage must be at least 1 V higher than the output voltage required by the lighting.
Dimensions	118 mm long by 76 mm wide by 27 mm high (excluding DIN fixing)

Table D.5: Specifications for Gardasoft PP420 Light Controller [4].

Each ring of LEDs are configured as 6 LEDs in series in 5 parallel lanes, as displayed in figure D.13.
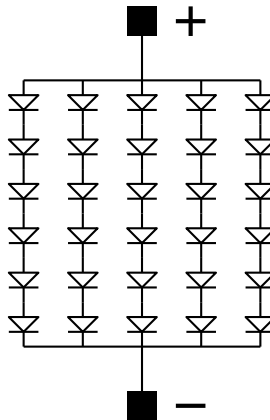

Fig. D.13: LED configuration for one colour ring.

Total voltage levels are calculated for each LED's typical voltage in table D.6. You can overdrive the LEDs when only pulsing the LEDs for a short time period. There should still be room for that as the PP420 is supplied with 48 V and can draw 10 A per channel when pulsed.

Typical voltage	2.2 V	2.85	2.85 V
Forward current @ Typical voltage	400 mA	250 mA	300 mA
Total voltage for 6 diodes	13.2 V	17.1 V	17.1 V
Total current for 5 lanes	2000 mA	1250 mA	1500 mA

Table D.6: Total voltage and current usage at typical levels. Forward currents are read from figure D.26 and D.27.

Each light needs one PP420 to support the three channels needed to control the three colours, so a total of three is needed.

An image of the finished PCB manufactured by Meodex can be seen in figure D.14.
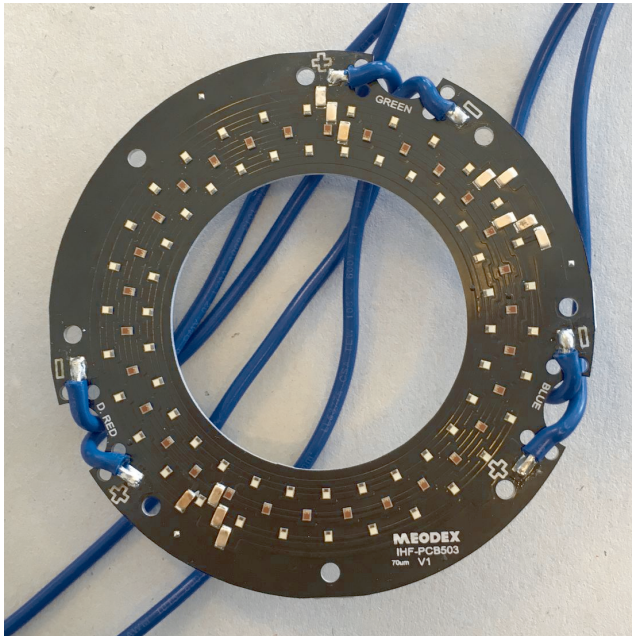
Fig. D.14: The finished ring light PCB.

3.4 Computer

The entry point of the system is the computer. It should have three USB3 ports for the cameras and one USB2 port for the Fooduino (see section 3.5). For connections it should also have two Ethernet ports, one for internet access and one for communication with the rest of the system. The BLUE LINE FANLESS FPC-6000 meet these requirements. To save the large amount of

data generated by the cameras, it requires a large SSD. The computer is there-fore fitted with a 1TB Samsung 850 Evo. For calculations on recording times see section 3.7. Important specifications can be seen in table D.7.

Computer Specifications	
Model Name	FPC-6000
CPU	i5-3610ME
RAM	8GB DDR3
Storage	1TB Samsung 850 Evo SSD
USB connections	2xUSB2.0 4xUSB3.0
Ethernet connections	2xRJ-45 Gigabit LAN
Dimensions (H x W x L)[mm]	93 x 215 x 272

Table D.7: Computer specifications for FPC-6000 [5].

3.5 Capture sequence

A total of five images should be captured for each viscera that passes the setup. Four images by the two side cameras and one by the top camera. A trigger signal is supplied whenever a clamp is present in front of the setup. This signal is registered by a Fooduino, which is an in-house timing controller board based on an Arduino Mega 2560. It has 8 inputs and 16 outputs and can be programmed as any other Arduino.

When the Fooduino receives an input signal it trigs the outputs according to the schedule depicted in figure D.15. The output trigs from the Arduino are all 200 µs long and only controls when to trig the camera and the light. The cameras exposure time is defined in the camera and the light pulse width is defined in the Gardasoft controller. The camera exposure trig signal is delayed by 50 µs to allow for some jitter and ramp up time in the Gardasoft controllers and LEDs. The delay between camera trig signals are 16.6 ms which corresponds to 60 fps.
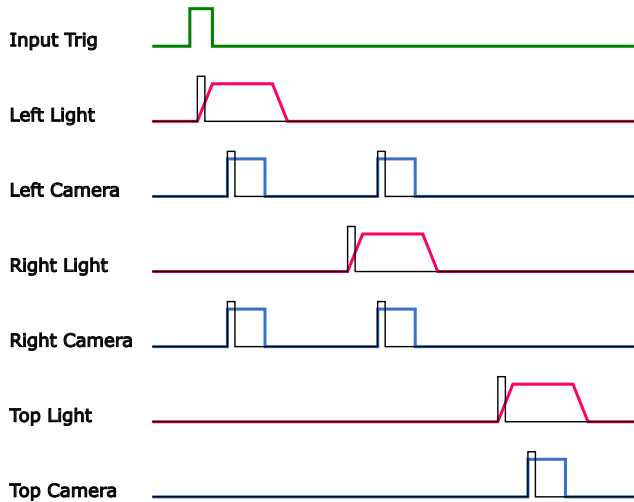
Fig. D.15: Image capture sequence controlled by the Fooduino. The black lines depict the output trigs from the Fooduino.

3.5.1 LED duty cycle

We want the exposure time to be as low as possible to minimize motion blur in the images. To do that we need as much light as possible from the LEDs. It is possible to overdrive the LEDs, without damaging them, as long as the duty cycle is low.

The cameras' exposure time is set to 750 μs. The LEDs are switched on 250 μs prior to the cameras and switched off 200 μs after to make sure the viscera is fully lit during the camera exposure. The pulse width of the LEDs are therefore 1200 μs and each LED is pulsed one time per bird as depicted in figure D.15. At 4 bps the LEDs are active for 4800 μs or 4.8 ms per second which corresponds to a duty cycle of $\approx$ 0.5%. That leaves plenty room to increase the voltage of the LEDs if the images are too dark.

3.6 Combining the hardware

A schematic of the entire system can be seen in figure D.16. The PC is the entry point for the system and where the captured images will be saved. The PLC is setup by Linco to generate trigs whenever a clamp passes in front of the cameras. This PLC also sends a time stamp and the shackle ID to the PC whenever a trig is send to the Fooduino. The shackle ID corresponds to a specific shackle on the chain and starts from 0 whenever the chain has made a full rotation.

The Fooduino can be accessed from the PC via USB2. When it receives a trig from the PLC it starts the capture sequence presented in section 3.5. This

activates the light and the cameras in the correct order. When the cameras capture an image, they are sent the PC where the images are given another time stamp for when they were received.

The length of a USB 3.0 cable is usually between 3 to 5 metres before the transfer speed deteriorates, depending on the quality. This limitation means the computer must be installed near the camera station and in the evisceration room. All hardware, besides camera, lens and light that are in the tubes, are therefore installed in a water tight case.
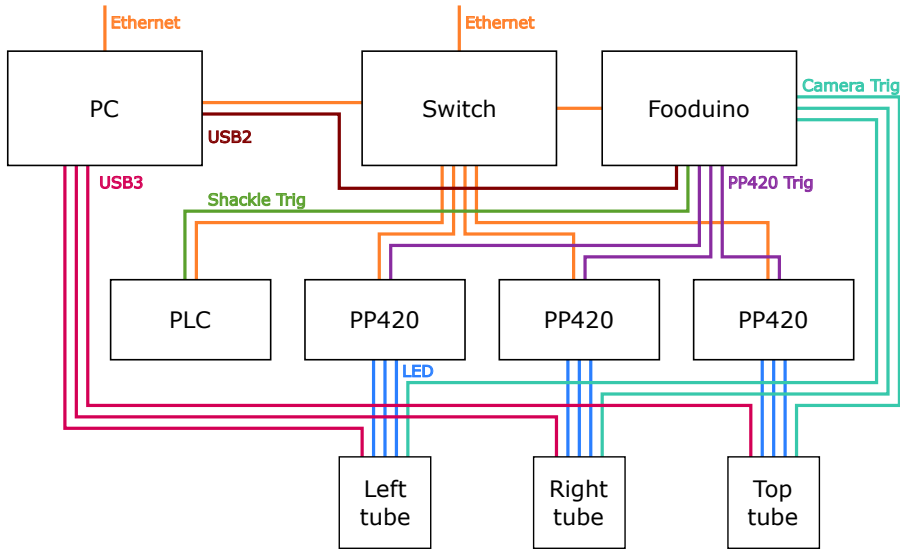
Fig. D.16: Schematics of the electrical parts and their connections in the system.

3.7 Trimming the dataset while recording

As discussed in section 2.2, should the system be able to automatically trim the data for healthy viscera sets. With 5 images per viscera, 4 times per second and 12 MP images the PC will receive 240 MB image data every second when saving RAW images. With a 1 TB SSD it will take roughly 69 minutes before it is full and only 0.5 % of the images will be unhealthy.

In order to automatically delete some of these healthy viscera sets we install an optical sensor after the viscera inspection site. The PLC controls this sensor and monitors the conveyor for empty clamps. The signals are depicted in figure D.17. When an empty shackle is detected the PLC sends a message to the PC with a time stamp and the shackle ID. The software to handle these messages are described in section 4.3.
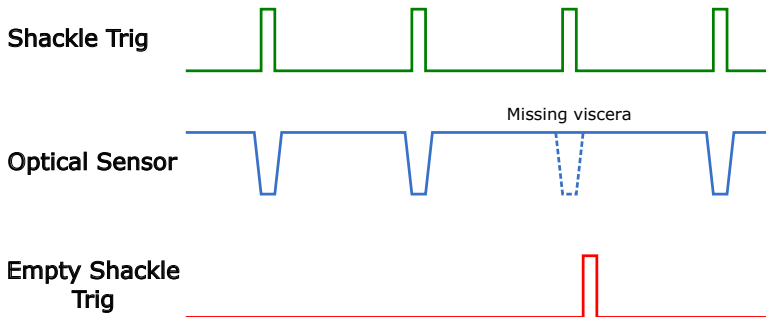
Fig. D.17: An optical sensor detects empty clamps and a trig is generated.

4 Software

This chapter describes the choices made regarding software and the programs developed for saving the images and for handling the large amount of images.

4.1 Image file names

The following name scheme is used to keep track of the recorded images. The images received from the cameras contains no meta data and the software only know when they were received and from which camera they came. This information is paired with the latest received PLC message that contains the time stamp for when the trig was received by the PLC and what the shackle ID was for that trig. Table D.8 shows an example of the file names for the five image files recorded for one viscera set.

<Camera>_<PC TS>_<PLC TS>_<Shackle ID>.raw
Left_20161115-092202-363_20161115-092458-956_000254.raw
Left_20161115-092202-415_20161115-092458-956_000254.raw
Right_20161115-092202-363_20161115-092458-956_000254.raw
Right_20161115-092202-415_20161115-092458-956_000254.raw
Top_20161115-092202-406_20161115-092458-956_000254.raw

Table D.8: File names for the five images captured from one viscera set. TS is short for time stamp.

The PLC time stamp is only updated every second so multiple viscera sets will get the same PLC_TS, but together with the shackle ID it will be unique for each viscera set. PC_TS is the time stamp for when the images arrive at the computer. The first image from two side cameras will be lit by the left

light source and the second image will be lit by the right light source. The received images are saved in three folders, one for each camera. The cameras are identified by their serial number, and thereby linked to the right folders. The PLC time seems to drift over time, so the PC time and PLC time does not align, but this is not an issue.

4.2   Image recorder

The main objective of this program is to save the images received from the three cameras. To do this it also listens for messages from the PLC. The latest shackle trig message from the PLC is kept in memory to be used in the filename when saving the images. No pre-processing are done is this program, the images are saved directly to the SSD in raw format. The program also monitors the free space on the SSD and stops the recording when it is full.

This program also receives the PLC messages sent when an empty clamp is detected. These messages are saved directly in a folder to be used by the clean-up software described in section 4.3. A snapshot of the recording software can be seen in figure D.28.

4.3   Remove unwanted images

This program monitors the recording folder and separates the images in healthy and unhealthy based on the empty shackle messages. When only saving unhealthy viscera sets, the healthy images are simply deleted when it is clear that they are not removed from the conveyor at the inspection site. The message received when an empty clamp is detected contains a PC time stamp, a PLC time stamp and a shackle id. The shackle id is the same for the empty shackle and for the images recorded for the removed viscera set. The PLC time stamp and PC time stamp are obviously different as the empty shackle is detected maybe 30 seconds after the viscera passed the cameras. The shackle id resets every time the chain has made a full rotation, but as we know roughly how many shackles there are on the conveyor between the capture point and the optical sensor, we can run back through the images and find the images captured for that empty shackle. Using the time stamps we can also verify that the time passed makes sense. We also use the time stamps to delete all images that are "unclaimed" by an empty clamp, when the images are older than a couple of minutes.

However, we might also want to save a similar amount of "healthy" birds, that is a viscera set not removed from the conveyor. The program therefore has an option to save a non-removed viscera set every minute. A snapshot of the clean-up software can be seen in figure D.29.

Some clamps are empty due to a missing broiler or if the evisceration machine fails. These will also generate an empty shackle message, but we

have no interest in saving images of empty clamps. The program therefore has another function where it scans three rows of pixels in the image and finds the maximum gradient. If the gradient is low, it indicates that we are looking at the dark background and the image is deemed empty. As the images are saved in RAW format, it can do this without loading the entire image, by simply reading a stream of bytes at three predefined positions in the image file. This makes the method very easy on both the CPU and SSD.

4.4 Captured images

The images are saved in raw bytes to avoid using CPU on lossless encoding and to make it easy to read bytes directly from the file for the empty clamp check described in section 4. For processing and viewing all images are converted into the PNG images. In that process all images are halved in both width and height as we use no interpolation of the Bayer pattern to generate the colour images. The resulting images are therefore 2048x1500 pixels. The cameras are installed sideways to better utilise the horizontal FoV. The presented images are therefore 1500x2048. Figure D.18 and D.19 shows the two images captured by the left camera and figure D.20 and D.21 shows the two images captured by the right camera. Note that the specular highlights are different between the two images. Figure D.22 shows the image captured by the top camera with a single light source.
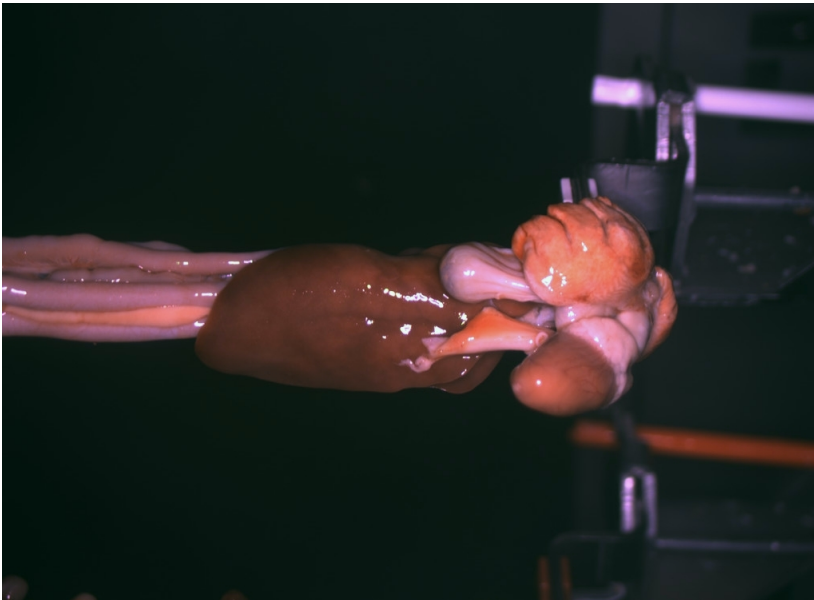

Fig. D.18: Left Camera, Left Light.

126

Fig. D.19: Left Camera, Right Light.

Fig. D.20: Right Camera, Left Light.

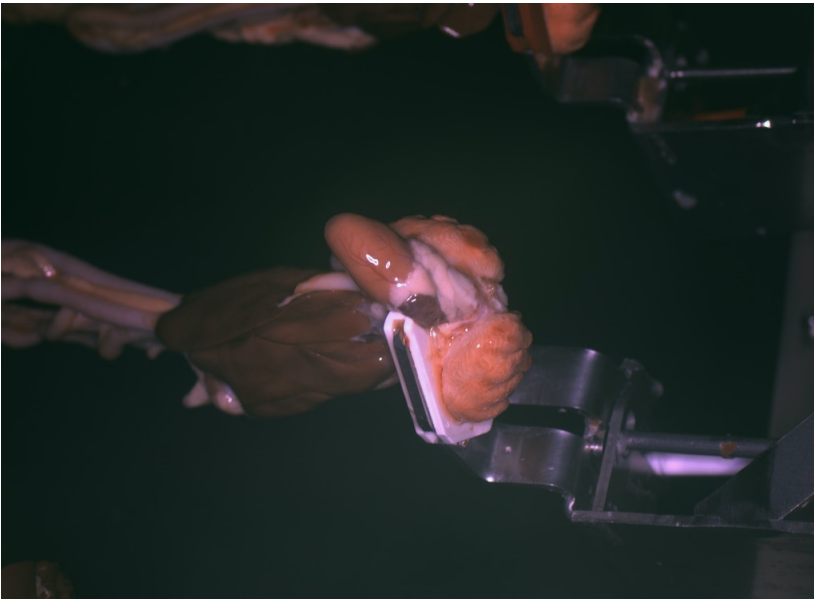**Fig. D.21:** Right Camera, Right Light.

Fig. D.22: Top Camera, Top Light.

5 Installation

Figure D.23 shows the hardware installed in a waterproof case at the Accepteance Test at Baader Linco. In the lower right corner are six connectors running to the three vision tubes. The PIC, marked with red, is passively cooled by the ribs the make of the entire case.
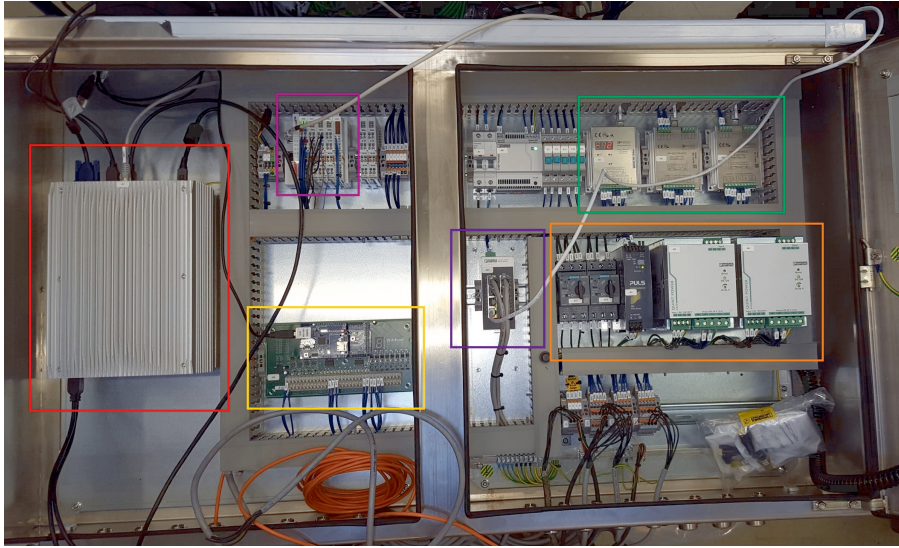

Fig. D.23: Image of the entire system mounted in a waterproof case at the test site. Red marks the computer, yellow the Fooduino, purple the PLC, blue the switch, green the Gardasofts and orange marks the power supply.

The system was installed at Danpo the XX th august 2016. The waterproof case were fastened above the conveyor at the vision location. Figure D.24 and D.25 shows the system running with a without the flash from the tubes. The orange wire in the top centre of the images is connected to the inductive sensor which triggers the cameras. Some extra shielding were put up around the setup to reduce the amount of flashing light, which could ignore the workers who occasionally are in the room during production. The system have proved able to operate automatically for weeks without manual interference.

Fig. D.24: Recording system running after installation.

Fig. D.25: Recording system running after installation mid flash.

6 Appendix

6.1 LEDs

Forward current graphs for the leds. Used to calculate the voltage and current usage at normal operation for the ring light in section 3.3.

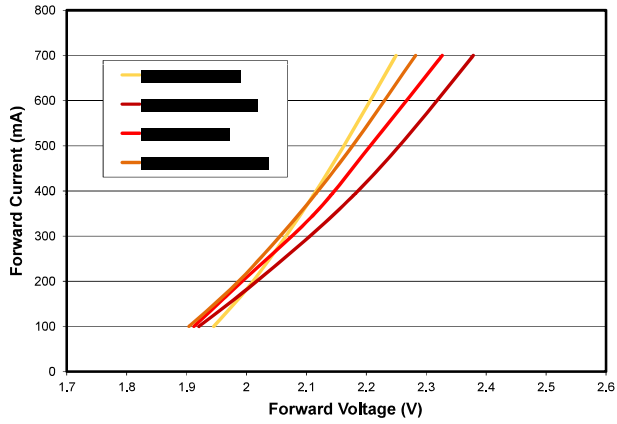Figure 4b: Typical forward current vs. forward voltage for
▬▬▬ , ▬▬▬ , ▬▬▬ and ▬▬▬ at T_j=25ºC.

Fig. D.26: Forward current per forward voltage. Image copied from datasheet [■].
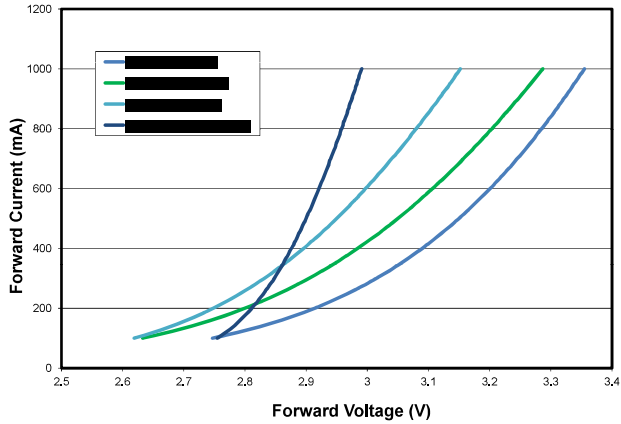
Figure 4a: Typical forward current vs. forward voltage for
▬▬▬ , ▬▬▬ , ▬▬▬ and ▬▬▬ at T_j=25ºC.

Fig. D.27: Forward current per forward voltage. Image copied from datasheet [■].

6.2   Software

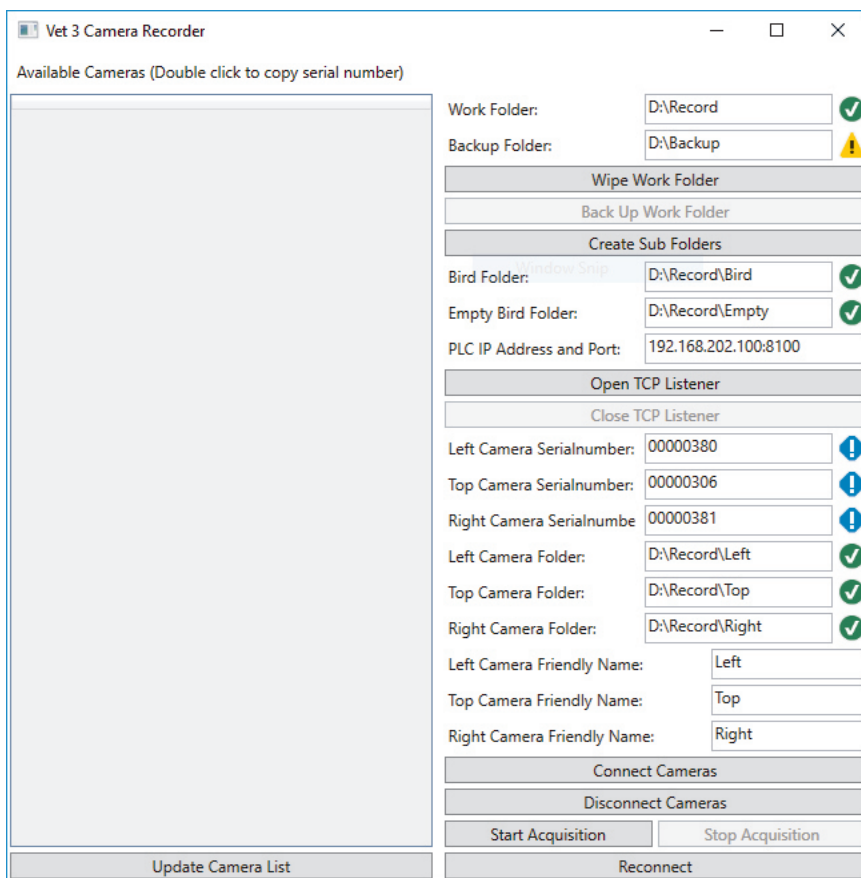Screenshots from the developed software described in section 4.

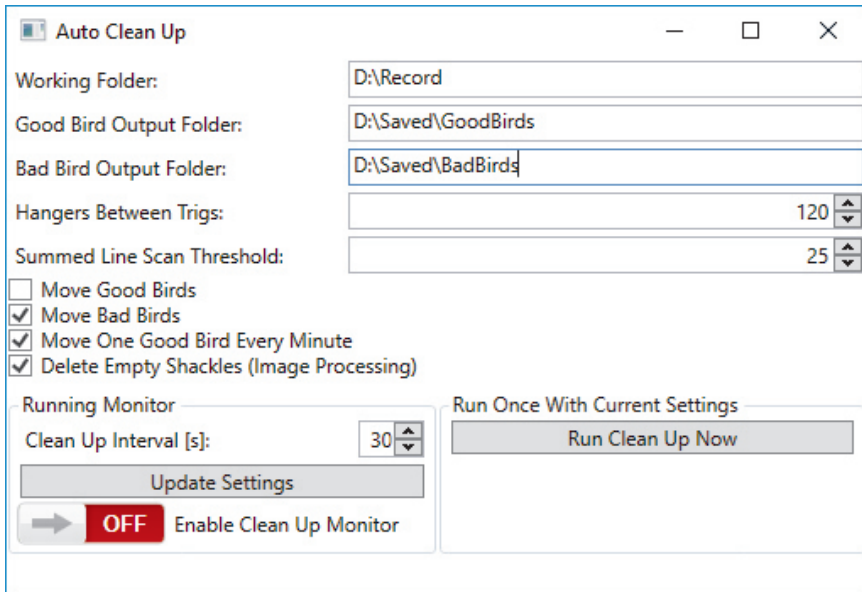**Fig. D.28:** Snapshot from the Image Recording software.

Fig. D.29: Snapshot from the Auto Clean Up software.

References

[1] EU, "REGULATION (EC) No 854/2004 OF THE EUROPEAN PARLIA-MENT AND OF THE COUNCIL of 29 April 2004 laying down specific rules for the organisation of official controls on products of animal origin intended for human consumption," *EC No 854/2004*, no. 854, p. 8, 2004.

[2] Imaging Solutions Group, "LW-AL-IMX253C-USB3," http://isgcameras.com/product/allegro-sony-imx253-color/, August 2017.

[3] Kowa, "LM25XC," https://www.kowa-lenses.com/en/applications/machine-vision/363/lm25xc, August 2017.

[4] Gardasoft Vision Ltd, "PP420 Series," http://www.gardasoft.com/Downloads/, August 2017.

[5] Blue Line, "Blue Line Fanless FPC-6000," http://www.blue-line.dk/produkter/industrial-computers-displays/fanless-pc/blue-line-fanless-fpc-6000, February 2018.

Paper E

Reaching behind Specular Highlights by Registration
of Two Images of Broiler Viscera

Anders Jørgensen, Jens Fagertun, and Thomas B. Moeslund

Abstract

The consumption of chicken is predicted to rise as a larger part of the earth's population wants protein. This calls for increased production speeds but the required post mortem inspection is becoming a bottleneck. The inspection of broiler viscera is a difficult task for computer vision, further complicated by the glossy wet organs which induce unpredictable specular highlights. This paper presents a method for producing a single image without specular highlights by combining texture information from two different images. The method achieved an average SSIM score of 0.96 over a test set of 100 samples. The result is visually pleasing images with correct textural information instead of specular highlights.

1   Introduction

In a world with increasing living standards, the demand for meat is growing. As chicken meat only requires one third of the water used to produces the same amount of bovine meat, the consumption of chicken is expected to rise even more [1]. The top 25 chicken meat producing countries have increased their production by almost 70 % from 2000 to 2013 [2]. This means more processing plants, but also increasing production rates. With line speeds already above 180 birds per minute(bpm), the task of manual health inspection is becoming difficult. All slaughtered animals must go through a post mortem inspection, required by EU legislation [3], where veterinarians inspect both the carcass and the viscera for diseases. This is a straining task that requires full concentration and their capabilities are already pushed to the limit. This call for automation of the inspection task with a computer vision system to aid the veterinarians.

Specular highlights are a well known challenge when designing vision systems, especially when capturing wet or glossy objects. Specular highlights are areas where the light source is perfectly reflected back to the camera. In RGB images specular highlights are seen as fully saturated pixels in either individual colour channels or in all channels at the same time. When pixels become fully saturated they no longer contain any useful colour information about the object.

In some cases it can be enough to detect and delete the specular highlights in software, but in other cases they can hide critical information and need to be handled in the hardware setup. A simple solution is to change the position of the light source so that its reflection does not point into the camera, but this will only work on objects with a well defined surface. Another solution is to use polarising filters in front of the light source and the camera lens, and filter out light reflected directly back from the object. But polarising filters blocks at least 50% of the light (by design), so with two filters the camera

receives less than 25% of the emitted light. This might be infeasible in setups with short exposure times, where the light source is already operating at maximum voltage.

This paper therefore presents a method that restores textural information behind specular highlights by combining two images with different lighting. The goal is to have a single image to use for further analysing. The scene object is broiler viscera captured in-line at a poultry processing plant during production. The captured viscera surfaces are wet and uneven, yielding some unpredictable specular highlights. A captured viscera set can be seen in figure E.1, where the largest organs have been outlined. Most visible diseases are located on the liver or the heart. The lower unmarked part (to the right in figure E.1) is the intestines.
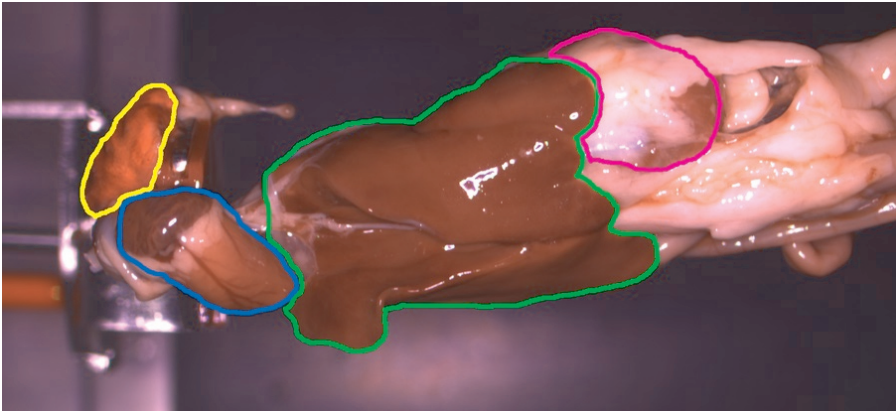

Fig. E.1: Example of a captured viscera set. Green marks the livers, blue the heart, yellow a lung and purple the gizzard. Rotated left to fit paper.

Figure E.2 shows two of the most common liver diseases found by veterinarians. In figure E.2a the broiler is diagnosed with perihepatitis, where the peritoneum get attached the the liver. It is seen as a white membrane on the surface of the liver. Figure E.2b shows a liver with necrotic hepatitis. The liver get spotted areas scattered over various parts of the liver. It is clear in this case that the specular highlights could hide texture information crucial for a correct classification of the liver.

(a) Perihepatitis.
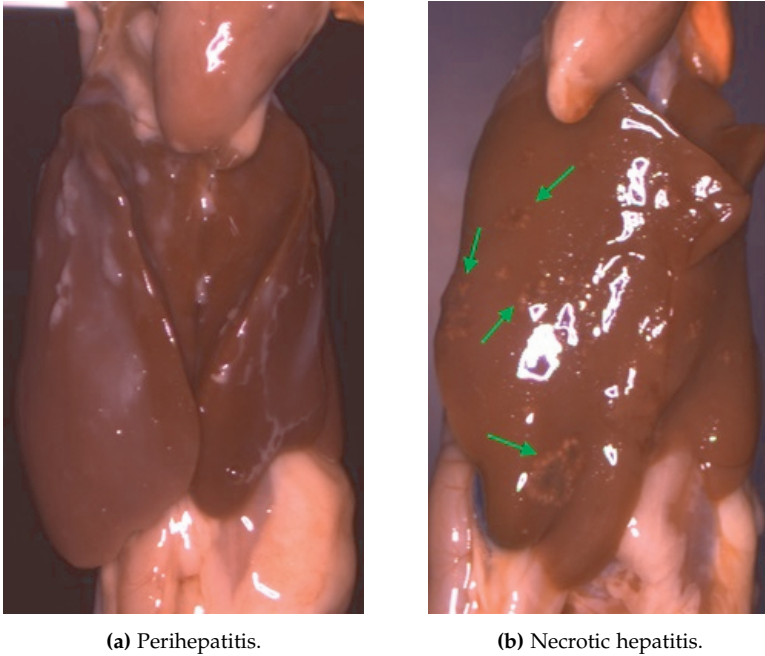
(b) Necrotic hepatitis.

Fig. E.2: Examples of two common liver diseases.

2 Related work

Computer vision inspection of poultry is not a new topic, however most focus have been on inspection of the carcass. A survey by [4] showed the majority of the work was done in detection of skin tumours [5, 6] and faecal contamination [7, 8].

Broiler viscera have received less attention. Early work by [9] used UV and colour imaging to detect splenomegaly in images of viscera. UV was helpful in segmenting the spleen which appear to have a similar colour as the liver in RGB images. [10] used multi-spectral imaging to classify chicken hearts into four diseases and normal. By selecting four spectral bands from the visible spectrum they achieved classification rates between 84 % and 100 % for the five classes. Through visible and near infra-red spectroscopy [11] could separate chicken livers into septicaemic and normal birds with 94 % and 98 % accuracy.

More recently [12] proposed a method for segmenting heart, liver and lung in RGB-D images of broiler viscera. [13] used a neural network to classify four liver diagnoses in RGB images and gained an overall true positive rate of 77.6 %.

Many different methods exist for the detection of specular highlights. Sat-

urated pixels can be detected by simple intensity thresholding, but more elaborate methods are needed to find specular highlights that are not saturated. [14] proposed a fast method for separating the diffuse and specular components of a single image by using a normalised version of the dichromatic model. Using a physics based approach [15] proposed a method that depends on the Fresnel term of the specular highlight for detection of said highlights.

Saturated specular highlights hide the texture information of the area beneath it. Restoration methods must either inpaint saturated areas or retrieve the data from another image. [16] developed a hardware setup with four flashes connected to the same camera. By combining the four images they are able to generate an image with reduced specular effects. [17] demonstrates a method where a video sequence was used to generate a single image without specular highlights. Using SIFT features the images were registered and for each image set the minimum value for each overlapping pixel were selected. A similar idea was used by [18] who proposed a method for generating a single high quality image of paintings photographed from multiple angles. By getting the median of the gradient from multiple images they use the Poisson equation to recover an image without highlights.

Combining colour values from multiple images, in which the object moves, requires a registration. Image registration or alignment generally consists of the following steps: feature detection, feature matching, model estimation and image resampling [19]. SIFT [20] are often used features for finding correspondence points or its speeded up version SURF [21]. [22] have trained a neural network to find and select features and another neural network is trained to generate descriptors.

3 Data set

The dataset consists of images of broiler viscera captured in-line at a poultry processing plant during normal operation. Two images are captured per viscera set. The first image is captured with the light shifted to the left, indirect light (IL), and the second image with the light directly from the camera's position, direct light(DL). While this gives different specular highlights between the images, it also means that the right side of the IL image is poorly illuminated. The DL image is the more visually pleasing image with a uniform illumination. The light sources are both ring shaped and consists of red, green and blue LEDs. Generally this generate a white light but some specular highlights appear in one colour channel only, due to the spatial distance between the LEDs.

The viscera is transported at ≈ 0.6 m/s on a conveyor belt and the two consecutive images are captured ≈ 17 ms apart, which mean the viscera moves

≈1 cm between the captured images. The exposure time is 750 μs to avoid motion blur. The image resolution is 1500x2048. An example image set can be seen in figure E.3.

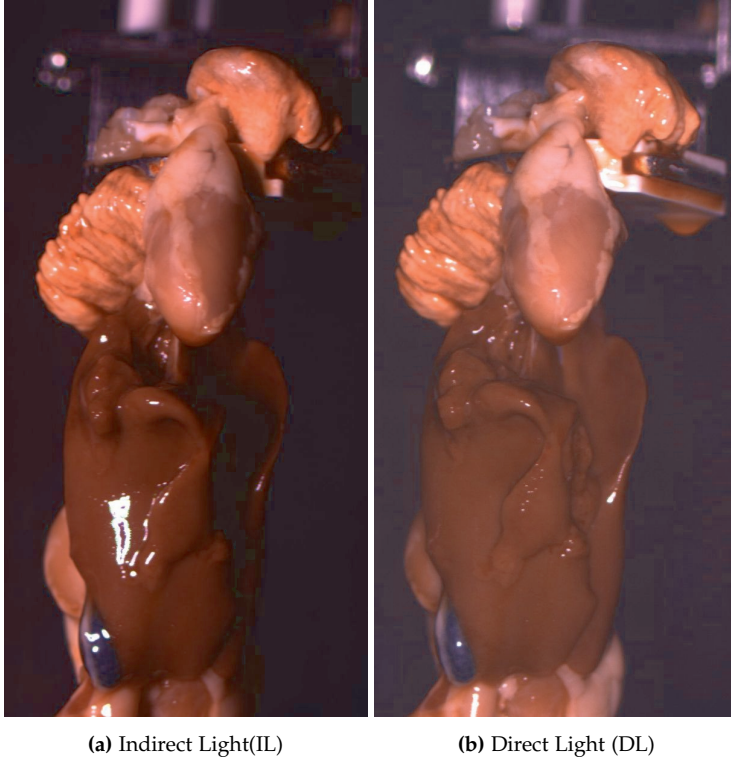(a) Indirect Light(IL)   (b) Direct Light (DL)

Fig. E.3: Example of the two images captured per viscera set. The images have been cropped to better present the important organs.

4 Method

As the most frequent diseases are present on the liver and heart these will be the focus of this method. The entire viscera set is useful for registration but the colour transformation will only be performed on the liver and heart.

The first step is to detect the specular highlights to remove. This is done in the DL image as the viscera is more evenly illuminated from this camera's point of view and therefore the image to use for further processing. The next step is to register the images. This happens automatically with out the need for manual registration. The full IL image is resampled to match the DL image. The last step is the colour transformation. As the lighting differs

greatly in each image this is done individually for each processed highlight. The pipeline is outlined in figure E.4.

Fig. E.4: Pipeline describing the individual steps in the system.

4.1 Specular highlight detection

Specular highlights are detected via a rule based approach. The viscera parts are primarily red with a low response in the green spectrum so pixels with a high response in the green channel are marked as specular highlights. The other rules are shown in E.1. The resulting mask is 255 in pixels where specular highlights are detected, otherwise 0.

$$
\begin{aligned}
r_{x,y} = g_{x,y} = b_{x,y} = 255 & \quad \text{if} \quad g_{x,y} > 175 \\
r_{x,y} = g_{x,y} = b_{x,y} = 255 & \quad \text{if} \quad b_{x,y} \geq r_{x,y} \ \text{and} \ b_{x,y} > 200 \\
r_{x,y} = g_{x,y} = b_{x,y} = 255 & \quad \text{if} \quad g_{x,y} \geq r_{x,y} \ \text{and} \ g_{x,y} \geq b_{x,y} \\
r_{x,y} = g_{x,y} = b_{x,y} = 0 & \quad \text{else}
\end{aligned}
\tag{E.1}
$$

4.2 Image registration

The viscera are secured with a clamp by the gullet, just behind the lungs and heart. When the viscera moves along the conveyor belt, the vibrations can cause the organs to move independently as they are connected by deformable soft tissue.

In order to investigate the motion of the viscera between the two consecutive images, control points have manually been annotated in 15 viscera image sets, where anatomical landmarks were visible.

Figure E.5 shows the control points vertical position (y) plotted with respect to the control point's difference in its x position (Δx) between the two images. Linear regression have been used to fit a line to the control points for each set.
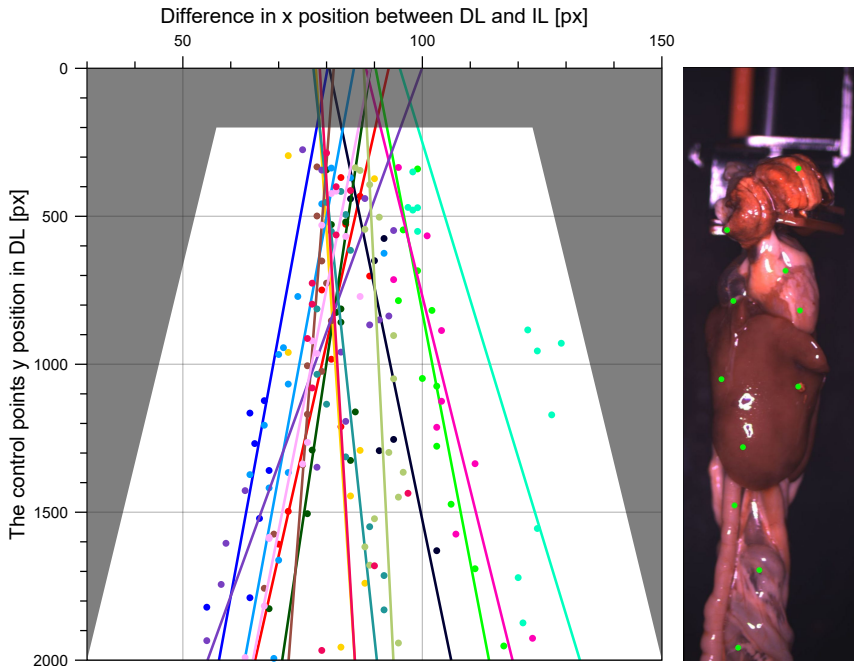
Fig. E.5: Manually selected control points plotted as the y position per difference in x position. The grey area indicates invalid matches as described in section 4.2. The green control points are also plotted in the image to the right.

Figure E.5 indicates that there is a correlation between Δx and y. The variation in Δx is smallest around $y = 500$, which corresponds to the location of the clamp. As y increases, so does the variation in Δx, yet control points from the same set deviates to the same side. This could indicate that the viscera swings back and forth in the clamp, like a pendulum, when the viscera moves along the conveyor belt. This was also observed at the processing plant. When a viscera set swings to the right the horizontal translation between the images will decrease and vice versa, as the conveyor belt moves to the left seen from the camera.

The difference in the y position (Δy) of the control points with respect to y in DL was also investigated. For all control points Δy was less than 10 pixels and there were no indication of any correlation with y. This means that the viscera does not swing high enough between the two captured images to generate a change in y that is greater than the vibrations from the conveyor belt.

For automatic control point detection is was chosen to use SURF [21] and the resulting matches for a single image set can be seen in figure E.6a. Most matches are found around the heart or at the intestines with only a few matches found on the liver. For SURF we have used a blob response

threshold of 700 and the number of octaves is 3.

Prior knowledge about the transform, illustrated in figure E.5, makes it possible to prune matches by looking at their direction. The grey area surrounding the control points in figure E.5 indicates the area where matches are no longer considered valid. For all matches y must be larger than y_{min} as the upper part of the image only contains a part of the hanger which is located in a plane further away from the camera than the viscera set. Matches here will have a very different offset, Δx, compared to the rest of the matches. Matches must also have a Δy less than Δy_{max}, otherwise they will be discarded. Δx_{min} and Δx_{max} are calculated from y as per the following formulas:

$$\Delta x_{min}(y) = -0.015y + 60 \tag{E.2}$$
$$\Delta x_{max}(y) = 0.015y + 120 \tag{E.3}$$

The pruned matches can be seen in figure E.6b. A total of 35 matches remain. Most are found on the lung and heart, the rest on the intestines and no matches are present on the liver. With so few features it is not possible to apply non-rigid registration methods [23]. Through testing it was found that the best compromise between precision and robustness was an affine transformation. It has the freedom to describe the major transformations between the two images and few enough parameters to be found almost consistently.

In some cases the remaining matches are too few and/or too inaccurate. Or the transformation fails because the local transformations are too large to be modelled with an affine transformation. A high amount of shear or rotation indicate a faulty transformation as this should not occur because the viscera are transported in a horizontal direction. This can be detected by inspecting the transformation matrix. The shear and rotation components (index 1,2 and 2,1) should not deviate more than ΔT_{rs} from 0 as this indicates a high amount of shear or rotation. In these cases it was chosen to fall back to a simple horizontal translation. Based on the control points plotted in figure E.5, the fixed translation have been set to $t_h = 85$ pixels.

Figure E.7 presents a successful registration as a checker board switching between the two images. It is clear that large variation in the intensity makes it difficult to extract matching SURF features. Especially around the heart, which is very dark on the right side (here top) due to it being the shady side.
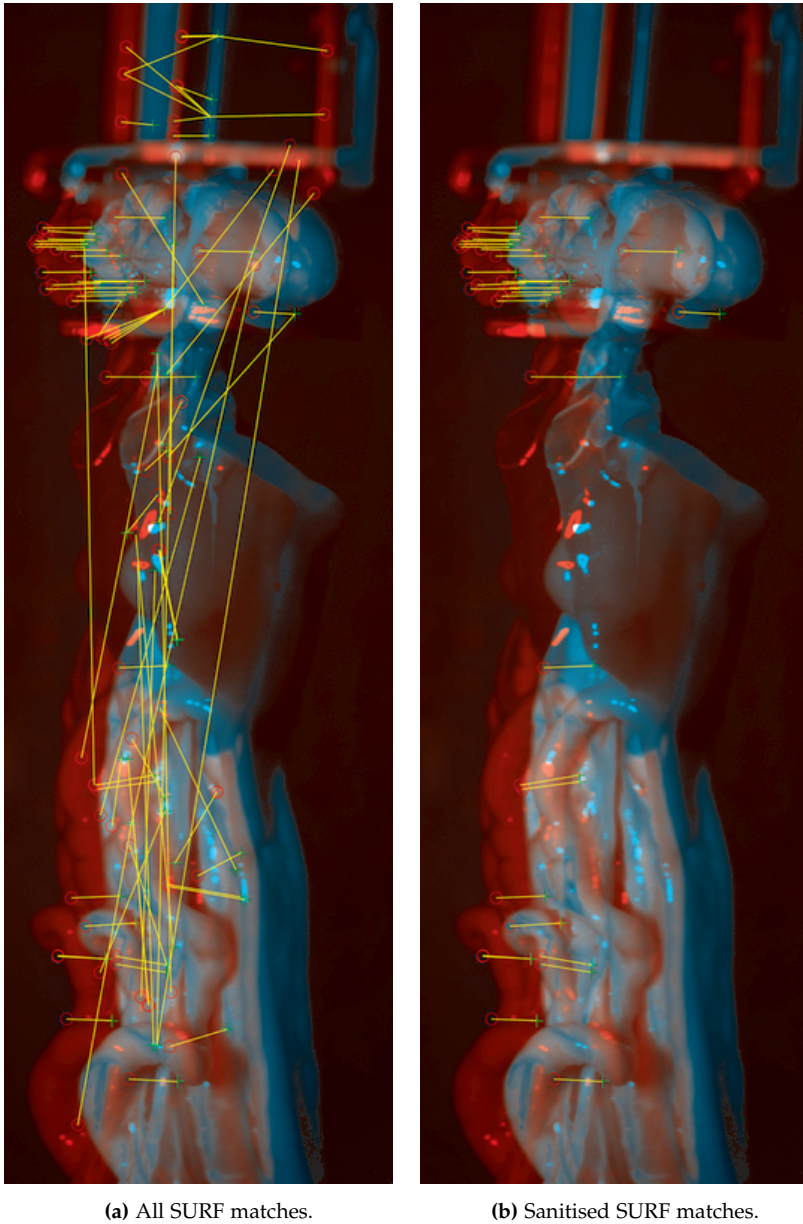
(a) All SURF matches.

(b) Sanitised SURF matches.

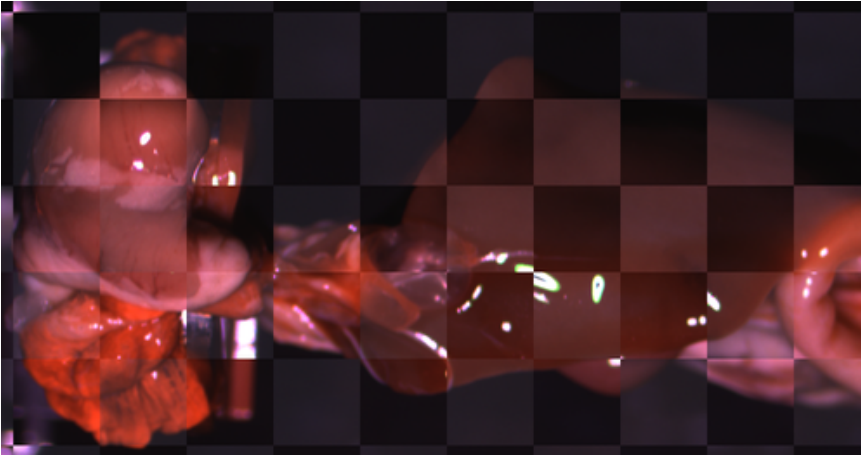**Fig. E.6:** SURF matches found in the two images.

Fig. E.7: Checker board of two registered images. It is easy to see the large variation in the intensity between the two images. The image is cropped and rotated to better fit the page.

4.3 Colour transformation

The colour transformation steps are illustrated in figure E.8. Other highlight removal methods use the minimum or median pixel value from multiple images to get their results. But using the minimum value will in this case result in a poorly illuminated image. The median value does not make sense for only two values as it would either be the minimum or the maximum value. Instead it is chosen to match the colours by finding the difference in colour in the area around the highlight by modelling it with a quadratic polynomial surface. The size of the area depends on the size of the specular highlight. A similar concept have been applied to greyscale images by [24] to stitch atomic force microscope images. The details of each step showed in figure E.8 will be described in the following.
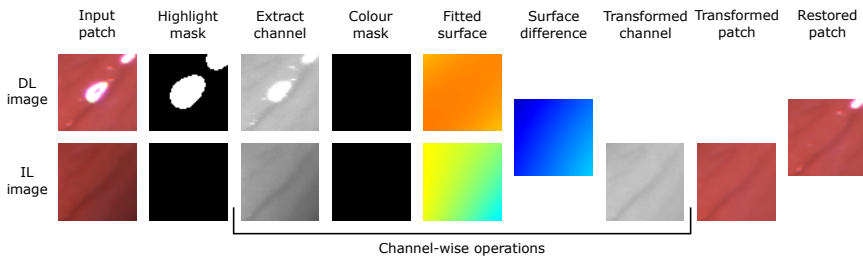

Fig. E.8: Steps in the colour transformation. Read from left to right.

Blob extraction is used on the highlight mask to process the highlights individually. Specular highlights with an area smaller than $A_{\min}$ are ignored.

Each blob is dilated with a circular kernel with a radius of r_{dk} to avoid arte-facts from colour blooming around the specular highlights.

The dilated blob is extracted from the images DL and IL with a b_b wide border around the blob's region of interest (ROI). The extracted image patches can contain other specular highlights so these are found and masked in the same way as described in section 4.1 and dilated the same way.

Each colour channel is extracted and smoothed with a Gaussian kernel with a standard deviation of σ. If the highlight is located in an area next to another organ or in an area with a shade, the abrupt change in intensity could result in a non-linearity that is poorly approximated by a polynomial surface. A mask is therefore created that filters pixels that differs from the central region by an intensity change ΔI larger than ΔI_{max}. As the central region in DL contains a highlight, the reference value for DL is calculated from the pixels surrounding the highlight. The result is a colour mask that is white were ΔI is larger than ΔI_{max}, otherwise black.

The highlight mask and the colour mask are used to exclude values when fitting the polynomial surface to the colour channel intensities. The surface is fitted using least squares. The resulting surfaces are depicted in figure E.8, where blue indicates low values and red indicates high values. The surface difference, in this particular example, shows that the lower right corner of the ROI is more different in intensity than the upper left, which can be confirmed by looking at the input images. The difference is added to the IL ROI and the three colour channels are combined to a new IL ROI image. Bear in mind that the surface difference can be negative when the IL ROI has a higher intensity than DL ROI. As the last step pixels from the new IL ROI are used to fill the highlight blob in the original image.

5 Results

The results are divided in two parts, qualitative and quantitative. The method has been tested on viscera image sets with specular highlights. As there are no ground truth for these images, the result will be evaluated subjectively. The method will also be tested on image sets with a manually drawn high-light mask. Here the results will be compared to the original image with difference maps.

The following settings have been used for the experiments:

$$t_h = 85, \quad b_b = 20, \quad \Delta I_{max} = 80, \quad \sigma = 5, \quad r_{dk} = 5$$
$$y_{min} = 200, \quad \Delta y_{max} = 10, \quad A_{min} = 10, \quad \Delta T_{rs} = 0.05$$

5.1 Qualitative Results

An example of the restoration results can be seen in figure E.9. Overall it looks good, the large highlights have been removed and pixels on the liver and heart have been restored.
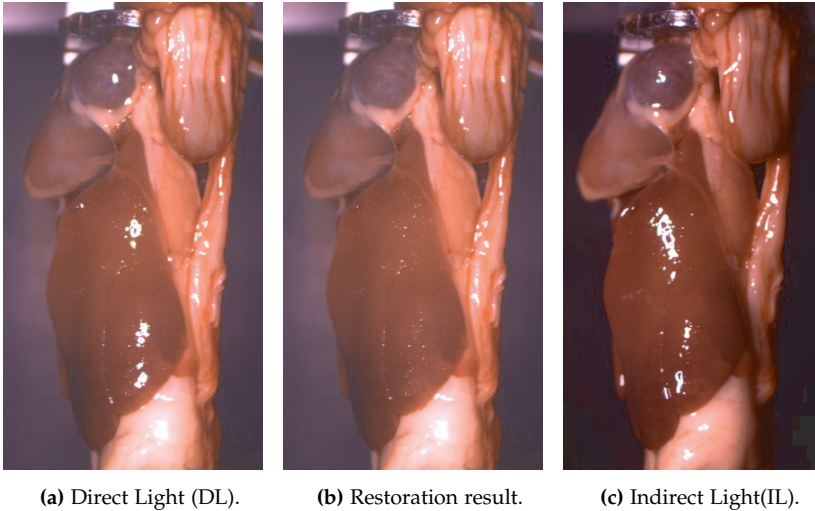

(a) Direct Light (DL).   **(b)** Restoration result.   **(c)** Indirect Light(IL).

Fig. E.9: Restoration results in (b), flanked by the two input images in (a) and (c).

Figure E.10 displays a close up of the restoration on a liver with necrotic hepatitis. The liver is speckled with bright areas, which are visual clues for the diagnosis. The textural information is correctly transferred from image IL into the area in image DL with specular highlights.

(a) Direct Light (DL).   **(b)** Restoration result.   **(c)** Indirect Light(IL).
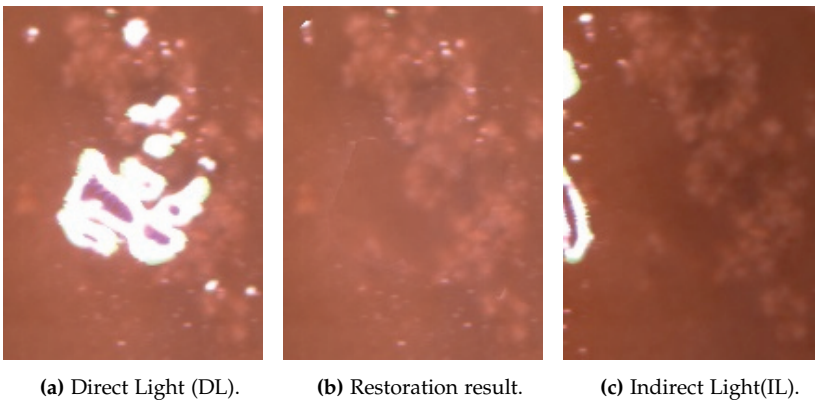
Fig. E.10: Close up of a liver with necrotic hepatitis. Textural information is correctly imported to the area with specular highlights.

Figure E.11 shows a close up of a healthy liver. The liver has some dark areas, which are damages caused during the evisceration. These damages are no concern for food safety, but is considered a quality issue. The dark areas are easily visible in figure E.11b, after the restoration.
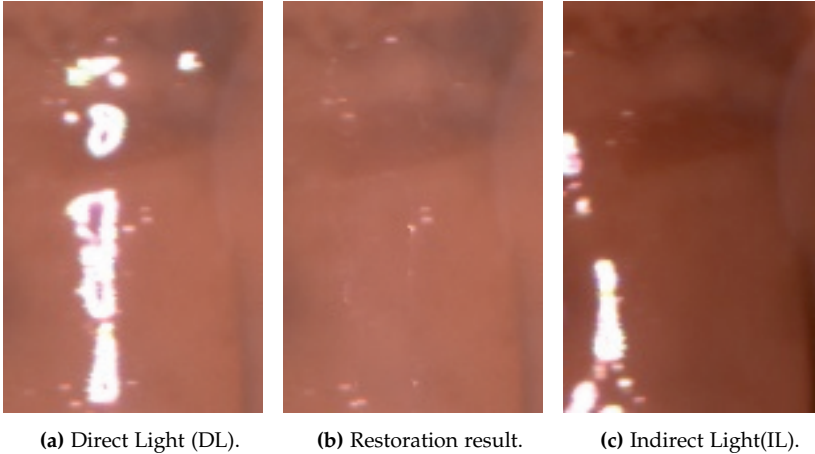

(a) Direct Light (DL).   **(b)** Restoration result.   **(c)** Indirect Light(IL).

Fig. E.11: Close up of the restoration of a healthy liver. The dark areas are easily visible in figure b.

In figure E.12 is the results of the restoration of highlights in the heart. The veins makes easy to spot if the registration is incorrect, but for this example it is judged correct.

(a) Direct Light (DL).   **(b)** Restoration result.   **(c)** Indirect Light(IL).
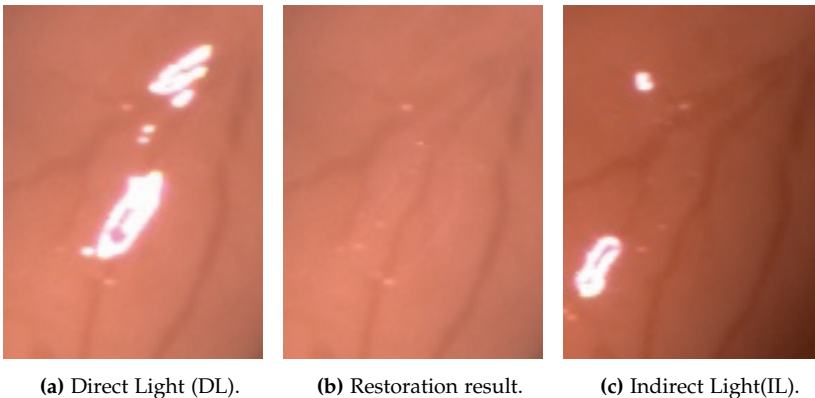
Fig. E.12: Restoration on the heart. The veins indicate that the registration is correct.

Figure E.13 shows a cut out from another processed viscera set. An issue becomes apparent when looking at the two highlights marked by green in

image E.13a. The texture information behind these highlights should be imported from image E.13c, but this area is underexposed and contains little to no information. This means that the method is basically raising noise to fit the colour in the DL image.
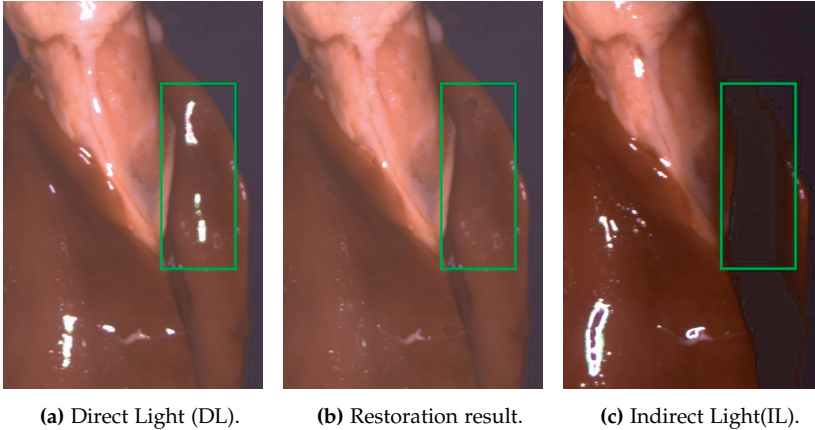

(a) Direct Light (DL).   (b) Restoration result.   (c) Indirect Light(IL).

Fig. E.13: Restoration result for heart and liver. The highlights marked by the green square in figure E.13a requires data from a poorly lit area in figure E.13c.

5.2 Quantitative Results

A total of 100 highlights were manually drawn in 30 viscera image sets. The results of the restorations are quantified with Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM) scores, [25]. Both scores are often used to evaluate the performance of inpaint methods or similar methods where the aim is to restore missing data, [26–28]. The scores are plotted in figure E.14. Mean PSNR is 41.80 ± 5.46 and mean SSIM is 0.96 ± 0.03.
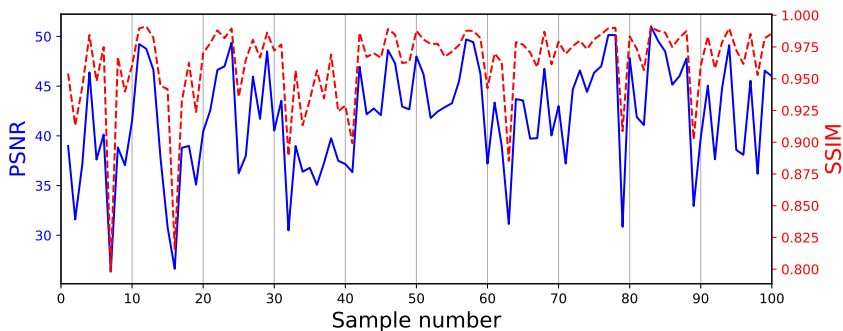

Fig. E.14: PSNR and SSIM scores for the 100 manually drawn highlights.

5. Results

Table E.1 shows the results for a few selected specular highlights. The PSNR and SSIM scores can be compared visually with the resulting image. By looking at sample 7 it is clear that a bad registration will have a negative impact on the result as the restored pixels comes from another area of the image. The inflammation on the liver in sample 9 becomes smaller in the result image due to an incorrect registration, which could be critical for the classification of the liver. Other highlights, like samples 21 and 12 are both registered and colour transformed correctly and requires close inspection to detect that pixels have been imported.
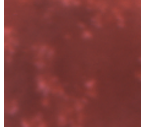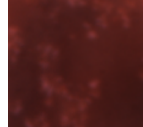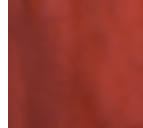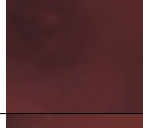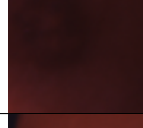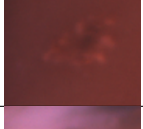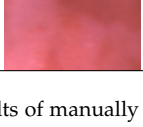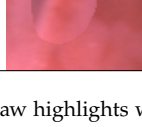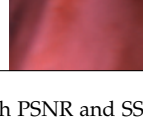
#	Mask	DL input	Result	IL input	PSNR	SSIM
21					42.57	0.98
3					37.14	0.94
12					48.75	0.99
9					37.05	0.94
79					30.86	0.91
7					26.36	0.80

Table E.1: Results of manually draw highlights with PSNR and SSIM Scores.

6   Conclusion

The proposed method aims to restore pixels behind specular highlights by combining two images into one image without highlights. The method showed promising results on a dataset that is far from trivial. In quantitative testing the method scored a mean PSNR of 41.80 ± 5.46 and a mean SSIM of 0.96 ± 0.03. Equally important did it produce visually pleasing images even when zooming in on the restored areas.

The livers contain very few features which together with the different light conditions make the registration a challenge. As the viscera rotates and wobbles as it moves along the conveyor belt there are clear cases were the results would benefit from a non-rigid registration. However as the images are captured just 17 ms apart, the affine transformation is enough to handle most situations, as the results indicate.

References

[1] M. Mekonnen and A. Hoekstara, "The green, blue and grey water footprint of farm animals and animal products," *Unesco-Ihe*, vol. 1, no. 48, p. 80, 2010.

[2] International Poultry Council, "Chicken meat production for top chicken producing countries," 2013.

[3] EU, "REGULATION (EC) No 854/2004," p. 8, 2004.

[4] Z. Xiong, A. Xie, D.-W. Sun, X.-A. Zeng, and D. Liu, "Applications of Hyperspectral Imaging in Chicken Meat Safety and Quality Detection and Evaluation: A Review," 2014.

[5] C. Xu, I. Kim, and M. S. Kim, "Poultry Skin Tumor Detection in Hyperspectral Reflectance Images by Combining Classifiers," pp. 1289–1296, 2007.

[6] K. Chao, C. C. Yang, and M. S. Kim, "Spectral line-scan imaging system for high-speed non-destructive wholesomeness inspection of broilers," *Trends in Food Science and Technology*, vol. 21, no. 3, pp. 129–137, 2010.

[7] S. C. Yoon, B. Park, K. C. Lawrence, W. R. Windham, and G. W. Heitschmidt, "Line-scan hyperspectral imaging system for real-time inspection of poultry carcasses with fecal material and ingesta," *Computers and Electronics in Agriculture*, vol. 79, no. 2, pp. 159–168, 2011. [Online]. Available: http://dx.doi.org/10.1016/j.compag.2011.09.008

[8] G. W. Heitschmidt, B. Park, K. C. Lawrence, W. R. Windham, and D. P. Smith, "Improved Hyperspectral Imaging System for Fecal Detection on Poultry Carcasses," *Transactions of the ASABE*, vol. 50, no. 4, pp. 1427–1432, 2007. [Online]. Available: https: //naldc.nal.usda.gov/download/12864/PDF

[9] Y. Tao, J. Shao, K. Skeeles, and Y. R. Chen, "Detection of splenomegaly in poultry carcasses by UV and color imaging," *Transactions of the Asae*, vol. 43, no. 2, pp. 469–474, 2000.

[10] K. Chao, Y. . R. Chen, W. R. Hruschka, and B. Park, "Chicken Heart Disease Characterization by Multi–spectral Imaging," *Applied Engineering in Agriculture American Society of Agricultural Engineers*, vol. 99, no. 171, pp. 99–106, 2001. [Online]. Available: https://pdfs. semanticscholar.org/41ca/cc5d4928663935a7523c8320e925cde9f604.pdf

[11] C. Hsieh, Y. R. Chen, B. P. Dey, and D. E. Chan, "Separating Septicemic and Normal Chicken Livers by Visible/Near–infrared Spectroscopy and Back–propagation Neural Networks," *Transactions of the ASAE*, vol. 45, no. 2, pp. 459–469, 2002. [Online]. Available: https://pdfs. semanticscholar.org/14f8/a35f6976796d981d6b6045f0d150bf47c8a7.pdf

[12] M. P. Philipsen, A. Jørgensen, S. Escalera, and T. B. Moeslund, "RGB-D Segmentation of Poultry Entrails," *TIX International Conference on Articulated Motion and Deformable Objects*, vol. 9756, pp. 168– 174, 2016. [Online]. Available: http://link.springer.com/10.1007/ 978-3-319-41778-3_17

[13] A. Jørgensen, J. Fagertun, and T. B. Moeslund, "Diagnosis of Broiler Livers by Classifying Image Patches." Springer, Cham, jun 2017, pp. 374–385. [Online]. Available: http://link.springer.com/10.1007/ 978-3-319-59126-1_31

[14] D. An, J. Suo, X. Ji, H. Wang, and Q. Dai, "Fast and High Quality Highlight Removal from A Single Image," pp. 1–11, 2015. [Online]. Available: http://arxiv.org/abs/1512.00237

[15] E. Angelopoulou, "Specular Highlight Detection Based on the Fresnel Reflection Coefficient," in *Proceedings of the IEEE International Conference on Computer Vision*, 2007.

[16] R. Feris, R. Raskar, K. H. Tan, and M. Turk, "Specular highlights detection and reduction with multi-flash photography," *Journal of the Brazilian Computer Society*, vol. 12, no. 1, pp. 35–42, 2006.

[17] S. M. Z. A. Shah, S. Marshall, and P. Murray, "Removal of specular reflections from image sequences using feature correspondences," *Machine Vision and Applications*, vol. 28, no. 3, pp. 409–420, 2017.

[18] A. Buades, G. Haro, and E. Meinhardt-Llopis, "Obtaining High Quality Photographs of Paintings by Image Fusion," *Image Processing on Line*, vol. 5, pp. 159–175, 2015.

[19] S. Saxena and R. K. Singh, "A Survey of Recent and Classical Image Registration Methods," *International Journal of Signal Processing Image Processing and Pattern Recognition*, vol. 7, no. 4, pp. 167–176, 2014.

[20] D. Lowe, "Object recognition from local scale-invariant features," *Proceedings of the Seventh IEEE International Conference on Computer Vision*, pp. 1150–1157 vol.2, 1999. [Online]. Available: http://ieeexplore. ieee.org/document/790410/

[21] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, jun 2008. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1077314207001555

[22] H. Altwaijry, A. Veit, and S. Belongie, "Learning to detect and match keypoints with deep architectures," in *British Machine Vision Conference (BMVC)*, York, UK, 2016. [Online]. Available: http://vision.cornell. edu/se3/wp-content/uploads/2016/08/learning-detect-match.pdf

[23] L. Zagorchev and A. Goshtasby, "A comparative study of transformation functions for nonrigid image registration." *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 15, no. 3, pp. 529–538, 2006.

[24] M. Vestergaard, S. Bengtson, M. Pedersen, C. Rankl, and T. B. Moeslund, "Stitching Grid-wise Atomic Force Microscope Images," *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, vol. 3, no. Visigrapp, pp. 110–117, 2016. [Online]. Available: http://www.scitepress.org/DigitalLibrary/ Link.aspx?doi=10.5220/0005716501100117

[25] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 13, no. 4, 2004. [Online]. Available: https://ece.uwaterloo.ca/~z70wang/publications/ ssim.pdf

[26] S. Zhang, L. Jiao, F. Liu, and S. Wang, "Global Low-Rank Image Restoration With Gaussian Mixture Model," *IEEE Transactions on Cybernetics*, pp. 1–12, 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7959620/

[27] S. Arumugaperumal, B. Sivagami, and K. PazhaniKumar, "An advanced scratch removal method for Fingerprint biometrics," in *2011 3rd International Conference on Electronics Computer Technology*. IEEE, apr 2011, pp. 196–200. [Online]. Available: http://ieeexplore.ieee.org/document/5941886/

[28] S. Trambadia and H. Mayatra, "Image inpainting based on Discrete Wavelet Transform (DWT) technique," in *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*. IEEE, nov 2016, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/7916794/

References

Paper F

Diagnosis of Broiler Livers by Classifying Image Patches

Anders Jørgensen, Jens Fagertun, and Thomas B. Moeslund

Abstract

The manual health inspection are becoming the bottleneck at poultry processing plants. We present a computer vision method for automatic diagnosis of broiler livers. The non-rigid livers, of varying shape and sizes, are classified in patches by a convolutional neural network, outputting maps with probabilities of the three most common diseases. A Random Forest classifier combines the maps to a single diagnosis. The method classifies 77.6% livers correctly in a problem that is far from trivial.

1   Introduction

To stay competitive with low salary countries, poultry processing plants in the western world, strive to increase their slaughter rates. With line speeds already over $180bpm$ (birds per minute), the manual health inspection is now becoming the bottleneck. EU legislation require its member states to perform a post mortem inspection [1] and though no minimum inspection time is specified, it calls for enough time to do a proper inspection. On top of that, local laws in some countries require up to three veterinarians per slaughter line [2].

Post mortem inspection includes both carcass and viscera. If either shows visual signs of a disease both must be discarded. In Sweden and other European countries [2], a mirror is added to the inspection site to allow the inspectors to see the opposite side of the viscera and carcasses. But this also effectively doubles the number of interest points that the inspectors must focus on, making their job twice as hard.

In Denmark, a normal flock has about 1% of the birds discarded due to diseases [3]. Spotting that one bird out of 100 can be a straining task and the inspectors are therefore required to have frequent breaks. This calls for an automated computer vision system to aid the veterinarians and help the processing plants to stay competitive. Such a system could automatically identify unhealthy viscera and discard them. This paper focuses on broiler livers and proposes a method for classifying the three most common liver diseases.

2   Related work

Computer vision inspection of poultry is a relatively new scientific field, whereas food inspection in general is more mature [4, 5]. Recent work by Panagou et al. [6] displays a non-destructive technique for determining the microbiological quality of beef. Elmasry et al. [7] used hyperspectral near

infrared imaging for predicting pH and tenderness and proved it could be an alternative to traditional measuring methods, though it will require some effort to move the system in-line. Feng et al. [8] developed a method for looking at bacteria by measuring the Total Viable Count in chicken breast fillets with a hyperspectral imaging system.

Others have looked at physical defects, like detecting anomalies on poultry carcases as described in this review by Xiong et al. [9]. One such anomaly can be skin tumours, which Xu et al. [10] detects by employing Principal Component Analysis, Discrete Wavelet Transform and Kernel Discriminant Analysis and combining the classification results. Nakariyakul et al. [11] proposed a new Adaptive Branch and Bound algorithm for fast feature selection also when detecting poultry skin tumours. They found their method to be roughly 10 times faster that other Brand and Bound methods.

Speed is an important factor when inspecting poultry at the slaughter lines in a processing plant. Chao et al. [12] presented an in-line system for identifying wholesome and unwholesome chicken. Tests were conducted at a line speed of $140bpm$, yet they concluded that their method should be capable of running at $200bpm$ also. Yoon et al. [13] used a hyperspectral line-scan camera for detecting fecal matter and ingesta on poultry carcases. The system was tested at $140bpm$ and $180bpm$, where the latter was deemed the upper limit for the system.

When inspecting viscera the first challenge is to detect and segment the organs. Viscera do not have the same rigidity as the carcass and are often transported in a clamp on a conveyor belt, causing it to swing and rotate as it moves along. Work by Amaral et al. [14] proposed a weighted atlas auto-context algorithm to segment pig viscera in RGB images. Similarly Philipsen et al. [15] proposed a method for segmenting heart, lung and liver in poultry viscera using RGB-D images. The extra dimension proved to give a small performance boost over RGB, but also increased the complexity of the image recording.

3 Data set

The data set, recorded for this work, consists of images of poultry viscera captured in-line just after evisceration at a poultry processing plant. The RGB images' resolution is 2048x1500.

The focus of this paper is the broiler liver, which is the organ causing the largest number of rejects [3]. In collaboration with a veterinarian working in the field, a total of $1,476$ images have been graded into the categories listed in Tab. F.1. An example image from each class can be seen in Fig. F.1. A healthy (**H**) liver has a smooth surface, but can still vary in colour between different shades of red/brown. Cobblestone (**C**) typically affects the entire

Data set		Data set after split	
Liver diagnosis	# Total	# Patch Classifier	# Overall
Healthy (**H**)	542	240	302
Cobblestone Liver (**C**)	458	232	226
Perihepatitis (**P**)	270	125	145
Necrotic Hepatitis (**N**)	206	115	91

Table F.1: Poultry viscera in dataset. "Patch Classifier" shows the number of images used for the patch extraction. "Overall" indicates that the images are used in evaluation of the entire pipeline.

liver where it looks like dark trenches on the surface. Perihepatitis (**P**) and Necrotic Hepatitis (**N**) can affect just smaller areas of the liver. With **P** the peritoneum gets infected and attaches it self to the liver. **N** appears as small spots which can affect the entire liver or only a small part.

The liver is segmented by the method developed by Philipsen et al. [15]. For this work the method has been adapted to RGB images instead of RGB-D images. This yields some decent but imperfect results, as it can be seen in Fig. F.1, where the green line indicates the border of the segmentation. Some of the heart might be segmented as liver or sometimes the white peritoneum fails to be segmented with the liver. Improving the segmentation method is not within the scope of this article.
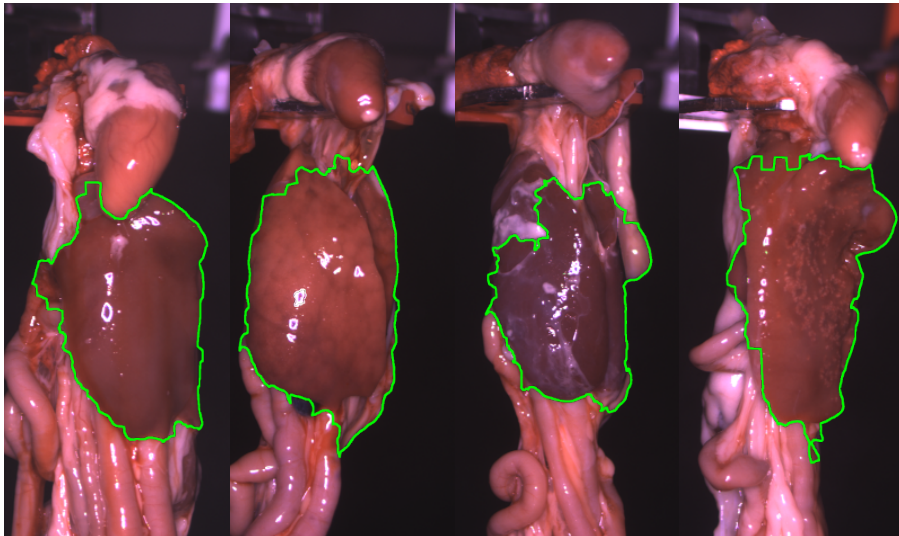
Fig. F.1: The four liver diagnoses. From left to right: Healthy liver, Cobblestone liver, Perihepatitis, Necrotic Hepatitis. The green border marks the segmentation found by the segmentation algorithm. The third liver from the left is not fully segmented in the top left part.

3.1 In-class variation.

The grading is discrete in the sense that a liver belongs to one and only one of the four classes, and there is no severity grading within each class. This is the way the viscera is currently being graded by veterinarians at the processing plant. As a result there is a large in-class variation as it can be seen in Fig. F.2, where all livers are diagnosed with **C**. The red/brown tone of the liver can also change widely between broilers. All livers in Fig. F.2 could be healthy if judged by colour alone and not texture.
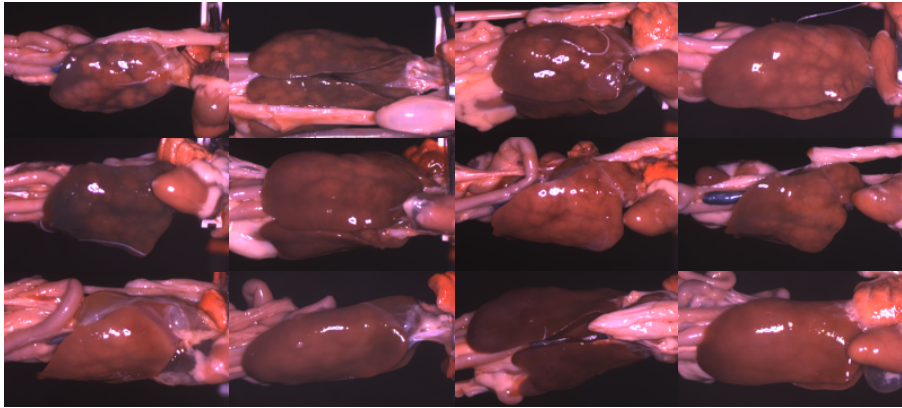
Fig. F.2: Variation in the class **C**. Top left is the most severe case, bottom right is the least severe.

4 Method

Because a disease does not necessarily cover the entire liver it is often the case that different areas of the same liver can be classified differently. The presented method works by finding areas affected by a disease with a patch classifier and based on these results classify the entire liver.

The overall classification consists of the following steps. First a highlight mask is created as described in Sect. 4.1. A convolutional neural network (CNN) trained on image patches constructs a probability map for each class, see Sect. 4.2. The maps are combined with a weight mask, see Sect. 4.3, made from the liver segmentation and the highlight mask. Features are extracted from the weighted maps and fed to a Random Forest classifier giving the overall prediction for the liver, see Sect. 4.4. An overview of the pipeline can be seen in Fig. F.3.
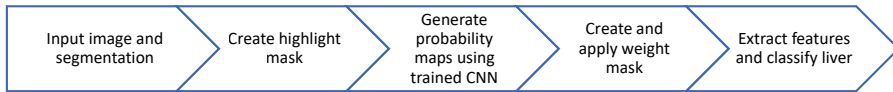
Fig. F.3: Overview of the method pipeline.

4.1 Highlight detection

The images are captured in-line at a poultry processing plant. When a set of viscera passes the camera, led diodes flash at high intensity to generate enough light to illuminate the viscera and enable the camera to use a low exposure time, in order to limit motion blur. As a result this generates some unwanted highlights on the wet glossy surfaces of the liver. Three examples can be seen in Fig. F.4. In some images the highlights also appears as speckles. This happens because the evisceration machine in some cases forms small bubbles on the viscera's surface as they are extracted from the carcass.
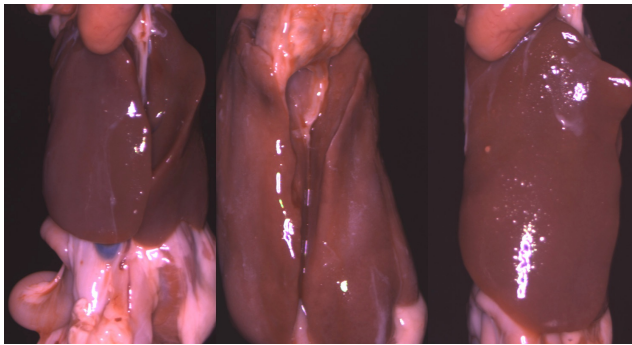
Fig. F.4: Examples of livers with highlights. In the image to the right some highlights appear as speckles.

The colour of pixels affected by highlights depends more on the light source than on the object they are portraying. These pixels are therefore considered as noisy and must be removed before further analysis.

The light source consists of red, green and blue LEDs. Because of the spatial distance between the LEDs some highlights occur only as a result of the reflection from one colour e.g. the green. Figure F.5 shows an example, and the results of the highlight removal.

Fig. F.5: Highlights can occur individually in all colour channels. The 2nd and 4th image show the results of the highlight removal.

The following rules are used to create the highlight mask, which is 0 in highlight areas, otherwise 255:

$$r_{x,y} = g_{x,y} = b_{x,y} = 0 \quad \text{if} \quad g_{x,y} > 175$$
$$r_{x,y} = g_{x,y} = b_{x,y} = 0 \quad \text{if} \quad b_{x,y} \geq r_{x,y} \ \text{and} \ b_{x,y} > 200$$
$$r_{x,y} = g_{x,y} = b_{x,y} = 0 \quad \text{if} \quad g_{x,y} \geq r_{x,y} \ \text{and} \ g_{x,y} > b_{x,y}$$
$$r_{x,y} = g_{x,y} = b_{x,y} = 255 \quad \text{else}$$

Bright green does not naturally occur in the viscera and any high values of green must therefore be caused by a reflection. The blue colour only appears in darker areas, so bright blue pixels are also discarded as highlight. The threshold values have been found through empirical studies.

4.2 Convolutional neural network patch classifier

It has been chosen to classify the input images in patches. The livers are non-rigid bodies that varies in shape and size and are unevenly affected by the diseases. This makes it difficult to extract the same features for each liver. Dividing the liver into patches allows for classes with less variation than full image livers and each patch will have the same shape and size. Similar approaches have been employed in [16] and [17], with good results. The patches are used in a CNN which will act as the patch classifier. The CNN will find the best features for discriminating the four classes.

A set of patches have been manually extracted from each class, to train and evaluate the CNN. An extra set of patches, miscellaneous (**M**), have also been extracted. This class contains areas that cannot be represented by the other classes. This may be other organs like the heart or intestines, that mistakenly have been included in the segmentation. It can also be post mortem bruises or damages to the liver that obscure the underlying class. All patches measures 75x75 pixels. This resolution was found to be large enough to cover the texture patterns of all diseases, and small enough to be easy placeable within the livers. Table F.2 shows some examples of the extracted patches.

To avoid data cross contamination, a set of images have been randomly selected for each class from the full data set, as shown in Tab. F.1. These images will only be used for training and evaluation of the CNN. Table F.3 shows the count of extracted patches. The number of images used for extraction varies based on the size and severity of the disease. For **C**, **P** and **N**, it has not been possible to extract the same amount of patches as for **H**. Patches for the class **M** have been extracted from images from the other four classes.
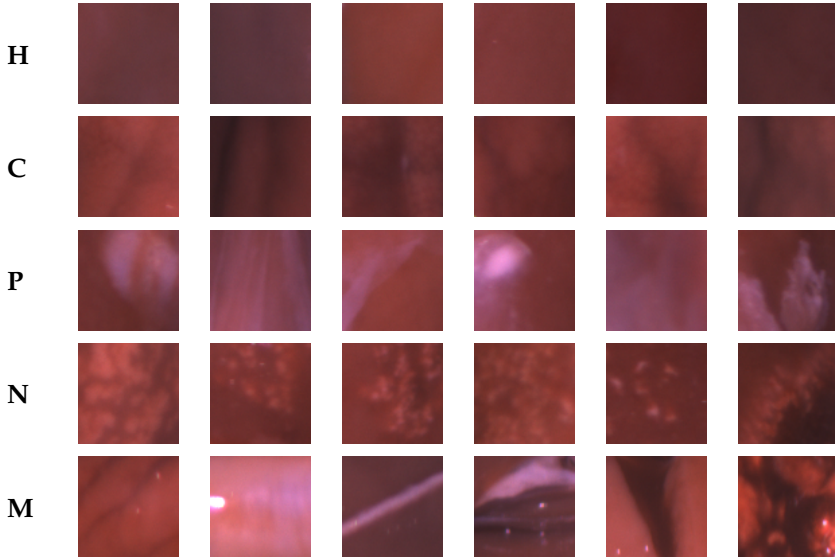
Table F.2: Six examples from each extracted class for the patch classifier.

Diag.	Full Images		Extracted Patches		
	Train/Val.	Test	Train	Val.	Test
H	180	60	306 (11016)	75 (2700)	100
C	152	80	235 (8460)	75 (2700)	100
P	89	36	205 (7380)	75 (2700)	100
N	77	38	200 (7200)	75 (2700)	100
M	-	-	323 (11628)	75 (2700)	100

Table F.3: Patches extracted from the full images. Test patches are extracted from the Test images. **M** patches have been extracted from images from all other classes. The number in the parentheses is with augmentation.

The CNN is implemented as a small sequential network with Keras [18] and Theano [19]. A representation of the network, as it is implemented in

Keras, can be seen in Fig. F.6. A softmax activation layer is used to generate the output, which is a vector with normalised probabilities for each class. The code is available at `https://bitbucket.org/andjor/2017_broiler_livers`.

| Conv2D(32,(3,3)) leakyReLU MaxPool(2,2) | Conv2D(32,(3,3)) leakyReLU MaxPool(2,2) | Conv2D(64,(3,3)) leakyReLU MaxPool(2,2) | Dense(64) leakyReLU Dropout(0.5) | Dense(5) Activation (softmax) |
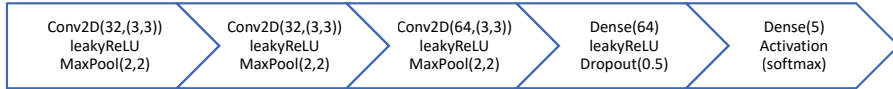
Fig. F.6: Structure of the CNN, as it is implemented in Keras [18]. More information about the layers is present in the Keras documentation.

All training and validation patches are augmented with 36 rotations in increments of $10°$. The validation set is used to pick the epoch with the best model. Looking at the validation loss in Fig. F.7, it can be seen that the smallest validation loss happened at epoch 24.
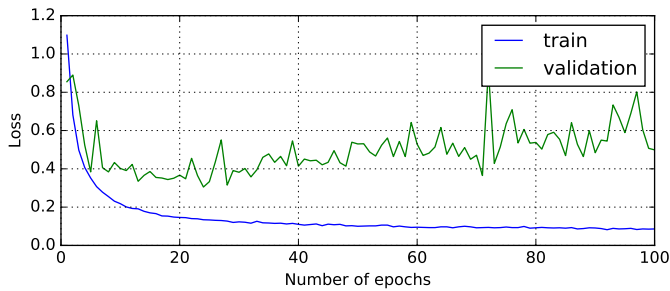
Fig. F.7: Training and validation loss for the CNN. A snapshot of the model is saved at the smallest validation loss, which is epoch 24.

4.2.1 Generating probability maps.

The CNN is used to predict patches sampled from the input images. The bounding box from the segmentation is used as region of interest in the input image. To further reduce the calculation time, the CNN is only applied to every n'th pixel in both the x and y direction, effectively down-sampling the generated probability maps by a factor of n compared to the original. A map will be generated for each class, where each element holds the probability that the corresponding region belongs to that class.

4.3 Creating the weight mask

As the patch classifier classifies pixels based on a surrounding square region it is important to justify how trustworthy that region is. Some pixels in the

region can be outside the liver mask or the region can contain a high number of highlight pixels. To take this into account a weight mask is created by anding the liver segmentation mask and the highlight mask. The weights are then calculated by moving a sliding window, 75x75 pixels, across the joint mask with the same step size used in the patch classifier. For each step the weight, α, of the region is calculated as

$$\alpha = \begin{cases} 1 - \frac{b}{b_{\max}} & \text{if } b < b_{\max} \\ 0 & \text{else} \end{cases} \tag{F.1}$$

$$b_{\max} = 0.1 \cdot w \cdot h$$

where b is the number of black pixels (highlight or not liver) in the patch and w and h is the width and height of the patch. If 10% or more of the patch is black, the weight is set to zero.

An example of the generated probability maps, with the weight mask applied, can be seen in Fig. F.8.
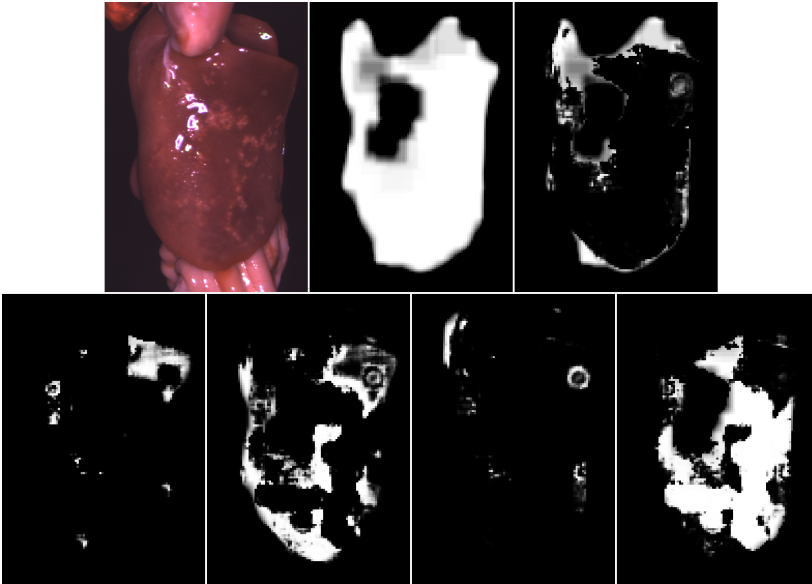
Fig. F.8: Probability maps for an input image diagnosed with **N**. Input image top left. Weight mask is top middle. Top right is the probability map for **M**. In the bottom from the left to the right: **H**, **C**, **P** and **N**.

4.4 Random forest liver classifier

The classification of the liver is based on the weighted probability maps. As it can be seen in Fig. F.8, will a liver often respond in all four classes and the

M class. This can be because of misclassified patches or simply because the disease only affects a small part of the liver.

For a data driven approach, features are extracted from each probability map. These are features like the count of non zero probabilities in the map, the sum of all probabilities or the ratio of non zero probabilities compared to the map size. A total of 8 features are extracted per probability map, 40 features per liver. All are fed to a Random Forest Classifier [20], responsible for the liver classification. Random Forest is chosen as it works well with non normalised features. The number of trees is 100 and the minimum split size is 1.

5   Results

The performance of the CNN patch classifier, calculated on the test patches, can be seen in Tab. F.4. Each class contains 100 sample patches. The total true positive rate is 95.4%.

		Prediction					True
		H	**C**	**P**	**N**	**M**	Positive
	H	98	2	0	0	0	98.0%
	C	5	91	2	2	0	91.0%
True	**P**	0	1	98	0	1	98.0%
	N	1	0	0	98	1	98.0%
	M	1	4	3	0	92	92.0%

Table F.4: Classification matrix for the CNN. 100 sample patches in each class.

For the results of the entire method the step size n is set to 5, which means the algorithm will moving 5 pixels when generating the probability maps and the weight mask. This was chosen as a good balance between speed up and resolution of the probability maps. The patches are still highly correlated as only 6.6̄% pixels will be "new data" when moving the classifier one step in the horizontal direction.

The images in the "Overall" column in Tab. F.1 was used to evaluate the performance over the entire pipeline in Fig. F.3. To measure both accuracy and precision, the classification rates are measured over 10 five fold runs, using stratified sampling. The average true positive rate is 77.6%. Tab. F.5 shows the confusion matrix average and one standard deviation for the 10 five fold runs.

		Prediction				True
		H	**C**	**P**	**N**	Positive
	H	242.5 ± 3.47	27.2 ± 1.93	32.3 ± 2.71	0.0 ± 0.00	80.3%
	C	37.1 ± 3.57	176.3 ± 4.37	10.1 ± 1.66	2.5 ± 0.71	78.0%
True	**P**	34.1 ± 3.28	7.9 ± 0.74	103.0 ± 3.06	0.0 ± 0.00	71.0%
	N	10.1 ± 0.57	8.6 ± 0.52	1.0 ± 0.47	71.3 ± 0.48	78.4%

Table F.5: Classification matrix for the overall system. Average $\pm$ one standard deviation.

The processing time per liver image is $3 - 5s$ depending on the size of the liver. The tests were performed on a laptop with Intel i7-4720HQ CPU, 16GB ram and a Nvidia 960m GPU.

6 Conclusion

The patch classifier correctly classifies 95.4% of the test patches, see Tab. F.4. **C** has the most misclassifications, many fall in **H**, which is consistent with the fact that **C** slowly evolves from **H**, as shown in Fig. F.2. Many early stage **C** livers will be confused with **H** livers. **M** has the second most misclassifications, primarily spread between **C** and **P**. **C** might, in severe cases, look like bruises, and **P** can be all white much like the intestines.

A good performance of the patch classifier does not directly translate to a good performance of the overall system. The patches were hand picked and it is more important that the patches fully represent the variation in their classes than getting a high performance on very specialised classes. It is the authors' beliefs that **H** is well represented, but **C**, **P** and **N** could make use of more samples. It was not possible to extract more patches for this work, without leaving too few images for the overall performance test. For future works it might help to use other augmentation techniques on the training data.

The overall system had an average classification error of 22.4%. Some of these errors are no doubt a result of the patch classifier, not classifying patches correctly. In cases where diseases like **P** and **N** only affects a small part of the liver, it makes the probability maps very susceptible to misclassifications. This makes it difficult for the Random Forest classifier to set an appropriate threshold. Some of this could be remedied with better features for the probability maps.

It is not known how much influence the imperfect segmentation have on the classification results. Manually annotating all livers would be a slow process and one could argue that the current segmentation is closer to what it would be in a real scenario. Yet to isolate the true performance of this method, a perfect segmentation is needed.

Some errors are due to the large variation in the classes. The livers in Fig. F.9 are all diagnosed with **P**, but they look very different. For future work it might be beneficial to categorise the diagnoses into smaller subclasses. This could be helped by having a more fine-grained grading that includes the severity of the diagnoses. With this, better features for the probability maps and more training data it should be possible to improve the performance of this method.
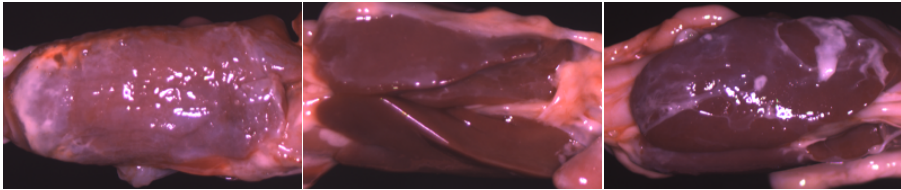
Fig. F.9: All livers are diagnosed with **P**, yet they look very different.

References

[1] EU, "REGULATION (EC) No 854/2004 : laying down specific rules for the organisation of official controls on products of animal origin intended for human consumption," p. 8, 2004.

[2] U. Löhren, "Overview on current practices of poultry slaughtering and poultry meat inspection," pp. 1–58, 2012.

[3] C. Claudi-magnussen and H. Daugaard, "Automatiserede hjælpeværktøjer til kødkontrol på kyllingeslagterierne," Danish Meat Research Institute, Tech. Rep., 2011.

[4] B. Park and R. Lu, Eds., *Hyperspectral Imaging Technology in Food and Agriculture*, ser. Food Engineering Series. New York, NY: Springer New York, 2015.

[5] H. Huang, L. Liu, and M. O. Ngadi, "Recent developments in hyperspectral imaging for assessment of food quality and safety." *Sensors (Basel, Switzerland)*, vol. 14, no. 4, pp. 7248–7276, 2014.

[6] E. Z. Panagou, O. Papadopoulou, J. M. Carstensen, and G. J. E. Nychas, "Potential of multispectral imaging technology for rapid and non-destructive determination of the microbiological quality of beef filets during aerobic storage," *International Journal of Food Microbiology*, vol. 174, pp. 1–11, 2014.

[7] G. Elmasry, D. W. Sun, and P. Allen, "Near-infrared hyperspectral imaging for predicting colour, pH and tenderness of fresh beef," *Journal of Food Engineering*, vol. 110, no. 1, pp. 127–140, 2012.

[8] Y. Z. Feng and D. W. Sun, "Determination of total viable count (TVC) in chicken breast fillets by near-infrared hyperspectral imaging and spectroscopic transforms," *Talanta*, vol. 105, pp. 244–249, 2013.

[9] Z. Xiong, A. Xie, D.-W. Sun, X.-A. Zeng, and D. Liu, "Applications of Hyperspectral Imaging in Chicken Meat Safety and Quality Detection and Evaluation: A Review," 2014.

[10] C. Xu, I. Kim, and M. S. Kim, "Poultry Skin Tumor Detection in Hyperspectral Reflectance Images by Combining Classifiers," pp. 1289–1296, 2007.

[11] S. Nakariyakul and D. P. Casasent, "Fast feature selection algorithm for poultry skin tumor detection in hyperspectral data," *Journal of Food Engineering*, vol. 94, no. 3, pp. 358–365, 2009.

[12] K. Chao, C. C. Yang, and M. S. Kim, "Line-scan spectral imaging system for online poultry carcass inspection," *Journal of Food Process Engineering*, vol. 34, no. 2011, pp. 125–143, 2011.

[13] S. C. Yoon, B. Park, K. C. Lawrence, W. R. Windham, and G. W. Heitschmidt, "Line-scan hyperspectral imaging system for real-time inspection of poultry carcasses with fecal material and ingesta," *Computers and Electronics in Agriculture*, vol. 79, no. 2, pp. 159–168, 2011.

[14] T. Amaral, I. Kyriazakis, S. J. Mckenna, and T. Ploetz, "Weighted atlas auto-context with application to multiple organ segmentation," *Proc. WACV*, 2016.

[15] M. P. Philipsen, A. Jørgensen, S. Escalera, and T. B. Moeslund, "RGB-D Segmentation of Poultry Entrails." Springer International Publishing, 2016, pp. 168–174.

[16] L. Hou, D. Samaras, T. Kurc, and Y. Gao, "Patch-based Convolutional Neural Network for Whole Slide Tissue Image Classification," p. 7, 2015.

[17] Y. Xu, Z. Jia, Y. Ai, F. Zhang, M. Lai, and E. I-Chao Chang, "Deep Convolutional Activation Features for Large Scale Brain Tumor Histopathology Image Classification and Segmentation," pp. 947–951, 2015.

[18] F. Chollet, "Keras," https://github.com/fchollet/keras, 2015.

[19] Theano Development Team, "Theano: A Python framework for fast computation of mathematical expressions," *arXiv e-prints*, vol. abs/1605.02688, 2016.

[20] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

Paper G

Graded Data and Rerun of Results from Paper F

Anders Jørgensen

The report has been used internally at IHFood

Abstract

A data set of 3984 images was graded by three veterinarians, to obtain ground truth. Grading diseases can be difficult and very subjective, so the three veterinarians all graded the same data. Analysis of the graded data set showed a great difference in the grades, which emphasises the care that should be taken when evaluating an automatic classification system.

The data was used to rerun the results for paper F. With the new data set the classification error is now down to 14.4 % from 22.4 %.

1 Obtaining ground truth

We had veterinarians to grade the images for obtaining ground truth. Organic material is difficult to grade and personal experience and stances can lead to fluctuations between graders. Two veterinarians have many years of experience from the poultry industry, while the third has experience from other parts of the sector.

To obtain a high level of certainty we had three veterinarians grade the same 3984 viscera sets from the images recorded with the camera system developed in appendix D. A workshop was held with the purpose of aligning the veterinarians and to teach them to use the grading software. The grading was then done individually over the following one and a half month.

We asked them to categorise the viscera in five liver diseases and one heart disease. These were agreed upon with one of the veterinarians at an earlier meeting. These diseases are the most frequently occurring in the production.

Disease name	Description
Necrotic hepatitis	Coloured spots or speckles on the liver.
Perihepatitis	The peritoneum gets attached to the liver which gets a white tint.
Cobblestone liver	The liver gets a cobblestone-like surface.
White liver	The liver gets very bright due to extra fat in the liver.
Liver changes	Shape and/or colour changes in the liver.
Heart inflammation	Any kind of inflammation in and around the heart or pericardium

Table G.1: Disease classes used in the grading software. Examples can be seen in figure G.1.

For each disease there is 4 levels of occurrence: No, Low, Medium and High. These levels were used to grade the severity of the disease based on area covered and intensity. The purpose is to construct the cost function of a future classification system on the number of levels separating the misclassi-

fied samples from the ground truth. Predicting a *High occurrence* disease as *No occurrence* should be more costly for the classifier than predicting a *Low occurrence* as *No occurrence*.

Besides these six disease categories were four auxiliary classes. These are for situations where the viscera doesn't fit in any of the defined disease categories and are all true or false, see table G.2.

Class name	Description
Nonsense	Does not make sense to grade the viscera. E.g. it is missing or shredded by the machinery.
Manual Inspection	Maybe sick, maybe healthy. They would be able to tell if they held it in their hands. A way to say that this cannot be graded from an image.
Unknown	Should be removed, but don't know what it is, and manual inspection wouldn't help.
Other	Sick with some other known disease or condition.

**Table G.2:** Auxiliary classes used in the grading software.

It is also possible to write a comment for all viscera sets. If the occurrence is *No* for all diseases and all auxiliary classes are false, then the viscera is regarded as healthy.

2 Grades

The results of the grading is analysed by looking at the grades for the three veterinarians (from here on referred to as vets) and by looking at the samples where they agree. The vets are referred to individually as Vet 1, Vet 2 and Vet 3, where Vet 2 and 3 have the most experience from the poultry industry.

Figure G.2 shows the distribution of grades in four main groups. Either the viscera is healthy, sick, should go to manual inspection or nonsense. It is clear that Vet 1 has graded the highest amount of viscera sick and fewest healthy. Vet 2 uses manual inspection and nonsense more than the other two and Vet 3 has the fewest sick. Agree denotes the samples where the three vet all have graded a sample in the same group. They agree that 1081 viscera sets are healthy and 1073 viscera sets are sick, not necessarily the same disease.
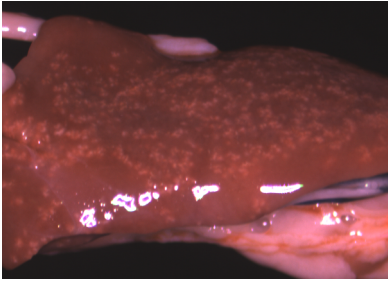
(a) Necrotic hepatitis.
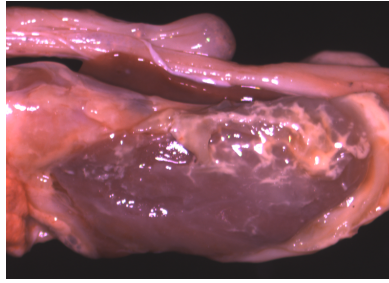
(b) Perihepatitis.
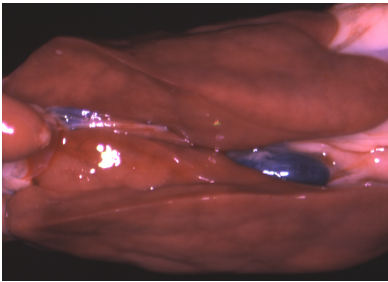
(c) Cobblestone Liver.
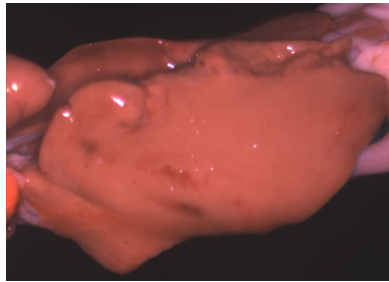
(d) White liver.
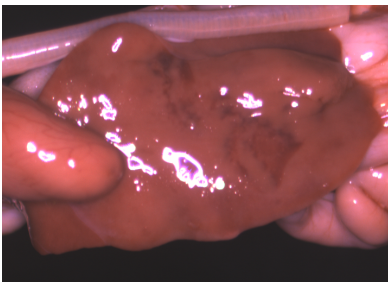
(e) Liver changes.
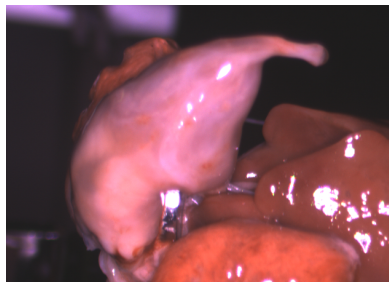
(f) Heart inflammation.

Fig. G.1: Examples of the six most frequent diseases seen in the viscera.
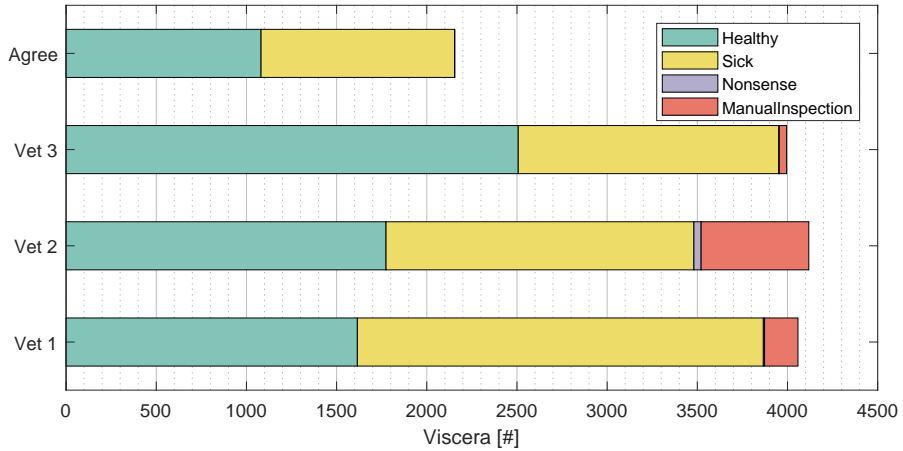
Fig. G.2: Overview of the categories graded by the three veterinarians. **Agree** is the number of sets where all veterinarians agrees. The viscera count is larger than the data set sample size, as some sets can be both sick and set for manual inspection for another disease.

Figure G.3 shows how the sickness categories are distributed. A viscera set can have multiple diseases so the total numbers here are typically higher than the sick portion in figure G.2. Agree is here lower than 1073 as figure G.3 shows the number of samples where the vets agree on the disease and not just whether the viscera is sick in general. They do not have to agree on the occurrence level.
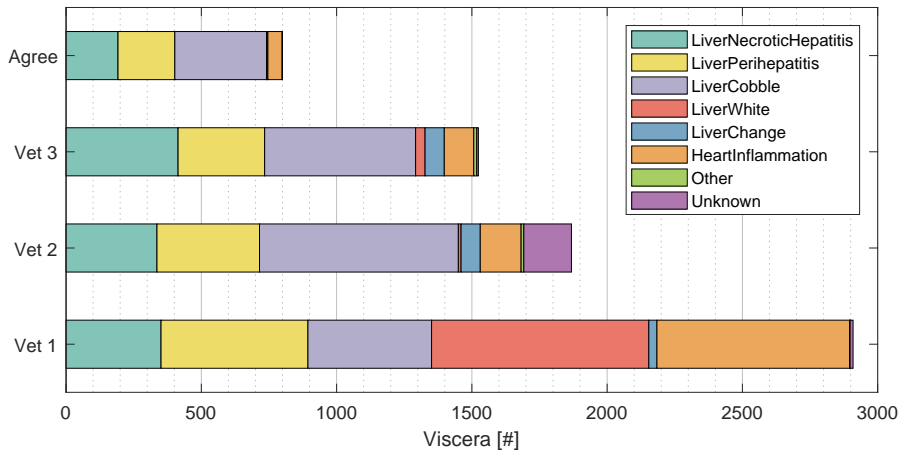
Fig. G.3: Distribution of the diseases in the data set. A viscera set with multiple diseases will be present in multiple categories.

Vet 1 has more sick samples than the other two vets and especially LiverWhite and HeartInflammation, which might be explained by having less

experience from the industry. Some of these can also be explained by the fact that Vet 1, more often than the others, grades a viscera set with more than one disease. This can be seen in figure G.4 which shows that Vet 1 has nearly 600 samples with more than one disease whereas Vet 2 and Vet 3 have 151 and 78, respectively. Vet 2 uses unknown and manual inspection more than the others and this might explain for some of the difference as well.
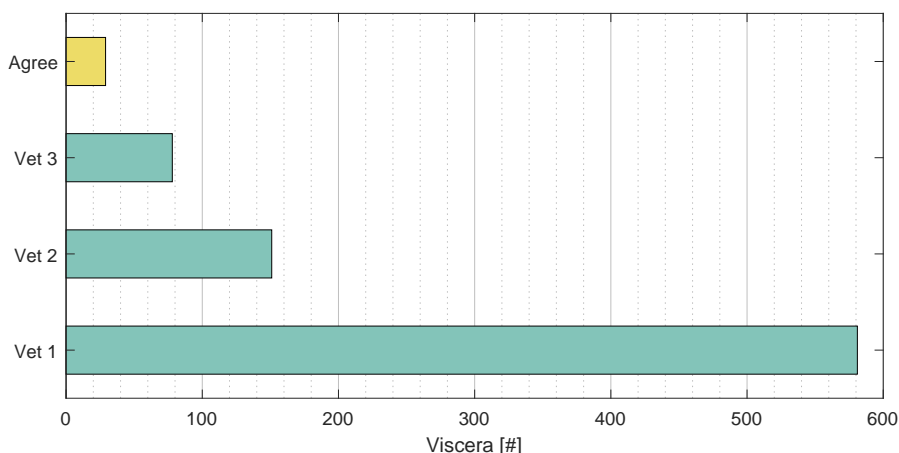
Fig. G.4: Viscera set categorised with more than one disease.

It is clear from the graphs that there is a large difference between the grades. From the 3984 graded samples only 2154 are agreed upon as either healthy or sick, even though the vets individually have 3866, 3481 and 3953 samples in these two categories. Looking at the large mismatch of the classes LiverWhite and HeartInflammation, more could probably have been done to align the vets with more workshops during the grading period. But that doesn't change the fact that grading these viscera is a difficult task.

3 Rerunning paper F with graded data

The data used in the paper F where among the samples graded by the three veterinarians. For the paper it was graded by the author in collaboration with a veterinarian. To test the algorithm with the new graded data set all samples where the vets did not agree on the disease were removed. Table G.3 shows the remaining samples after the clean up.

Grade	Old data set [#]	new data set [#]	Remaining [%]
Healthy	302	177	58.6 %
Cobblestone liver	226	156	69.0 %
Perihepatitis	145	78	53.8 %
Necrotic hepatitis	91	73	80.2 %

Table G.3: Cleaned up dataset where the three veterinarians agree on the disease.

Table G.4 show the result as they were presented in paper F. The overall classification error was 22.4 %.

		Prediction				True
		H	**C**	**P**	**N**	**Positive**
True	**H**	242.5 ± 3.47	27.2 ± 1.93	32.3 ± 2.71	0.0 ± 0.00	80.3%
	C	37.1 ± 3.57	176.3 ± 4.37	10.1 ± 1.66	2.5 ± 0.71	78.0%
	P	34.1 ± 3.28	7.9 ± 0.74	103.0 ± 3.06	0.0 ± 0.00	71.0%
	N	10.1 ± 0.57	8.6 ± 0.52	1.0 ± 0.47	71.3 ± 0.48	78.4%

Table G.4: Results from paper F. Classification matrix for the overall system. Average $\pm$ one standard deviation.

Using the regraded samples, we get the results in table G.5. The true positive rate has increase for all classes and the overall classification error has dropped to 14.4 %. Especially Necrotic hepatitis shows a large increase in performance from 78.4 % to 94.5 % true positives.

		Prediction				True
		H	**C**	**P**	**N**	**Positive**
True	**H**	153.7 ± 1.34	13.3 ± 1.16	9.9 ± 1.10	0.1 ± 0.32	86.8%
	C	17.4 ± 2.12	133.0 ± 1.89	4.5 ± 1.08	1.1 ± 0.32	85.3%
	P	16.8 ± 0.79	3.0 ± 0.00	58.2 ± 0.79	0.0 ± 0.00	74.6%
	N	1.0 ± 0.00	3.0 ± 0.00	0.0 ± 0.00	69.0 ± 0.00	94.5%

Table G.5: Results with the new data set. Classification matrix for the overall system. Average $\pm$ one standard deviation.

No changes were made to the classifier and all features are the same. The performance is calculated from 10 five-fold runs as in the paper. It is not a surprise that the results get better, as we have effectively removed samples where the veterinarians were uncertain. Yet we argue that this result is fairer as the algorithm is now compared to multiple veterinarians.

Paper H

Classify broiler viscera using an iterative approach on noisy labelled training data

Anders Jørgensen, Jens Fagertun, and Thomas B. Moeslund

Abstract

Poultry meat is produced and slaughtered at higher and higher rates and the manual food safety inspection is now becoming the bottleneck. An automatic computer vision system could not only improve the slaughter rates but also lead to a more consistent evaluation. This paper presents a method for classifying broiler viscera into healthy and unhealthy, in a data set recorded in-line at a poultry processing plant. The results of the on-site manual inspection are used to automatically label the images during the recording. The data set consists of 36,228 images of viscera.

The produced labels are noisy, so the labels in the training set are corrected through an iterative approach and ultimately used to train a convolutional neural network. The trained model is tested on a ground truth data set labelled by experts in the field. A classification accuracy of 86 % was achieved on a data set with a large in-class variation.

1 Introduction

The consumption of poultry meat has seen a large increase from 2.88 kg on average per person in 1961 to 14.99 kg in 2013 [1]. To keep up with the high demand, broilers are slaughtered at rates up to 13,500 birds per hour on a single conveyor [2]. The processing line is highly automated to achieve these rates, yet the food safety inspection is still done manually by veterinarians. Both the carcass and the viscera are inspected and with nearly four birds per second this is a task that requires constant attention. The inspectors are therefore relived every 30 minutes for a break to avoid straining themselves, but these shifts cause inconsistency as the inspection can be very subjective. An automatic computer vision system could assist the veterinarians and result in a better and more consistent food safety inspection.

This paper focuses on the inspection of the viscera. The goal of a developed system is to detect and remove unhealthy sets of viscera from the conveyor. But getting labelled images is expensive and time consuming as the viscera must be graded by veterinarians. We developed a recording system utilising the existing work procedures in the slaughter house to attain automatically labelled data albeit with with a certain amount of noise in the labels. In an iterative approach the mislabelled samples where detected and relabelled and used to train a convolutional neural network (CNN). A small subset of data was graded by experts and used for validation and test.

2 Related work

Not much work have been done on the topic of inspection of poultry viscera. Early work showed that multi-spectral images of the heart could be used to classify four different diseases with true positive rates between 84 % and 100 % [3]. The used data set contained 125 hearts, 25 in each category. [4] achieved a classification rate of 96 % in detecting septicaemia in chicken livers by utilising the NIR spectrum. Their data set contained 200 livers used for training and 100 for testing. [5] used UV light to detect splenomegaly in poultry carcasses. 57 images were in the data set on which they achieved a 95 % classification rate. All data sets were fairly small and the heart and liver were separated from the rest of the viscera before the classification, which is impractical in real life scenarios. [6] describes a method for diagnosing broiler livers still attached to the viscera set. A data set of 1,476 images was used to separate the livers into four categories and achieved true positive rates between 71.0 % and 80.3 %.

Getting labelled data to use in supervised learning can be a challenge, especially when experts are needed to obtain ground truth. In some situations labels can be gathered effortlessly, for example via Google image search or Instagram hash tags, but these labels will often we inaccurate. In a survey from 2014 [7] the author describes three ways to handle noisy labels; Noise-cleansing, Noise-tolerant and Noise-robust. Noise-cleansing models tries to remove or filter the noisy samples before training. Noise-tolerant models attempt to model the noise as part of the training, which requires some prior knowledge about the label noise. Noise-robust methods does neither of the previous, but trains with the noisy labels.

[8] proved that deep learning methods are fairly good at handling uniform noise, by showing a classification performance over 90 % on the MNIST data set, with 100 noisy labels for each correct label. CIFAR-10 and ImageNet showed similar trends but performance was hit harder by the number of noisy labels. All experiments, however, showed that the best performance was achieved with no noisy labels.

Many methods have been developed to utilise noisy training set. With the rise of deep learning, data sets needs to be larger which often introduce more noise, as it gets more time consuming and costly to grade the samples manually. [9] proposed a technique using a complementary neural network and removed samples misclassified by both networks. The method gave a minor performance boost for all four tested data sets. [10] augmented a neural network with a noise model that is learned simultaneous with the base model. The noise model effectually denoises the noisy samples through backpropagation, which results in a more accurate model. The augmented network performed considerably better than the base model, especially at noise level
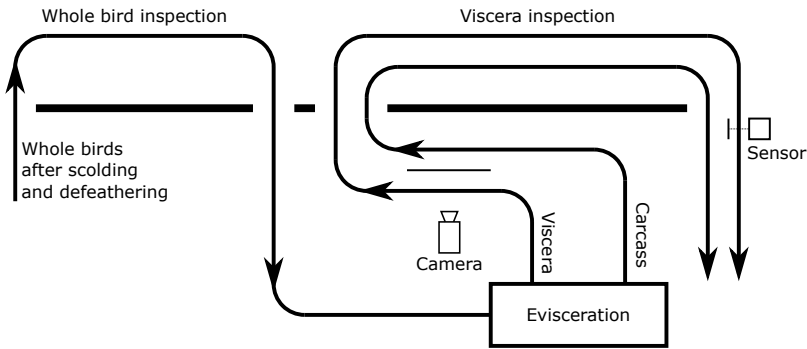
Fig. H.1: Overview of the conveyor around the inspection and evisceration.

above 30 %. The method presented in this paper utilises the noise-robustness of neural networks to identify and relabel the mislabelled samples in the training set.

3   Data set acquisition

The data acquisition consists of two parts, first image recording and later automatic image labelling. In order to utilise the labels already given to the viscera by the on-site inspectors, one must understand the normal working procedure at the evisceration and inspection part of the processing line. This has been sketched in figure H.1.

The birds are first inspected for diseases that are detectable on the carcass. This can be bruises, eczema and fractures among others. If the broiler is unfit for consumption it is removed from the line. The broilers are then transported into another room for evisceration. From here viscera and carcass, now separated, move back into the inspection room where the viscera are inspected for diseases. The evisceration happens in a different room due to high noise levels and as the use of water on the scolded carcasses causes a high humidity. As the camera and flash must not disturb the inspectors the recording system is installed in this room just after the evisceration.

As the viscera sets pass the inspection the unhealthy sets are removed and the healthy sets stay on the conveyor. In order to utilise this knowledge a beam break sensor is placed on the viscera conveyor after the inspection. If the sensor detects that a hanger is empty is means that the viscera has been removed and the corresponding image can be labelled "unhealthy".

3.1 Data handling

The images are recorded with a LW-AL-IMX253C-USB3 camera from ISG which is a global shutter colour camera capable of capturing images at a resolution of 4096x3000. The broilers are slaughtered at a rate of around 3.75 birds per second, which means the camera generates 45 MiB every second. As the grade is not known by the time the viscera passes the camera all viscera sets are captured and saved. However, only around 0.7 % of the broilers will be unhealthy [11] and many of these will be removed during the inspection of the carcass. To capture a more balanced data set the system automatically deletes most of the images captured of viscera that are not removed from the conveyor. All images of viscera that are removed are saved.

It must be assumed that each category contains a certain amount of noise. If the inspectors fail to remove an unhealthy viscera it will be labelled as "healthy". The risk of it being saved is small, however. It can also happen that the inspectors removed some healthy viscera sets by mistake or because it looks suspicious on the conveyor. Bear in mind that they only have around 1/4th of a second to make the decision and they will likely rather remove one too many than one too few. These will be labelled "unhealthy". If a viscera set falls out of the hanger between the camera and the sensor, it will also be categorised as "unhealthy".

Lastly, for some unhealthy viscera, the disease might be present in an area that is not visible in the image. These will end up in the unhealthy category but should for this purpose be regarded as healthy.

3.2 Resulting data set

The data set consist of 13,542 "not removed" images and 22,686 "removed" images captured in the period October 21th 2016 to December 10th 2017. Six examples of each class can be seen in figure H.2. It is clear that there is a large variation within both the healthy and unhealthy class.

A small set of the images were graded by veterinarians, which contains 1208 unhealthy and 1086 healthy images. 500 images from each class will be used in the test set and the rest will be used for validation. The final split can be seen in table H.1. The training images uses the noisy labels given during the acquisition.

	Training	Validation	Test
Healthy (**H**) = Not removed	12456	586	500
Unhealthy (**U**) = Removed	21478	708	500

Table H.1: Number of samples in each set. Only the validation and the test set have been graded by veterinarians.

Fig. H.2: Examples of an images in the data set. Top row is healthy, bottom row is unhealthy.

3.3 Preprocessing

The images have been captured over a period of a year and it was noticed that the images got brighter over time. This is primarily due to limescale accumulating on the glass in front of the lens which reflect light back into the lens. The images contain a large portion of the black background and all images are therefore corrected by simple subtraction to have a background colour of 30, 30, 30 in R, G, and B, respectively.

4 Method

To obtain a CNN that can classify the viscera as healthy or unhealthy, we employ an iterative approach where we train using the noisy labelled data and after each iteration we grade the samples with the largest error. The next iteration is then trained on the same data, but with the new labels applied. A sample can also be removed from the training set if it contains no useful information for the classifier. This can be images where large parts of the viscera are missing or hanging half way out of the frame. These images were also removed from the validation and test set by the veterinarians during grading.

4.1 Training the CNN

All training is done in python with the CNTK framework from Microsoft [12]. The network is a simple neural network with three convolutional layers. The entire network are depicted in table H.2.

Convolution(5,5)
MaxPooling(3,3)
Dropout(0.2)
Convolution(5,5)
MaxPooling(3,3)
Dropout(0.35)
Convolution(5,5)
MaxPooling(3,3)
Dropout(0.5)
AvgPool(8,8)
Dense(2,1)

Table H.2: Neural network. Each convolution layer is followed by an relu activation. MaxPooling uses a stride of 2 in each direction.

For each iteration we train 300 epochs with a batch size of 64 and a constant learning rate of 0.005. Each epoch uses 100,000 samples each augmented with random crops and horizontal flips. The images are scaled to a resolution of 150x300. After each iteration we select a model with which to predict the training samples. This model must perform well on the validation set but we also want to avoid overfitting, as we want the model to fail on the training samples that are mislabelled. The selected model is therefore the first model that is less than 5 correct validation samples from the minimum validation error. As the validation set contains 1294 images this corresponds to a maximum difference of 0.39 %.

The incorrectly predicted training samples are ranked in descending order by the difference in their class predictions as in equation H.1.

$$P_D = P_W - P_C \qquad \text{(H.1)}$$

P_W is the probability of the wrong class and P_C is the probability of the correct class. P_D will always be positive for wrongly classified samples and will be 0 when both classes are equally likely at 0.5 and 0.5. The first 1000 samples are then graded by the author as either healthy, unhealthy or invalid. Invalid samples are excluded from the training set for the following iterations. This grading is done after each iteration. When a sample has been graded once it is omitted in the following grading iterations, so that no sample is graded twice.

The last iteration will be validated and tested in two ways. First in the same way as for the previous iterations to compare the results of the final manual grading round. The model will then be trained again as the purpose is no longer to identify mislabelled training samples, but to get the optimal performance from the network. To make sure that the model has converged it will train for 1000 epochs, 500 with a learning rate of 0.005, then 300 epochs with 0.003 and 200 epochs with a learning rate of 0.001. The performance on the validation set will be used for early stopping.

5 Results

A total of seven iterations have been run, where the first, iteration 0, was without any manually graded training samples. For the next six iterations a total of 6000 training samples were manually inspected. The amount of samples that were either relabelled or discarded can be seen in figure H.3, plotted for each iteration. The total number of relabelled samples is 2870 and 425 was discarded. The majority of relabels are from unhealthy to healthy which is in line with the assumption about the noise in the data set.
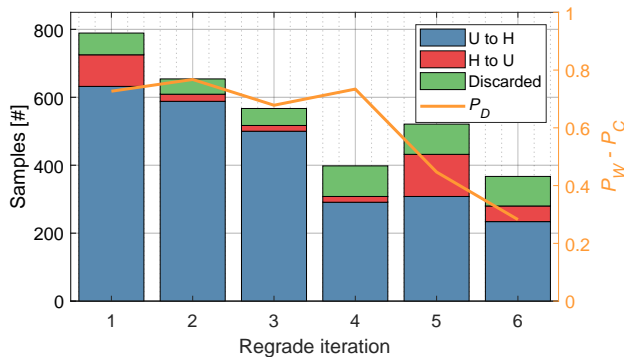

Fig. H.3: Regrades for each iteration. 1000 samples were graded for each iteration. P_D describes the certainty of the wrong prediction.

Figure H.3 also shows the average prediction difference, P_D, for the 1000 samples graded in each iteration. It can be seen that it drops rapidly for iteration 5 and 6, indicating that the model is no longer certain about the predictions. This also means that the samples get harder to grade as the viscera no longer are clearly unhealthy or clearly healthy. It was therefore chosen to stop grading after iteration 6.

Figure H.4 shows the minimum train and validation error for each iteration. The training error drops rapidly for the first 3 iterations and seems to

converge around 7 %. The validation error shows a more linear decline from 13.60 % at iteration 0 to 11.44 % at iteration 6.
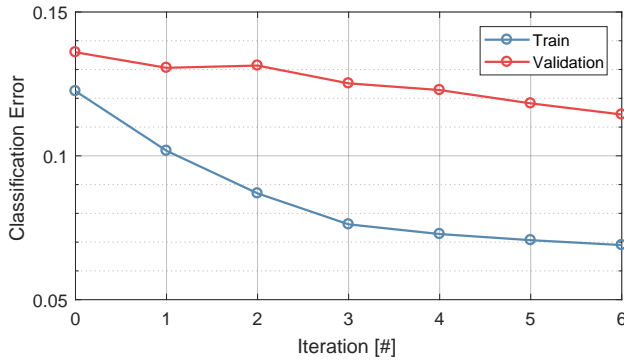

Fig. H.4: Minimum training and validation error for the seven iterations. All iterations have been trained for 300 epochs.

The results of the final training can be seen in figure H.5. This model uses the training labels from iteration 6, but trains for 1000 epochs instead of 300. The training error seems to converge around 6 % and the validation converges around 12.5 %.
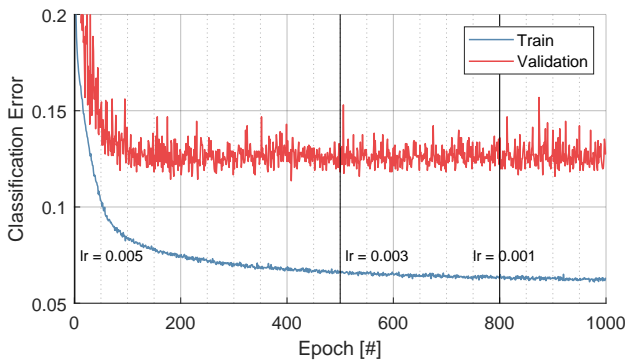

Fig. H.5: The last iteration trained for 1000 epochs.

The smallest validation error is 11.36 % at epoch 408, which is a small improvement compared to the model only trained for 300 epochs. The model at epoch 408 was the selected for testing and the classification error on the test set was 14.00 %.

6 Conclusion

Automatic classification of viscera is a difficult task as there is a large variation is both the healthy and the unhealthy samples. In this work we presented a method that achieved a classification accuracy of 86 %. Data graded by experts was used for validation and test only, while the training was perform on noisy labelled data. Of the 36,228 images in the data set only 6.33 % were graded by experts. Through the training iterations 17.68 % of the training samples were graded by the author and 8.46 % was relabelled.

A fully graded data set is always desirable, but grading images is a slow and expensive process, especially when a correct grading requires the use of experts. The author is no expert, but simply trained from the samples in the validation set. This proved good enough to spot and regrade crude errors in the automatically labelled training set, but when the incorrectly predicted samples became less and less obvious the grading became harder and harder.

The results indicate a clear improvement over the 6 grading iterations where the classification error on the validation set fell from 13.60 % to 11.44 %. Even though the input images are binary labelled, the content is very varying and there can be multiple reason to reject a viscera. This variation can probably explain some of the gap seen between the validation and the test error, if the two sets doesn't fully express the population variance.

References

[1] H. Ritchie and M. Roser, "Meat and Seafood Production & Consumption," 2018. [Online]. Available: https://ourworldindata.org/meat-and-seafood-production-consumption

[2] Marel Poultry, "The world of Poultry Processing," *Marel.com*, 2018. [Online]. Available: https://marel.com/files/pdf/world-of-stork-poultry-en.pdf

[3] K. Chao, Y. R. Chen, W. R. Hruschka, and B. Park, "Chicken Heart Disease Characterization by Multi-spectral Imaging," *Applied Engineering in Agriculture American Society of Agricultural Engineers*, vol. 99, no. 171, pp. 99–106, 2001.

[4] B. P. Dey, Y. R. Chen, C. Hsieh, and D. E. Chan, "Detection of septicemia in chicken livers by spectroscopy,," *Poultry Science*, vol. 82, no. 2, pp. 199–206, feb 2003.

[5] Y. Tao, J. Shao, K. Skeeles, and Y. R. Chen, "Detection of splenomegaly in poultry carcasses by UV and color imaging," *Transactions of the Asae*, vol. 43, no. 2, pp. 469–474, 2000.

[6] A. Jørgensen, J. Fagertun, and T. B. Moeslund, "Diagnosis of Broiler Livers by Classifying Image Patches," in *Lecture Notes in Computer Science*. Springer, Cham, jun 2017, vol. 10269 LNCS, pp. 374–385.

[7] B. Frenay and M. Verleysen, "Classification in the Presence of Label Noise: A Survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 845–869, may 2014.

[8] D. Rolnick, A. Veit, S. J. Belongie, and N. Shavit, "Deep learning is robust to massive label noise," *CoRR*, vol. abs/1705.10694, 2017.

[9] P. Jeatrakul, K. W. Wong, and C. C. Fung, "Data Cleaning for Classification Using Misclassification Analysis," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 14, no. 3, pp. 297–302, apr 2010.

[10] I. Jindal, M. Nokleby, and X. Chen, "Learning Deep Networks from Noisy Labels with Dropout Regularization," in *2016 IEEE 16th International Conference on Data Mining (ICDM)*. IEEE, dec 2016, pp. 967–972.

[11] U. Löhren, "Overview on current practices of poultry slaughtering and poultry meat inspection," *EFSA Supporting Publications*, vol. 9, no. 6, pp. 1–58, jun 2012.

[12] F. Seide and A. Agarwal, "Cntk: Microsoft's open-source deep-learning toolkit," in *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 2135–2135. [Online]. Available: http://doi.acm.org/10.1145/2939672.2945397

Paper I

Weight Estimation of Broilers in Images using 3D Prior Knowledge

Anders Jørgensen, Jacob V. Dueholm, Jens Fagertun, and Thomas B. Moeslund

Abstract

Cameras are already widely used for inspection and monitoring tasks in poultry slaughter houses. In this paper we evaluate the use of computer vision for broiler carcass weight estimation. We compare the use of 2D image features with 3D features extracted from a statistical shape model fitted to the image. The statistical shape model (SSM) is built from 45 3D scans captured from broiler carcasses collected at a slaughter house. The use of a 3D prior gave a minor yet significant 1.93 % reduction in mean absolute error compared to 2D features alone and achieved an overall mean average percentage error of 3.47 %. The algorithm can run real time and was tested on a dataset containing 136,472 images of broilers.

1   Intro

The production of poultry meat is rising and at the same time the production gets more automated. 40.6 billion chickens were slaughtered in 2000 which has increased to 65.8 billion in 2016 [1]. Broilers are slaughtered and cut up almost entirely by robots and automated equipment at speeds up to 13,500 birds per hour for one slaughter line [2]. To keep the production running smoothly the equipment needs measures like size and weight at multiple locations along the production line.

Broilers are typically weighed with a physical weigher installed as part of the conveyor. Such weighers are typically quite large and maybe requires the bird to be transferred off and back on the conveyor. Maintenance and replacements require that the line is stopped or bypasses the weigher.

The average weight of a flock can be used to adjust the cut-up equipment that separates the chicken into breast, legs, and wings. The individual size and weight are used to direct the broiler to the right cut-up station. By setting the equipment correctly the factory minimizes waste and avoids faulty cuts. Shortly after being slaughtered very light or small birds can be removed, as these are likely sick or underdeveloped. Trying to eviscerate these birds can cause the equipment to damage the intestines, causing faecal contamination on the following birds [3].

Many slaughter houses already have multiple camera systems placed along the processing line for inspection and monitoring tasks. They are installed in a non-intrusive manor, so they are relatively inexpensive to install and easy replaceable. A 2D image only shows the size of the bird, but because broilers are bred to be similar there is a high correlation between size and weight [4].

However, some variation may not be explained in a 2D image, e.g. dimensional changes can happen in the direction towards the camera. This paper investigates whether adding 3D features from 3D prior will increase the per-

formance of the weight estimation. The prior knowledge consists of 45 3D scans of broilers gathered at a poultry processing plant. A statistical shape model generated from the 3D scans is fitted to the 2D images to extract 3D features for the weight estimation.

2   Related works

Weight estimation from images is useful in many applications, especially in situations where physical weighing is not possible or feasible. In production of fresh lettuce where you can't uproot the plants to weigh them, [5] showed that it is possible to estimate the fresh weight using an image applied with both morphology and pixel-based methods.

A lot of work has been done in the field of estimating the weight of livestock in a non-intrusive fashion. The weight is an important parameter in rearing, but physical weighing requires large scales for animals like cattle and pigs and it can be a cumbersome affair to weigh hundreds of animals. Work by [6] shows that live weight of individual pigs can be estimated with an individual accuracy of 96.2 % using a camera installed over the pig pen. The body area of the pig is found by fitting an ellipse to the pig's back and used in a transfer function to estimate the weight every minute.

A similar top view approach was used to estimate the live weight of broilers but using the Kinect sensor to acquire an RGB-D image. Using a combination of 2D and 3D features, [7] achieved a relative error of 7.8 % across all broilers and ages. The system was installed in a commercial production setting and operated fully automatically.

In meat production environments, [8] showed that good results could be achieved with an RGB sensor and polynomial regressions when estimating the weight of salmon. Early work by [9] demonstrated a structured light setup capable of determining the weight of flatfish as they passed the camera and laser on a conveyor. A similar approach was used by [10], but on herring. The structured light gives a measure of the depth that is useful when estimating volumes, which for similar objects are highly correlated with the weight.

3   Approach

The goal is to estimate the weight of broilers depicted in 2D images captured in-line at a poultry processing plant. The 2D pipeline is part of an already existing pipeline for inspection of broilers. A flow chart of the system described in this paper is depicted in figure I.1.

A pre-processing step involves capturing multiple 3D scans of broilers to combine in a statistical shape model. This step is performed off-line and

should only be done once.

The in-line process starts with the image acquisition that captures a regular 2D image. From this image 2D landmarks are extracted, which are used to fit the SSM to the broiler in the image. From the fitted SSM 3D features are extracted and combined with 2D features which is used to estimate the weight of the broiler. An example of a captured image can be seen in figure I.9.
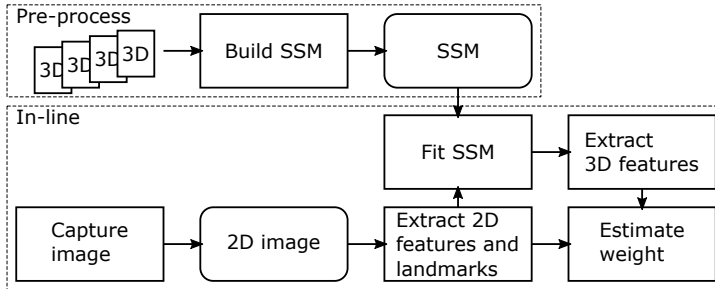

Fig. I.1: Flow chart of the weight estimation approach describe in this paper. The pre-process step is a one time only operation.

4 Statistical shape model generation

Fitting statistical models to images gained traction with the invention of Active Shape Models(ASM) [11] and Active Appearance Models(AAM) [12]. ASM and AAM have an inbuilt prior from the shapes used to construct the models. Using ASM [13] developed an automatic method for measuring femoral cartilage thickness in low-contrast images, where previous methods were only semi-automatic. [14] demonstrated the ability to fit a 3D face model to a single image of an unknown face using no manually annotated key points. [15] took it a step further when they introduced a system capable of fitting a 3D face model to a 24 fps video stream in real time using 75 automatically located landmarks.

4.1 Creating 3D scans of broilers

45 birds have been collected at a poultry slaughter house and recorded using an Canon EOS 5Ds camera over the course of 3 weeks to ensure diversity. The recording setup consisted of a hanger in the centre of the room and the camera was rotated around the chicken. One full rotation with the camera placed higher than the chicken and one with the camera placed lower than the chicken. Between 90 and 120 images where captured per bird. See figure I.2 for a sketch of the camera angles.
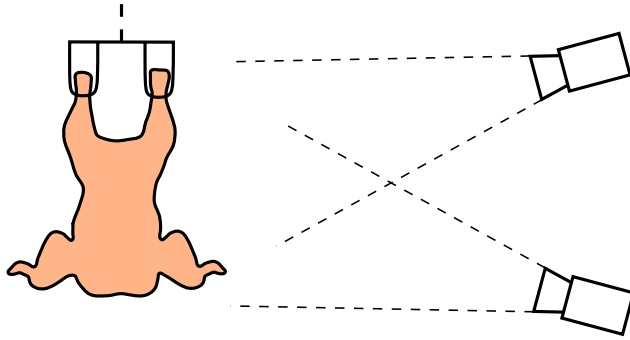
Fig. I.2: Sketch of the setup used for capturing images for the 3D scan generation. The camera is approximately 1 meter from the bird.

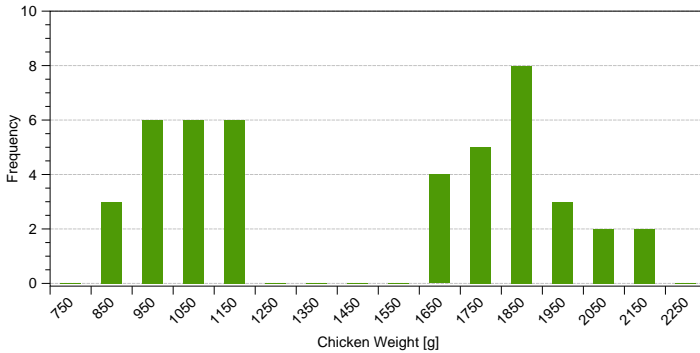The distribution of the collected birds can be seen in figure I.3.

Fig. I.3: Distribution of the weight of all birds used in the SSM.

The histogram in figure I.3 shows that there is a span in the weights between 1200 g and 1600 g where no birds have been recorded. The operators were asked to pick some average birds but also some extremes and it seems that they have picked too many extremes, without selecting any in between.

The recorded images were fed to the commercially available software, ContextCapture, which generated the 3D scans. Pieces of coloured tape were attached to the hanger two centimetres apart to ensure the correct scale of the bird in the 3D scan. Examples of three 3D scans rendered in Blender [16] can be seen in figure I.4. Each scan contains between 180,000 and 330,000 vertices.

(a) Bird 8. 1160 g. **(b)** Bird 26. 1770 g. **(c)** Bird 33. 1139 g.
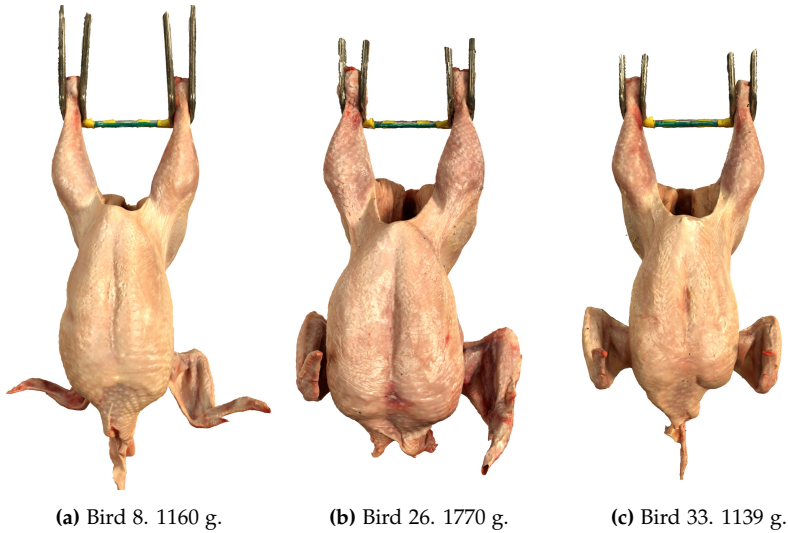
Fig. I.4: Three 3D scans created by ContextCapture. Rendered in Blender.

4.2 Registering the 3D scans

The 3D scans were manually trimmed to remove the wings, the knees and the neck skin as these parts can be very different between birds and therefore difficult to register. It is also assumed that especially the neck skin and the knees have a small and near constant impact on the total weight. The scans were then smoothed to remove small bumps on the skin, which otherwise would create very fluctuating normals along the surface.

One scan was chosen as the template and scaled down to around 25,000 vertices using the surface simplification method [17]. The template was then registered to the other scans with a non-rigid iterative closest point method called N-ICP-A [18]. The algorithm assigns an affine transformation to each vertex and starts with a stiff template to find the global alignment, then gradually reduces the stiffness to allow more localized deformations. The stiffness is used to regularize the deformation and controlled by penalizing the difference between neighbour vertices' transformations.

The registration produces a few very skewed faces, especially around the end of the legs due to badly registered vertices. To remove these faulty faces, it was chosen simply to remove a few vertices around the end of both legs. The same vertices must be removed from all scan to keep the correspondence between the models. Unconnected vertices are also removed, again from all scans. As the last step all scans are refaced with the faces from the template. The resulting meshes for the three birds in figure I.4 can be seen in figure I.5.

The area around the groin proved difficult to register as there is a large

variation between the birds in this area. This can also be seen in figure I.4, where there is a big height difference between the backside of the thigh and the end of the breast bone between the birds. As a result, vertices around the cloaca were the viscera are eviscerated were removed.
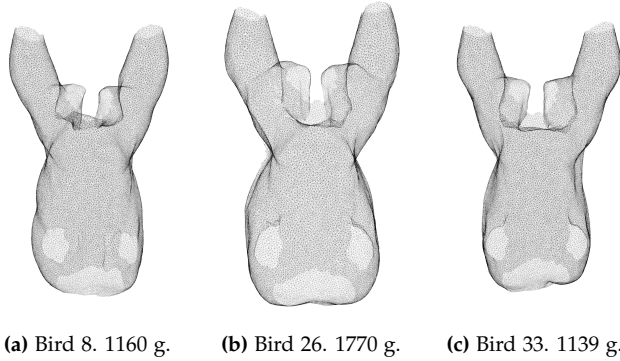

(a) Bird 8. 1160 g.　　(b) Bird 26. 1770 g.　　(c) Bird 33. 1139 g.

Fig. I.5: Template scan registered to the three birds in figure I.4.

The registered and trimmed scans are now used to generate the SSM. All scans are aligned with Procrustes Analysis without scaling and flipping. The mean of all scans is then subtracted from the individual scans before Principal Component Analysis (PCA) is used to describe the variation in the data. Studying the explained variances show that the first 7 components contains more than 95 % of the total variation. The mean shape can be seen in figure I.6.

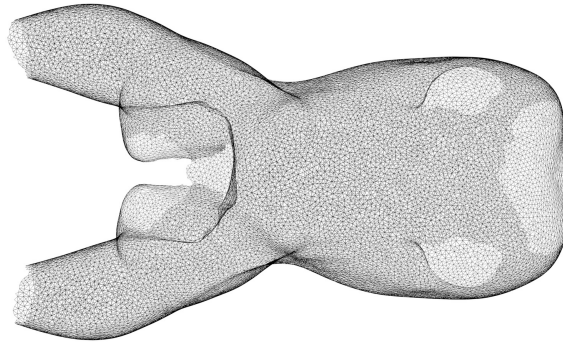**Fig. I.6:** Mean of the SSM. Rotated for better viewing.

New samples can now be generated using equation I.1.

$$\underset{3n \times 1}{\mathbf{x}} = \underset{3n \times 1}{\bar{\mathbf{x}}} + \underset{3n \times m}{\mathbf{P}} \cdot \underset{m \times 1}{\mathbf{b}} \tag{I.1}$$

where **x** is a new sample generated from the SSM, x̄ is the mean shape, **P** is the eigenvectors found with PCA and **b** defines a set of parameters that deform the SSM. n is the number of vertices in the model and as the vertex coordinates are stored as $x_1, y_1, z_1, x_2, y_2, z_2, ...x_n, y_n, z_n$, the number of rows in **x**, x̄ and **P** becomes $3n$. m is the number of principal components.

5 Fitting the SSM to a 2D image

All birds are presented the same way, hanging in the legs with the breast facing the camera. This allows us to lock the yaw and pitch rotation of the SSM reducing the degrees of freedom in the fitting problem. As the broilers are transported sideways on the conveyor, some differences in roll must be expected between the images.

The fitting is done with automatically located landmarks. 8 points are extracted from the bird in the 2D image. This is the left and right wing-pit, shoulder, hip and groin. The 8 points are depicted in figure I.10. The broilers can look very different, but these points were chosen as they could be found most consistently. 8 corresponding vertices are selected in the SSM. These vertices are handpicked, as if the z-values are set to zero and all vertices are compressed to the image plane. The vertices are only selected once as they should always correspond to the same landmarks in the images.

When the landmarks are extracted from the image, Procrustes Analysis is used to align the points to the landmarks in the SSM, again without scaling. The SSM and image landmarks are now aligned and ready for the model to be fitted. As the orientation of the SSM is now locked, only the x and y values are used for the fitting which gives a total of 16 values in **x**. Equation I.1 will now have the dimensions showed in equation I.2. Only the 7 first principal components are used.

$$\underset{16 \times 1}{\mathbf{x}} = \underset{16 \times 1}{\bar{\mathbf{x}}} + \underset{16 \times 7}{\mathbf{P}} \cdot \underset{7 \times 1}{\mathbf{b}} \tag{I.2}$$

$$0 = \mathbf{Pb} - (\mathbf{x} - \bar{\mathbf{x}}) \tag{I.3}$$

Equation I.3 is an overdetermined problem and hence it will have multiple solutions. We want to constrain values in **b** to ±3 standard deviations. This interval contains 99.73 % of all seen values and it is therefore uncertain how the model will behave outside this interval. Equation I.3 will therefore be solved with an optimizer. The minimization is done with Scipy's [19] minimize method. The algorithm used is L-BFGS-B [20] and each scalar in **b** is bound to ±3 std. Initial values for **b** is zeros. The resulting fit for three broilers can be seen in figure I.7.
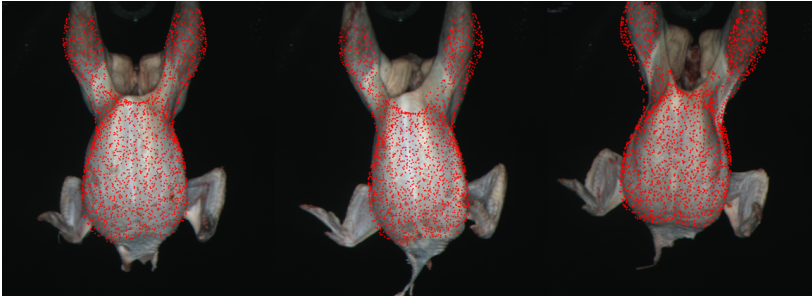
Fig. I.7: The SSM fitted to three different images. Red dots are vertices in the model.

6   Features

Using the fitted SSM 3D features can now be extracted for each bird. One feature is the total volume, which is calculated as the sum of all tetrahedras formed by the faces of the SSM and the origin which is placed roughly in the canter of the SSM. The volume of each tetrahedra is calculated by finding the determinant of the 3 vectors spanning from origin to the three vertices of each face and divide it by 6. The determinant is signed, so the faces pointing inwards will give negative volumes and faces pointing outwards give positive. If a 3D model is manifold this means that the total volume will be correct even though some tetrahedra will cross surface of the models. However, the SSM is non-manifold so the volume will not be what one might expect from a chicken, yet it will still grow proportional to the size of the model.

The remaining features are either areas or distances on the surface of the SSM. Because there is point correspondence between all fitted SSMs, surface areas can be calculated from the same set of vertices every time, it is just the location of these vertices that have changed. The surface area is found by summing the area of all faces spanned by the vertices. Because these surfaces are in 3D they should grow as the volume of the broiler grows. The area features include the left and right breast, left and right upper thigh and a band around each drum. Figure I.8 shows the faces used to calculate the surface area of the right breast.
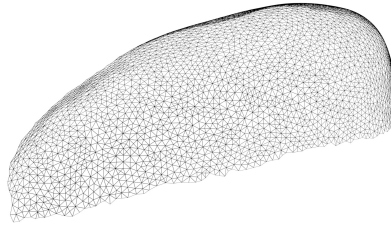
Fig. I.8: Faces used to calculate the surface area of the right breast.

The same principle is used to calculate distance features. Pre-specified vertices make a path that is used to calculate the distance between two vertices. This path is the shortest in the mean shape, but because the vertices move individually when the SSM is fitted to the 2D image, it is not necessarily the shortest path in the fitted SSM. It is however much faster to calculate the distance of a fixed path than searching for the shortest path between two vertices for every fit. The distance features include the path from the collarbone to the bottom of the breast, the circumference of each drum and the path from each wing-pit to the centre of the breast. In total 12 3D features are extracted from the fitted SSM.

2D features are extracted directly from the 2D image and are generally areas and distances like the 3D features. These are features like the area of the breast, area of the drums and the length from wing-pit to wing-pit. In total 23 2D features are extracted.

All features are augmented twice by finding the square root and by squaring each of them. This is done to introduce some non-linearities. With augmentation there are 42 3D features and 69 2D features.

7 Data acquisition

2D images of broilers are recorded with a Jai BB-141GE RGB camera installed at a slaughter house in Chesterfield, England. The images are recorded after defeathering and evisceration where the broilers are transported sideways hanging in their legs with the breast facing the camera. Their weight is measured with a LINCO 520 Weigh Transfer [21] mechanical weight with an accuracy of ± 0.25 %. The weight is paired with the images and will function as ground truth. An example of a captured image can be seen in figure I.9.

Fig. I.9: Image recorded with the 2D camera. Cropped for better viewing.
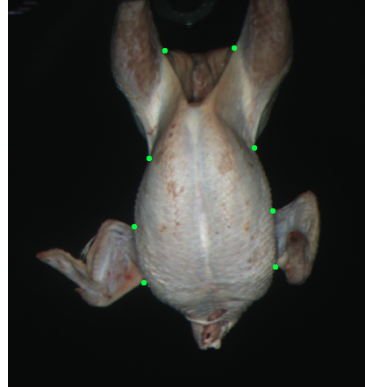

Fig. I.10: Landmarks used to fit the SSM to the image.

All 2D images have been recorded over the course of four days. Images of broilers weighing less than 800 g or more than 2200 g are discarded, so the remaining images are in the same weight range as the broilers used for the SSM. The remaining 136,472 images are randomized, and 102,412 images are used for training and 34,060 images are used for testing. The weight distribution of the birds in the training and the test set can be seen in figure I.11.
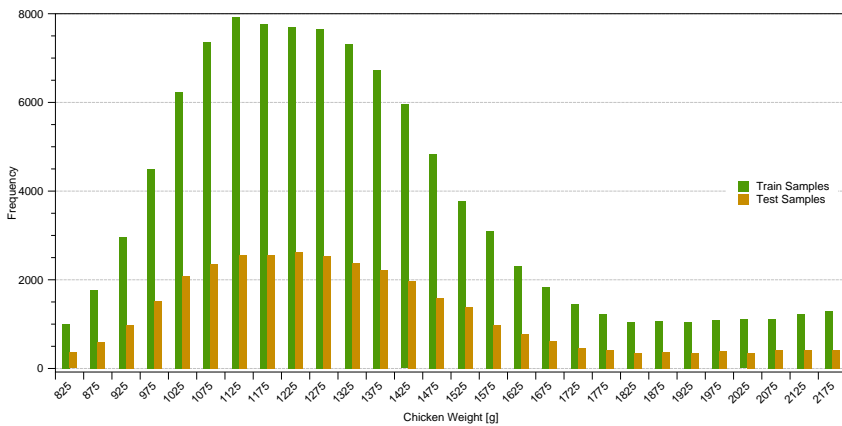

Fig. I.11: Distribution of the weight of all birds in the data set. Total bird count is 136,472.

8   Results

8.1   Weight estimation performance

Performance are measured for 2D features only (2D), 3D features only (3D) and for a combination of 2D and 3D features (2D3D). The number of features used to train the regression models are listed in table I.1. All features are normed by subtracting the mean and dividing by the standard deviation. Both the mean and the standard deviation are calculated using only the training samples.

The regression model chosen for comparison is a linear robust regression. The robust variant means the linear model is fitted iteratively and for each iteration data points are weighted based on their residual value. Outliers will be down weighted and therefore have a smaller effect on the fit.

Model	# Features	MAE	MAPE
2D	23 (69)	47.22	3.53%
3D	14 (42)	63.49	4.72%
2D3D	37 (111)	**46.37**	**3.47%**

Table I.1: Results for the the linear regression models. Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are reported.

The resulting errors of the regression models are listed in table I.1. The error gets smaller by combining 2D and 3D features and MAE is reduced by 1.80 % from 2D to 2D3D. An unpaired t-test was performed on the 2D and 2D3D residuals to investigate whether this reduction is significant, which the resulting p-value of 0.0129 strongly indicates.

As there are no studies done on the same dataset we will compare our results to related weight estimation papers. Weight estimation of herring by [10] achieved an R^2 of 0.980 using structured light to extract 3D features and had a sample size of 179. They used 2D and 3D features where our method achieved an R^2 of 0.963 using 2D and pseudo 3D features on 34,060 test samples. In work by [6], weight estimation of pig, they got an R^2 of 0.962 working on multiple frames extracted from a video. They tracked the weight of individual pigs as it grew over 13 days where we sample and estimate the weight once.

8.2   Timing

The system should be able to operate in-line at production speeds up to 13,500 birds per hour. That is 266 ms of available processing time per bird. The 2D pipeline already take a part of this, so the SSM fitting and 3D features extraction must be fast. In this work we have the SSM fitting and 3D feature extracting running off-line on the previously recorded images with landmarks already extracted. The code is implemented in python 3.6 using

numpy 1.14 and numba 0.36.2 for speed-up where possible. Timing showed that the average processing time per image is 7 ms. The test was run on an i7-4770K running at 3.7 GHz. This is far below the 266 ms, so there should be time to add the 3D features to the already existing analysis.

9   Conclusion

The results showed that the mean absolute error fell by 1.80 % by adding 3D features from the SSM to the 2D features extracted directly from the image. A t-test was used to ensure that the results were significant.

The SSM was fitted to the 2D image using only 8 landmarks. This meant the SSM could be fitted very quickly, in just a couple of milliseconds, allowing the entire fit and 3D feature extraction to run real time. The soft curves of the broiler and the bland texture made it difficult to extract more than these eight landmarks and especially the landmarks around the wings proved to be volatile in some cases. The angle of the wings could deviate a lot on some birds, which have a large impact on the wing-pit and shoulder landmarks. Figure I.4 is a good example of how the wings can look different between birds. All 3D features depend on the fit, so a bad fit would lead to bad 3D features. For future work it would be interesting to find a more robust way to fit the SSM to the 2D image.

Constructing a statistical shape model is a time-consuming task. Capturing 90-120 images of each broiler is a tedious job and building the 3D scans is computationally expensive and therefore also time consuming. After this comes the manual process of inspecting and trimming the scans before the they can be registered to each other with yet another computationally expensive method.

Many steps in this process can however be automated which will make it easier to expand the SSM with more broilers and therefore more variance. For this paper we built the statistical model from 45 3D scans of broilers, which we used to estimate the weight of over 100,000 birds. 45 3D scans can clearly not represent the variation in that many birds, but the results indicate that the weight estimation error can be reduced by adding prior knowledge and it is our belief that an SSM with more 3D scans can improve the performance further.

References

[1] Food and Agriculture Organization of the United Nations, "Total poultry production," Online Statistics Database, 2017.

[2] Marel Stork, "The world of poultry processing," Online Brochure, 2016.

[3] I. Guerrero-Legarreta, *Handbook of poultry science and technology. Volume 1, Primary processing.* John Wiley, 2010.

[4] Aviagen, "ROSS 308 BROILER: Performance Objectives," 2014. [Online]. Available: http://en.aviagen.com/assets/Tech_Center/Ross_Broiler/Ross-308-Broiler-PO-2014-EN.pdf

[5] D.-H. Jung, S. Hyun Park, X. Zhe Han, and H.-J. Kim, "Image Processing Methods for Measurement of Lettuce Fresh Weight," *J. of Biosystems Eng*, vol. 40, no. 1, pp. 89–93, 2015. [Online]. Available: http://dx.doi.org/10.5307/JBE.2015.40.1.089

[6] M. Kashiha, C. Bahr, S. Ott, C. P. Moons, T. A. Niewold, F. O. Ödberg, and D. Berckmans, "Automatic weight estimation of individual pigs using image analysis," *Computers and Electronics in Agriculture*, vol. 107, pp. 38–44, 2014.

[7] A. K. Mortensen, P. Lisouski, P. Ahrendt, A. Krogh Mortensen, P. Lisouski, and P. Ahrendt, "Weight prediction of broiler chickens using 3D computer vision," *Computers and Electronics in Agriculture*, vol. 123, pp. 319–326, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.compag.2016.03.011

[8] M. O. Balaban, G. F. Ünal Şengör, M. G. Soriano, and E. G. Ruiz, "Using Image Analysis to Predict the Weight of Alaskan Salmon of Different Species," *Journal of Food Science*, vol. 75, no. 3, pp. E157–E162, apr 2010. [Online]. Available: http://doi.wiley.com/10.1111/j.1750-3841.2010.01522.x

[9] F. Storbeck and B. Daan, "Weight estimation of flatfish by means of structured light and image analysis," *Fisheries Research*, vol. 11, no. 2, pp. 99–108, apr 1991.

[10] J. R. Mathiassen, E. Misimi, B. Toldnes, M. Bondø, and S. O. Østvik, "High-speed weight estimation of whole herring (clupea harengus) using 3D machine vision," *Journal of Food Science*, vol. 76, no. 6, pp. 458–464, 2011.

[11] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active Shape Models-Their Training and Application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, jan 1995. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1077314285710041

[12] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *European Conference on Computer Vision*, pp. 484–498, 1998. [Online]. Available: http://link.springer.com/10.1007/BFb0054760

[13] S. Solloway, C. J. Taylor, C. E. Hutchinson, and J. C. Waterton, *Quantification of articular cartilage from MR images using active shape models*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1996, pp. 400–411. [Online]. Available: https://doi.org/10.1007/3-540-61123-1_156

[14] Lie Gu and T. Kanade, "3D Alignment of Face in a Single Image," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1 (CVPR'06)*, vol. 1. IEEE, 2006, pp. 1305–1312. [Online]. Available: http://ieeexplore.ieee.org/document/1640900/

[15] C. Cao, Y. Weng, S. Lin, and K. Zhou, "3D shape regression for real-time facial animation," *ACM Transactions on Graphics*, vol. 32, no. 4, p. 1, jul 2013. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2461912.2462012

[16] Blender Online Community, *Blender 2.79*, Blender Foundation, Blender Institute, Amsterdam, 2017. [Online]. Available: http://www.blender.org

[17] M. Garland and P. S. Heckbert, "Surface simplification using quadric error metrics," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques - SIGGRAPH '97*, 1997, pp. 209–216. [Online]. Available: http://portal.acm.org/citation.cfm?doid=258734.258849

[18] B. Amberg, S. Romdhani, and T. Vetter, "Optimal step nonrigid icp algorithms for surface registration," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–8.

[19] E. Jones, T. Oliphant, P. Peterson *et al.*, "SciPy: Open source scientific tools for Python," 2001–, [Online; accessed 08-12-2017]. [Online]. Available: http://www.scipy.org/

[20] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A Limited Memory Algorithm for Bound Constrained Optimization," *SIAM Journal on Scientific Computing*, vol. 16, no. 5, pp. 1190–1208, sep 1995. [Online]. Available: http://epubs.siam.org/doi/10.1137/0916069

[21] Linco Food Systems, "Linco 520 Weigh Transfer," 2018. [Online]. Available: http://baader.ca/files/products/poultry/Leaflet_A17_520_Weigh_Transfer_01.pdf