Detection of Sugarcane Crop Rows From UAV Images Using Semantic Segmentation and Radon Transform

Renato Rodrigues da Silva

Universidade Federal de Uberlândia
Faculdade de Computação
Programa de Pós-Graduação em Ciência da Computação

Uberlândia
2020

Renato Rodrigues da Silva

Detection of Sugarcane Crop Rows From UAV Images Using Semantic Segmentation and Radon Transform

Dissertação de mestrado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Prof. Dr. André Ricardo Backes
Coorientador: Prof. Dr. Mauricio Cunha Escarpinati

Uberlândia
2020

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

S586 Silva, Renato Rodrigues da, 1990-
2020 Detection of Sugarcane Crop Rows From UAV Images Using Semantic Segmentation and Radon Transform [recurso eletrônico] / Renato Rodrigues da Silva. - 2020.

Orientador: André Ricardo Backes.
Coorientador: Mauricio Cunha Escarpinati.
Dissertação (Mestrado) - Universidade Federal de Uberlândia, Pós-graduação em Ciência da Computação.
Modo de acesso: Internet.
Disponível em: http://doi.org/10.14393/ufu.di.2020.736
Inclui bibliografia.

1. Computação. I. Backes, André Ricardo,1981-, (Orient.). II. Escarpinati, Mauricio Cunha,1976-, (Coorient.). III. Universidade Federal de Uberlândia. Pós-graduação em Ciência da Computação. IV. Título.

CDU: 681.3

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em Ciência da Computação
Av. João Naves de Ávila, nº 2121, Bloco 1A, Sala 243 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
Telefone: (34) 3239-4470 - www.ppgco.facom.ufu.br - cpgfacom@ufu.br

ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciência da Computação				
Defesa de:	Mestrado Acadêmico, 34/2020, PPGCO				
Data:	07 de dezembro de 2020	Hora de início:	13:32	Hora de encerramento:	15:32
Matrícula do Discente:	11822CCP010				
Nome do Discente	Renato Rodrigues da Silva				
Título do Trabalho:	Detection of Sugarcane Crop Rows From UAV Images Using Semantic Segmentation and Radon Transform				
Área de concentração:	Ciência da Computação				
Linha de pesquisa:	Ciência de Dados				
Projeto de Pesquisa de vinculação:	-				

Reuniu-se, por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Marcelo Zanchetta do Nascimento - FACOM/UFU; Hemerson Pistori - UCDB; Mauricio Cunha Escarpinati - FACOM/UFU (coorientador) e André Ricardo Backes - FACOM/UFU, orientador do candidato.

Os examinadores participaram desde as seguintes localidades: Hemerson Pistori - Campo Grande/MS; Marcelo Zanchetta do Nascimento, Mauricio Cunha Escarpinati e André Ricardo Backes - Uberlândia/MG. O discente participou da cidade de Uberlândia/MG.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. André Ricardo Backes, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

Aprovado.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.

Documento assinado eletronicamente por **André Ricardo Backes**, **Professor(a) do Magistério Superior**, em 08/12/2020, às 10:01, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Marcelo Zanchetta do Nascimento**, **Professor(a) do Magistério Superior**, em 08/12/2020, às 10:14, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Mauricio Cunha Escarpinati**, **Presidente**, em 08/12/2020, às 15:51, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Hemerson Pistori**, **Usuário Externo**, em 09/12/2020, às 14:28, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2436269** e o código CRC **BABC7424**.

Referência: Processo nº 23117.072932/2020-02 SEI nº 2436269

Criado por ludimila.andrade, versão 11 por ludimila.andrade em 08/12/2020 09:12:31.

I dedicate this work to all type 1 diabetics who, like a tightrope walker, have to keep the balance between managing their lives while handling all the other responsibilities this world demands.

A special feeling of gratitude to my love Pâmela, whose words of encouragement and push for tenacity were crucial, and my sister Susana who had always believed in me. This work is for you guys, too.

Acknowledgements

First of all, I thank God, who have lit my way during this path. I thank all the professors of the course, who were very important in my academic life, especially those who have been true mentors to me: Mauricio Cunha Escarpinati, my co-advisor, for his companionship, friendship and all the advice given along these years; and André Ricardo Backes, who, in addition to being my advisor, is a great friend and has always been a key player in this journey. I also thank my friends and all the people who are important to me and who have always been by my side, especially my parents and wife. Thanks, Pâmela for the patience and for giving me strength at all times. Finally, I thank all those who, in some way, were and are close to me, making this life more and more worthwhile.

"Life is like riding a bicycle. To keep your balance, you must keep moving."
(Albert Einstein)

Resumo

Nos últimos anos, os VANTs (Veículos Aéreos Não Tripulados) têm se tornado cada vez mais populares no setor agrícola, promovendo e possibilitando o monitoramento de imagens aéreas tanto no contexto científico, quanto no de negócios. Imagens capturadas por VANTs são fundamentais para práticas de agricultura de precisão, pois permitem a realização de atividades que lidam com imagens de baixa ou média altitude. O cenário da área plantada pode mudar drasticamente ao longo do tempo devido ao aparecimento de erosões, falhas de plantio, morte e ressecamento de parte da cultura, intervenções de animais, etc. Assim, o processo de detecção das linhas de plantio é de grande importância para o planejamento da colheita, controle de custos de produção, contagem de plantas, correção de falhas de semeadura, irrigação eficiente, entre outros. Além disso, a informação de geolocalização das linhas detectadas permite o uso de maquinários autônomos e um melhor planejamento de aplicação de insumos, reduzindo custos e a agressão ao meio ambiente. Neste trabalho, abordamos o problema de segmentação e detecção de linhas de plantio de cana-de-açúcar em imagens de VANTs. Primeiro, experimentamos uma abordagem baseada em Algoritmo Genético (AG) e Otsu para produzir imagens binarizadas. Posteriormente, devido a alguns motivos, incluindo a relevância recente da Segmentação Semântica, seus níveis de abstração e os resultados inviáveis obtidos com AG, estudamos e propusemos uma nova abordagem baseada em Semantic Segmentation Network (SSN) em duas etapas. Primeiro, usamos uma SSN para segmentar as imagens, classificando suas regiões como linhas de plantio ou como solo não plantado. Em seguida, utilizamos a transformada de Radon para reconstruir e melhorar as linhas já segmentadas, tornando-as mais uniformes ou agrupando fragmentos de linhas e plantas soltas. Comparamos nossos resultados com segmentações feitas manualmente por especialistas e os resultados demonstram a eficiência e a viabilidade de nossa abordagem para a tarefa proposta.

Palavras-chave: Linhas de Plantio, Cana-de-açúcar, Segmentação, CNN, VANT, Transformada de Radon.

Detection of Sugarcane Crop Rows From UAV Images Using Semantic Segmentation and Radon Transform

Renato Rodrigues da Silva

Universidade Federal de Uberlândia
Faculdade de Computação
Programa de Pós-Graduação em Ciência da Computação

Uberlândia
2020

Abstract

In recent years, UAVs (Unmanned Aerial Vehicles) have become increasingly popular in the agricultural sector, promoting and enabling the application of aerial image monitoring in both scientific and business contexts. Images captured by UAVs are fundamental for precision farming practices, as they allow activities that deal with low and medium altitude images. After the effective sowing, the scenario of the planted area may change drastically over time due to the appearance of erosion, gaps, death and drying of part of the crop, animal interventions, etc. Thus, the process of detecting the crop rows is strongly important for planning the harvest, estimating the use of inputs, control of costs of production, plant stand counts, early correction of sowing failures, more-efficient watering, etc. In addition, the geolocation information of the detected lines allows the use of autonomous machinery and a better application of inputs, reducing financial costs and the aggression to the environment. In this work we address the problem of detection and segmentation of sugarcane crop lines using UAV imagery. First, we experimented an approach based on Genetic Algorithm (GA) associated with Otsu method to produce binarized images. Then, due to some reasons including the recent relevance of Semantic Segmentation in the literature, its levels of abstraction, and the non-feasible results of Otsu associated with GA, we proposed a new approach based on SSN divided in two steps. First, we use a Convolutional Neural Network (CNN) to automatically segment the images, classifying their regions as crop lines or as non-planted soil. Then, we use the Radon transform to reconstruct and improve the already segmented lines, making them more uniform or grouping fragments of lines and loose plants belonging to the same planting line. We compare our results with segmentation performed manually by experts and the results demonstrate the efficiency and feasibility of our approach to the proposed task.

Keywords: Crop-row, Sugarcane, Segmentation, CNN, UAV, Radon Transform.

List of Figures

List of Tables

Acronyms list

AUVSI Association for Unmanned Vehicle Systems International

CNN Convolutional Neural Network

DIP Digital Image Processing

DNN Deep Neural Networks

DSC Dice Similarity Coefficient

FCN Fully Convolutional Network

GA Genetic Algorithm

GNSS Global Navigation Satellite Systems

HSV Hue, Saturation, Value

HT Hough Transform

ICT Information and Communications Technology

IMU Inertial Measurements Unit

IoT Internet of Things

JSC Jaccard Similarity Coefficient

PA Precision Agriculture

RHT Random Hough Transform

RGB Red, Green, Blue

SHT Standard Hough Transform

SSN Semantic Segmentation Network

UAV Unmanned Aerial Vehicle

VI Vegetation Index

Contents

I hereby certify that I have obtained all legal permissions from the owner(s) of each third-party copyrighted matter included in my dissertation, and that their permissions allow availability such as being deposited in public digital libraries.

Renato Rodrigues da Silva

Chapter 1

Introduction

Sugarcane is one of the most planted cultures in the planet. Its planting practice and mechanization are the development tendency of the modern agro-industry (UCHIMIYA; SPAUNHORST, 2020). Brazil is the largest producer of sugarcane in the world. The country registered an area around 10,123.5 million hectares (Mha) of land planted with the crop in the 2018/2019 harvest. This area includes fields meant for both sugar and ethanol production (LIMA et al., 2020).

The main destination for ethanol is the biofuel industry supplying the Brazilian vehicle fleet with the mixture of anhydrous ethanol for gasoline and for the engines with flex fuel technology which make up an increasingly emerging market in Brazil and worldwide (LIMA et al., 2020). As expected, Brazil is also the largest producer of sugarcane ethanol worldwide and this production is expected to have a substantial increase in the coming years as the sugarcane ethanol sector contributes significantly to the national economy (BRINKMAN et al., 2018).

Despite all the economic benefits that sugarcane and other crop cultures bring, the massive expansion of agriculture also leads to some social and ecological issues. Some of them due to drastic deforestation, social conflicts, land disputes and the degradation of the environment caused by the spread of pesticides and agricultural inputs. A great ally for helping solving some of these problems is the Precision Agriculture (PA) (MCBRATNEY et al., 2005; MILELLA; REINA; NIELSEN, 2018).

PA is a modern concept for managing agricultural activities and it has been increasingly adopted by producers in several countries. This concept is associated with research, information gathering, and the use of various technologies to analyze and monitor the conditions of planted areas in a more precise and efficient way. It is based on observation, measurement, monitoring, and rapid decision making in the face of the variability that planting crops can present (LINDBLOM et al., 2017).

The main objective of the research in the field of PA is to define a support system for the necessary decisions aiming at a better management of the crop to optimize the use of resources and inputs, while increasing the financial return (MCBRATNEY et al.,

2005; MILELLA; REINA; NIELSEN, 2018). More recently, research in this area has had a major positive impact on the growth of agricultural production (REN et al., 2020). Methodologies for improving seed quality, more efficient irrigation systems and soil quality control are just a few examples of the techniques that benefit from this research (JR; DAUGHTRY, 2018).

Technological advances in the use of unmanned aerial vehicles (UAVs) have also opened up new opportunities in the PA sector. This type of equipment allows for more effective monitoring and greater agility in cultivation. Sensors coupled to an UAV are able to collect large amounts of information about the plantation. In addition, they enable more frequent data collections and less cloud interference due to their lower flight altitude (SILVA et al., 2017; SOARES; ABDALA; ESCARPINATI, 2018; SOUZA; ESCARPINATI; ABDALA, 2018; FUENTES-PEÑAILILLO et al., 2018). The use of UAVs has also fostered the development of new and more efficient digital image processing techniques to analyze images acquired by their sensors (SILVA et al., 2017; SOARES; ABDALA; ESCARPINATI, 2018; SOUZA; ESCARPINATI; ABDALA, 2018). Most of these techniques aim to estimate the growth of the crop or to identify other important agronomic characteristics, such as nitrogen stress, water stress, new diseases, known pests and vegetation indexes.

After the initial planning and effective sowing, the scenario of the planted area may change over time due to the appearance of failures, erosion, death, and drying of part of the plantation, tipping of plants, animal interventions, among others. This makes the identification of crop lines, and how they are arranged in a region, an important task within PA. With this information, it is possible, for example, to better plan the application of inputs, thus reducing financial costs and the aggression to the environment.

Recently, Convolutional Neural Networks (CNNs) have emerged as a powerful approach to computer vision tasks. Its use has been widespread in the most diverse areas of research and it has presented relevant results in applications of classification, object detection, and facial recognition (LIU et al., 2020; SIMONYAN; ZISSERMAN, 2014; KANG et al., 2014). They have, for example, been used with great success in identifying pests in agricultural environments of complex soil structures (CHENG et al., 2017), in the detection of weeds (FERREIRA et al., 2017), detection of plant diseases (FERENTINOS, 2018) and even in the detection of flowers (DIAS; TABB; MEDEIROS, 2018).

In this work we address the problem of crop line detection and segmentation in aerial images of sugarcane plantations obtained by UAVs. First, we experimented an approach based on Genetic Algorithm associated with Otsu method to produce binarized images that were then reconstructed using a Radon transform. Then, due to some reasons including the recent relevance of Semantic Segmentation in the literature, its levels of abstraction, and the non-feasible results of Otsu associated with GA, we studied and proposed a new automatic segmentation approach based on SSN consisting in two steps. First, we use a Convolutional Neural Network (CNN) to segment the planted area into regions of

crop lines (region of interest) and unplanted soil (background). Then, we use a refinement process which aims to reconstruct and to improve the previous detected lines. This is performed in order to make the detected crop lines more uniform and to connect line fragments and isolated plants that originally belonged to the same crop line.

The remainder of this dissertation is organized as follows: In the chapter 2 we describe the fundamental concepts relevant to the understanding of the working. In the chapter 3 we present a review of the related architecture and the state of the art in detecting crop lines. In chapter 4 we describe our methodology, showing the datasets as well as the techniques used in the proposed approach. In the chapter 5 we show our experiments and obtained results and finally, the chapter 6 concludes this dissertation.

1.1 Motivation

The identification of crop lines, and how they are arranged in the planted area, in low and medium altitude images obtained by Unmanned Aerial Vehicle (UAV) is an important problem within PA. The lower cost of obtaining images by the UAV also allows farmers to monitor them more frequently. This is important because after the initial planning and effective sowing, the scenario of the planted area may change over time, such as the appearance of failures, erosion, death and drying of part of the plantation, tipping of plants, animal interventions, among others. Thus, this process of detecting the lines is important for planning the harvest, estimating the use of inputs, controlling costs, estimating production, counting plants and early correction of sowing failures. In addition, the geolocation information of the crop allows a better planning of application of inputs, thus reducing financial costs and less aggression to the environment as sugarcane represents a great percentage of all plantation crop worldwide.

Another important point is the fact that the harvest may come from autonomous vehicles and machinery. The geolocation of the crop-rows are crucial for these machines to drive themselves through the field. Plus, additional information such as which parts of the crop-row have gaps help the machinery to know which parts of the row do not need to receive inputs, suppressing their spread and thus saving money and, more importantly, lessening the degradation of the environment as some of theses substances can be harmful. Figure 1 shows an example of sugarcane crop where the crop lines were detected by an expert. Also, the exact location of the line can minimize the stump trampling and soil compaction in the seedling zone done by the own machinery. In the beneath layer we can see the image captured by the UAV. In the above layer we can see the segments of the crop lines. The green segments represent the part of the line where there is plants and the gaps of the line are described in different colors depending on the extension of the gap. The red color segments represent small gaps. Orange represents the medium segments. The yellow, large gaps.

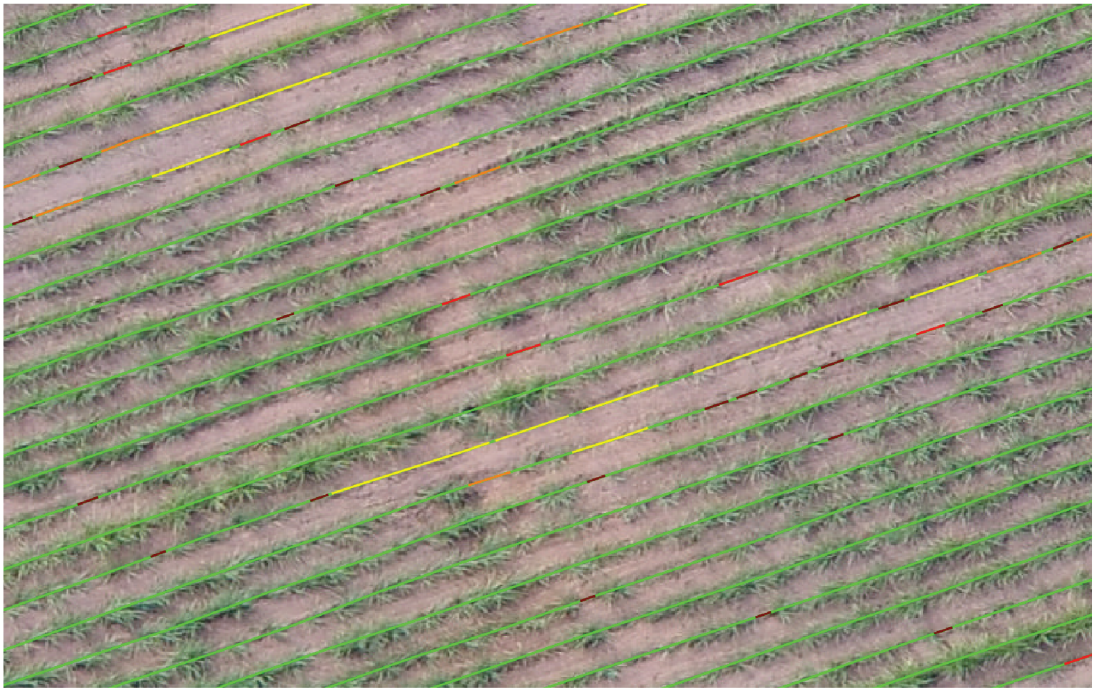


Figure 1 – Example of crop-row identification performed manually by an expert.

In addition, in Figure 2 it is possible to see an example of an autonomous tractor equipped with a device that allows PA applications. The device shows the crop-rows which it is being guided by. It also makes it possible for the tractor to suppress input spreads in case of crop gaps. This device is the CommandCenter™ Premium produced bu John Deer and compatible to some of their machinery. It also allows the sharing of information between machines working in the same field. This is a practical example of how PA can benefit from the results of this work.

Figure 2 – Example of an autonomous tractor equipped with a device showing the crop-rows which it is being guided by. CommandCenter™ Premium produced bu John Deer, extracted from https://www.agriexpo.online/prod/john-deere/product-169419-2710.html

.

1.2 Problem Description

Crop Row detection is not a trivial task and demands a hard effort to be performed manually by humans. Large plantation areas may have thousands of hectares and effective techniques to perform this task using images taken by UAVs in low or medium altitude are crucial. The frequency of satellite imagery is not enough to perform some planning related to the agricultural processes. In this case, images captured by UAVs are more suitable as it also cannot have its sensors blocked by clouds. For farmers to keep track of the Vegetation Index (VI) of the crop as well as the changes in the crop lines, frequent UAV imagery is very important.

Although the technology related to UAVs has matured in recent years, the current market still has some limitations when it comes to software solutions prepared to work with the images collected by this kind of equipment. Most of the software available today for the agricultural image processing area is still part of a group of programs and techniques built and designed for images from vehicles that capture images at much higher altitudes, such as satellite and airborne imaging. This is expected since for years these were the only image capture options on the market. With the reduction of the operational cost of UAVs and its greater use, there is a need to develop new methodologies capable of dealing with low or medium altitude images as proposed in this work.

1.3 Hypothesis

This research explores basically two hypotheses:

1. An effective binarization can be obtained from a two-class trained CNN. This segmentation is effective in sugarcane crops producing a result of quality when compared with binarization performed manually by experts.

2. Using the binarized output images from the CNN, it is possible to perform a Radon transform to achieve a refinement process that targets the reconstruction (gap filling) and enhancement of the previous detected lines. This is performed in order to make the detected crop lines more uniform and to link row fragments and isolated plant areas that originally belonged together.

1.4 Contributions

The main contributions of this work are :

- ❏ Development of a CNN training capable of classifying crop images in two classes (crop rows and background) generating binarized images. The efficient binarization can reduce drastically the cost of the post-processing step.

- ❏ Development of a framework that receives sugarcane crop images, binarize them using the CNN already trained and then perform a post-processing step based on Radon transform to perform a refinement process reconstructing and enhancing the lines making them more uniform and to linking row fragments that are supposed to belong together at the same line/row.

1.5 Thesis Organization

The remainder of this dissertation is organized as follows:

- ❏ In Chapter 2 we present the concepts of PA, image segmentation, remote sensing as well as some techniques important to the understanding of the work;

- ❏ In chapter 3 we describe some of relevant state of the art works;

- ❏ Chapter 4 states the methodology of this project, including the image acquisition process and the datasets used, a description of our first approach based on genetic algorithm, and finally our prime approach based on semantic segmentation to binarize images, as well as the post-processing step to reconstruct lines and the metrics used for the work evaluation;

- Chapter 5 describes the experimental results obtained for each dataset. Discussions and comparisons between the results obtained by GA and CNN as well as the results obtained by the Radon transform.
- Finally, in Chapter 6 we present the conclusions of this research work, its results, contributions and future work.

Chapter 2

Fundamentals

2.1 Precision Agriculture

It is indisputable that the positive economic impacts that agriculture cause locally and globally are enormous. However the massive expansion of agriculture also has led to some social conflicts, land disputes, as well as some serious ecological issues. Drastic deforestation, such as soil erosion and pollution, water scarcity, are usually caused by the overuse of pesticides and other chemical inputs (LI et al., 2020). Consequently, technological advances in the agricultural context have been crucial to promote sustainable solutions in field productivity, economical farm incomes, more security in the food section and general economic growth (LI et al., 2020), as well as reducing agricultural harmful impacts in the environment.

One solution for helping solving some of these problems is to implement advanced agricultural technologies, and the precision agriculture technologies are some of the most important of them (MCBRATNEY et al., 2005; MILELLA; REINA; NIELSEN, 2018) as they enable the more precise use of agricultural inputs (LI et al., 2020). Hence, it is clear that PA is a great ally for increasing efficiency, productivity and profitability in field operations, enhancing food security, and minimizing the unintended impacts of overuse of inputs on agricultural production systems and affecting positively the environment (LI et al., 2020). One example of PA equipment developed for farm management and tasks such as high precision positioning systems, levelling, and precision seeding/fertilizer/irrigation/harvesting can bee seen in Figure 3.

The main objective of the researching in the field of PA is to define a support system for the necessary decisions aiming at a better management of the crop to optimize the use of resources and inputs, while increasing the financial return (MCBRATNEY et al., 2005; MILELLA; REINA; NIELSEN, 2018). More recently, research in this area has had a major positive impact on the growth of agricultural production (REN et al., 2020). Methodologies for improving seed quality, more efficient irrigation systems and soil quality control are just a few more examples of the techniques that benefit from research in

Figure 3 – Example of precision agriculture equipment developed for farm management and tasks such as high precision positioning systems, laser land levelling, and precision seeding/fertilizer/irrigation/harvesting, extracted from (LI et al., 2020).

PA (JR; DAUGHTRY, 2018).

Robust vision systems are an important technology for building autonomous machinery in precision agriculture. By using this systems, it is possible to automate time-consuming work that had to be done manually in the crop, while increasing harvest yield and reducing the dependency on herbicides, pesticides and other chemical inputs, minimizing waste (BOSILJ; DUCKETT; CIELNIAK, 2018).

Despite the scientific advances that has been achieved in genetics, chemistry and the fact that machinery have contributed strongly to the evolution of agricultural technology, there has to be a significant increase of agricultural products in order to follow up the rapid increase of the global human population. According to (SYLVESTER, 2018), this product increase has to be by 70% until 2050, when it is expected for the planet population to reach 9 billion people. Meanwhile, agricultural sector has to consider important environmental challenges, namely the climate change, the limited arable land territory in Earth, and also the increasingly need for freshwater (RADOGLOU-GRAMMATIKIS et al., 2020).

The Information and Communications Technology (ICT) services can be a great strategy to create viable solutions for these critical challenges. Specifically the advent of the Internet of Things (IoT) and especially the accelerated development of the UAV technology combined with image data gathering represent a very promising perspective in terms of PA solutions to handle the mentioned challenges. In general, PA aims at adopting ICT services to combine and process information provided by multiple sources that can extract useful conclusions regarding the soil understanding, creating more efficient ways to manage the crops (RADOGLOU-GRAMMATIKIS et al., 2020).

UAV equipment allows more effective monitoring and greater agility in cultivation. Sensors coupled to an UAV are able to collect large amounts of data about the plantation. In addition, they enable more frequent data collection and less cloud interference due to

their lower flight altitude (SILVA et al., 2017; SOARES; ABDALA; ESCARPINATI, 2018; SOUZA; ESCARPINATI; ABDALA, 2018; FUENTES-PEÑAILILLO et al., 2018). The use of UAVs has also fostered the development of new and more efficient digital image processing techniques to analyze images acquired by their sensors (SILVA et al., 2017; SOARES; ABDALA; ESCARPINATI, 2018; SOUZA; ESCARPINATI; ABDALA, 2018). In addition, most of these techniques aim to estimate the growth of the crop or to identify other important agronomic characteristics, such as nitrogen stress, water stress, new diseases, known pests and VI. In the next section the UAV technology is more explored and explained.

2.2 Unmanned Aerial Vehicles

The unmanned aerial vehicles, popularly known as UAVs, are vehicles capable of flying medium and low altitudes with no need of a on-board pilot. Commonly, the flight mission of UAV is pre-defined; however a pilot can also control it through remote teleoperation commands from the ground (RADOGLOU-GRAMMATIKIS et al., 2020). There are many kinds of UAVs and they can be categorized in many different ways. A way to classify them is by their type of wing: rotary and fixed (RADOGLOU-GRAMMATIKIS et al., 2020). Some examples of rotary-wing UAVs, a more common type commonly known as 'drones', can be seen in Figure 4, while some examples of fixed-wing UAVs are shown in Figure 5.

Even though the first UAV equipments were developed for primarily military operations, the fast evolution of technologies such as imaging sensors, Inertial Measurements Unit (IMU) (MIRZAEI; ROUMELIOTIS, 2008), synthetic aperture radar (CHAN; KOO, 2008) and Global Navigation Satellite Systems (GNSS) (DOW; NEILAN; RIZOS, 2009) resulted in the development of nonmilitary UAVs fostering the development of multiple areas, such as PA, geomatics, logistics and infrastructure monitoring (RADOGLOU-GRAMMATIKIS et al., 2020).

UAVs are important for understanding crop and soil variability as it is one of the oldest challenges faced by agriculture and researches in PA (LAGACHERIE; MCBRATNEY, 2006). The use of this technology is strongly increasing not only in the agricultural but in all as they are able to perform some air operations that manned airborne may struggle to do or are not feasible such as regular crop and soil analysis (RADOGLOU-GRAMMATIKIS et al., 2020). Their use even leads to economic savings and environmental welfare, also reducing the need of a human pilot in the air, thus reducing the risk of human lives according to (OUTAY; MENGASH; ADNAN, 2020). The authors also highlight that the drones market (considering the commercial and the civilian ones) is growing at a compound rate of 19%.

The Association for Unmanned Vehicle Systems International (AUVSI) in its economic

Figure 4 – Examples of rotary-wing UAVs, a more common type popularly known as 'drones', extracted from (RADOGLOU-GRAMMATIKIS et al., 2020).

report forecasted that only in the United States it is expected that more than 100,000 jobs will be created by 2025, bringing a positive economic impact of 82 billion dollars. Seven million UAVs are already in air space today for commercial use including the agricultural domain (OUTAY; MENGASH; ADNAN, 2020). In addition, Chamola et al. in their work (CHAMOLA et al., 2020) make an interesting discussion of the worldwide UAV deployment to manage the recent COVID-19 events.

Specifically speaking about the agricultural sector, in recent years, UAV technological advances have been enabling a more effective and faster monitoring through the crop lifetime. The use of sensors attached to the UAV allows it to acquire a wide range of information concerning the crop field. They can capture and help keep track of relevant characteristics such as the VI, level of nitrogen and hydric stress, new diseases, among others.

The lower cost of obtaining images by the UAV also allows farmers to monitor them more frequently. This is important since after the initial planning and effective sowing, the scenario of the planted area may change over time with the appearance of gaps, erosion, drying and even death of part of the crop, tipping of plants, animal interventions, among other aspects. The UAVs are crucial in the detection of crop lines because of the cost-efficiency and proximity to the ground. Tasks such as planning the harvest, estimating the use of inputs, controlling costs, estimating production, counting plants, early correction of sowing failures are strongly enhanced by the UAV monitoring. Tasks that not only

Figure 5 – Examples of fixed-wing UAVs, extracted from (RADOGLOU-GRAMMATIKIS et al., 2020).

reduce the financial costs but also lead to less aggression to the environment. Beyond the imagery, the geolocation information captured by them enable the use of autonomous vehicles and machinery. The UAV used to capture the imagery for this work can be seen in Figure 6. Figure 7 shows an example of sugarcane crop raw image seen from above. This image was captured by this UAV and has no edition or distortion applied, except resizing.

Figure 6 – fixed-wing UAV SX2 made by Sensix Innovations and responsible for capturing the imagery used in this work.

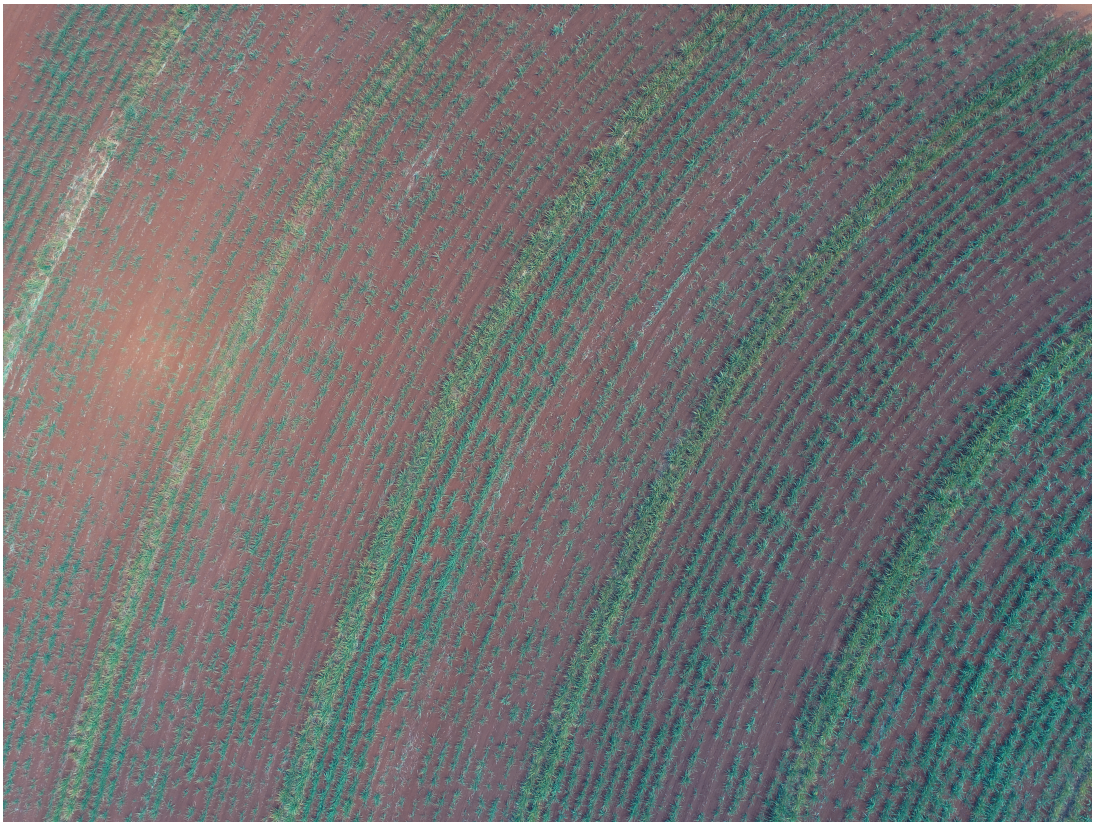


Figure 7 – Example of sugarcane crop seen from above. This image was taken by the unmanned aerial vehicle (UAV) shown in Figure 6.

2.3 Convolutional neural network

Convolutional neural networks have been showing relevant results in computer vision applications in the last years (LIU et al., 2020). Their use has been widespread in the most diverse areas of research and it has presented relevant results in applications of classification, object detection, image recognition, target classification and facial recognition (LIU et al., 2020; LIN; SHEN, 2018). They have, for example, been used with great success to identify pests in agricultural environments of complex soil structures (CHENG et al., 2017), in the detection of weeds (FERREIRA et al., 2017), detection of plant diseases (FERENTINOS, 2018) and even in the detection of flowers (DIAS; TABB; MEDEIROS, 2018).

This machine learning method is a trainable multi-layer network structure composed by stacks of other convolution neural networks. The basic structure of a CNN normally includes two layers. The first one, the convolutional layer, is the feature-extraction layer where the input of each neuron connects to the local acceptance domain of the previous layer and extract the local feature. The second common layer is the feature-mapping or fully connected layer that assumes that each computing layer of the network represents multiple feature maps, each of these maps is considered a plane, and the weights of all the neurons in each plane are equal (LIN; SHEN, 2018). Each of the these layers can be

divided in three basic stages, namely, convolution feature extraction, non-linear activation and down-sampling stage (LIN; SHEN, 2018).

However, there may be variations of CNN structures depending on the context in which it is being applied. Different types of layers can be used to build the network structure. Some of the most commonly used are the convolutional layer, the pooling layer and the fully-connected layer. For image processing purposes, for instance, the convolutional layer is responsible for applying convolutions using activation filter masks responsible for extracting the features of the image samples. The use of this type of layer is the reason for the name "convolutional neural network".

The filters are initially defined in a random way and have their values adjusted gradually at each iteration of the samples in the neural network (KANG et al., 2014). The pooling layer is responsible for progressively reducing the spatial size of the sample representation to reduce the amount of parameters, memory and computation in the network. Pooling layer operates on each feature map independently. Besides this advantage, this layer is intended to generate more robust features by reducing the sensitivity of the network to distortions present in the image. This way, a greater variety of images can be associated with the generated features, thus enhancing the classification (KANG et al., 2014). Figure 8 shows an illustration of convolution and pooling operations in CNN. Finally, the fully connected layer is responsible for performing regression and weight adjustments. The samples used as inputs to the neural network are initially divided into training and validation sets. Then, the validation set is compared with the training set in order to identify necessary weight adjustments for next iterations (KANG et al., 2014). These networks receive labeled samples as inputs. As these samples pass through the network layers by the epochs, features are extracted and the network learns, more generally, which features best represent each label (KANG et al., 2014). An example of architecture of CNN having fours convolution layers followed by pooling and one fully-connected layer can be seen in Figure 9.

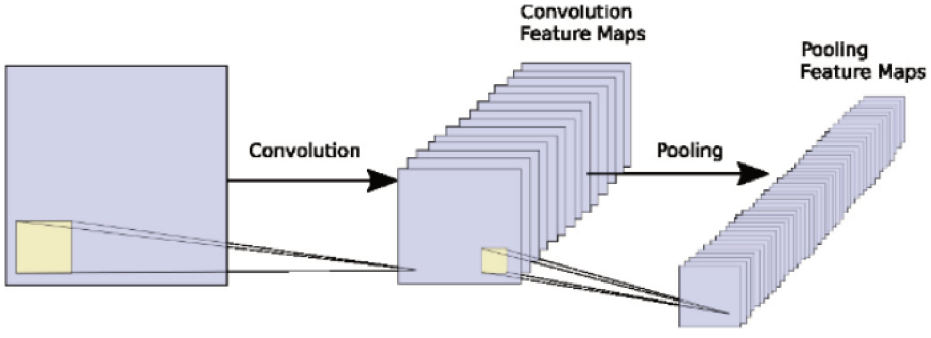


Figure 8 – An illustration of convolution and pooling operation in CNN, extracted from (DEWA; AFIAHAYATI, 2018).

In addition, it has been stated that the activation function is an important part of the convolution neural network model (LIN; SHEN, 2018). A nonlinear activation function is commonly used to map the calculated features to avoid insufficient expression problems caused by linear operation (LIN; SHEN, 2018). The activation function removes redundant data while preserving features by adding nonlinear factors. It retains "active neuron features" and maps out these features by nonlinear functions. This is the essence of the neural network to solve the complex nonlinear problem. Until now, a range of activation functions has been applied to build CNNs. Some examples are: Sigmoid (ILIEV; KYURKCHIEV; MARKOV, 2017) , Tanh (LIN; SHEN, 2018), Softplus (SENIOR; LEI, 2014), ReLu (NAIR; HINTON, 2010) and so forth.

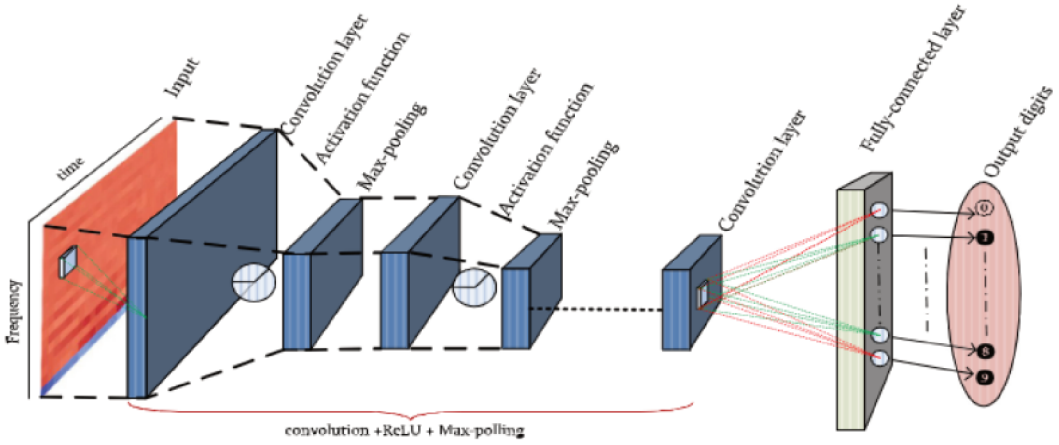


Figure 9 – Example of architecture of CNN having 4 convolution followed by pooling and one fully-connected layer, extracted from (ZADA; ULLAH, 2020).

Signmoid and Tanh are susceptible to problems of slow convergence speed and gradient dispersion problem due to the saturated non-linear activation functions. Thus, the most commonly used activation functions in CNN model are the unsaturated and nonlinear ones, such as ReLu, Softplus, Softsign (GLOROT; BENGIO, 2010). However, ReLu is even more used than the other two functions and it has multiple variations, such as Relu6 (LIN; SHEN, 2018), Elu (CLEVERT; UNTERTHINER; HOCHREITER, 2016), Leaky-Relu (ANTHIMOPOULOS et al., 2016), PRelu (HE et al., 2015), RRelu (XU et al., 2015), SiLU and dSiLU (ELFWING; UCHIBE; DOYA, 2018), etc, bringing relevant contributions to the neural network field (LIN; SHEN, 2018).

Relu is a function which demands simple calculation, easy to achieve and shows fast convergence speed. Hence, it is a feasible solution to lessen the vanishing gradient problem, also making the post-training characteristics more sparse and in accordance to the nature of biological neuron activation (LIN; SHEN, 2018). Figure 10 illustrates the ReLu activation function.

It is worth mentioning the existence of Deep Neural Networks (DNN), a now state of the art for object recognition on images, speech recognition, game playing and many

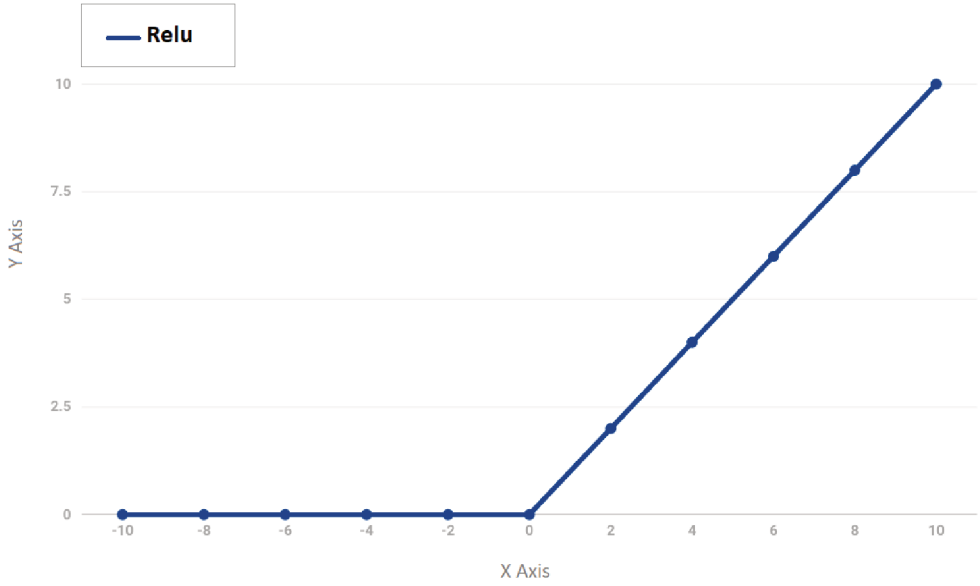


Figure 10 – ReLu activation function.

other complex tasks in which the ReLU activation function is also standard (ECKLE; SCHMIDT-HIEBER, 2019).

2.4 Image Segmentation

The aim of image segmentation is to subdivide an image into specific regions. Depending on the context applied, this can be a very complex task. This process may be continuously performed in the image until the regions or objects of interest are detected satisfactorily and grouped in their respective regions. Thus, it can be said that at the end of this process, the image will be a set of regions, connected by their borders, so that each pixel of the image is assigned a value referring to only one of the regions and there is no overlap of these regions.

Segmentation is considered one of the most difficult steps in Digital Image Processing (DIP) (NADIPALLY, 2019), since the precision of the detection directly impacts the result of the analysis phase, therefore the techniques chosen for segmentation can determine the final success or failure of the subsequent processes (GONZALEZ; WOODS, 2000).

2.4.1 Binarization

One of the most used segmentation techniques is the binarization, which is defined as the process of converting a RGB or gray-scale image that initially assumed various color intensity values. After this conversion, the binarized image will assign only two values

that are normally black and white (VERMA; PARIHAR, 2017). These two values labels the regions of the image as regions of interest and background (SEZGIN; SANKUR, 2004).

There are several ways to perform the binarization, but the simplest technique is the use of a threshold value to classify the pixels of the source image based on that value. All pixels with values greater than this threshold are defined as white or intensity 255, and all other pixels will receive an intensity value of 0 or black. The threshold value in a binarization can be set based on the region of the image or globally (VERMA; PARIHAR, 2017). One of the most common approaches to calculate a threshold is through the use of Otsu method. This algorithm basically assumes that the image has two classes of pixels following a bi-modal histogram (background and foreground). Then, the ideal threshold is reached by separating the two classes so that the combined intra-class variation is minimized (OTSU, 1979). Usually, finding a single correct threshold that satisfactorily represents the entire image can be quite difficult or even impossible. Thus, applying local thresholds to smaller regions of the image may be the best option (GONZALEZ; WOODS, 2000).

2.4.2 Semantic Segmentation

Another segmentation technique that has shown great efficacy is the semantic segmentation method (NEMOTO et al., 2020). It can be defined as the act of classifying each pixel of an image into a class (BRAS et al., 2020). Semantic segmentation is one of the most important tasks in machine learning and computer vision. This technique has been used with success in many applications, namely: autonomous driving, medical diagnosing, image editing, among others (FENG et al., 2020; STAN; THOMPSON; VOORHEES, 2020; BRAS et al., 2020). Its relevance has been increasingly growing in the last years due the resumption of convolution neural networks and its fast development (KRIZHEVSKY; SUTSKEVER; HINTON, 2012; HE et al., 2016).

Convolutional neural networks (CNNs) have received great attention, among other machine learning and semantic segmentation techniques. Nonetheless, unlike traditional machine learning methods, a CNN does not demand manual feature extraction on an image, as it automatically incorporates feature extraction engineering into the training steps. Features are extracted in convolutional layers, while the classification process is performed in the fully connected layers based on the feature values. CNNs were initially applied for object recognition (KRIZHEVSKY; SUTSKEVER; HINTON, 2012); lately, a range of semantic segmentation works in the literature are based on Fully Convolutional Network (FCN), which, according to (LONG; SHELHAMER; DARRELL, 2015).

An FCN is a network composed of only convolutional layers and it can be applied to semantic segmentation training (NEMOTO et al., 2020). It arose from the CNN concept (LONG; SHELHAMER; DARRELL, 2014) and replaces the fully connected layer with convolution layer to support arbitrary input sizes (SANG; ZHOU; ZHAO, 2020).

Despite the impressive advances these networks have made, there are still two main challenges associated with this task to be worked out. The first one is how to correctly capture rich contextual information and features to determining confusing classes. The second one is related to how to accurately recover feature map resolution to improve spatial performance (SANG; ZHOU; ZHAO, 2020). Figure 11 shows an example of a semantic segmentation process performed in some images, their results, as well as their classifications and respective percentage score per segment/label. In the sequence we explain three of the most used semantic segmentation networks (U-net, PSPNet and LinkNet). We chose these networks due to the great results they have been showing in literature. Further explanation regarding CNNs is provided in section 2.3

2.4.2.1 U-net

U-net was initially proposed by (RONNEBERGER; FISCHER; BROX, 2015) for tasks that needed precise segmentation, but had few examples available for training, such as medical images. It consists of two main segments: contracting and expanding paths, giving to the network the shape of an "U", which justifies its name.

The contracting path consists of the repeated application of blocks of two 3×3 convolution layers (each followed by a ReLU unit) and a 2×2 max-polling layer. After each block, the number of filters doubles so that the network can learn the more complex structures.

In the expanding path, each block consists of two 3×3 convolution layers (each followed by a ReLU unit) and a 2×2 up convolution layer. It also concatenates the high-resolution features maps from the respective step of the contracting path in order to ensure the proper reconstruction of the image. After each block, the number of filters halves. This is necessary due to the loss of the border region in each convolution.

Both contracting and expanding path present the same number of blocks. After expanding path, the resulting feature map passes through a 3×3 convolution layer where the number of feature maps is equal to the number of classes in the segmented image.

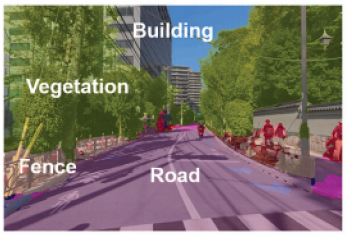
2.4.2.2 PSPNet

The Pyramid Scene Parsing Network (PSPNet) (ZHAO et al., 2017) has as its fundamental principle the use of global information from the image by extracting context information in each scene. Its architecture consists of a fully convolutional network, being the ResNet (HE et al., 2016) used in (ZHAO et al., 2017). From the feature map, four pooling layers of different sizes are applied, thus generating four feature sub-maps. Subsequently, the network applies a 1×1 convolution to reduce the maps' dimensionality, which are enlarged through a bilinear interpolation to return to the size of the original feature map. Finally, the feature maps are concatenated and a convolution is applied to obtain the prediction map.

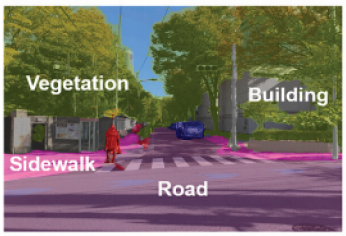
2.4.2.3 LinkNet

LinkNet (CHAURASIA; CULURCIELLO, 2017) is a SSN presenting good performance for both accuracy and processing time. It consists of two types of main residual blocks: encoder and decoder. First, the network applies a 7×7 convolution layer (stride of 2), which is followed by a 3×3 max-pooling layer. After that, the network starts the encoder blocks. Each encoder block consists of two groups of 3×3 convolution layer, which reduces the feature map by a factor of 2. The output of each group is added to its input, and there are a total of 4 encoder blocks in the network. In the sequence, the output of the last encoder block is fed to the group of 4 decoders. Each decoder consists of a 1×1 convolution layer, a 3×3 full convolution layer, which increases the feature map by a factor of 2, and another 1×1 convolution layer. After each decoder block, the output is added to the input of the respective encoder.

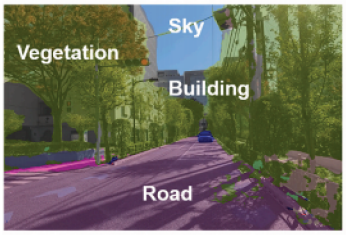
The network structure end with three layers that perform, respectively, 3×3 full convolution, which increases the feature map by a factor of 2; a 3×3 convolution layer; and again a 3×3 full convolution with an increase factor of 2.



Segments	%
Vegetation	43.39
Road	30.91
Building	9.62
Fence	6.63
Sky	4.17
...	...



Segments	%
Vegetation	45.32
Road	32.55
Building	9.92
Sidewalk	4.69
Sky	3.40
...	...



Segments	%
Vegetation	56.47
Road	25.93
Building	6.13
Sky	5.42
Fence	2.64
...	...

Mean value of the segmented components at the intersection

Segments	%
Vegetation	48.39
Road	29.80
Building	8.56
Sky	4.33
Fence	3.56
Sidewalk	2.15
Pole	0.97
Wall	0.58
Person	0.55
Bicycle	0.48
Car	0.36
Rider	0.15
Terrain	0.08
Traffic Sign	0.05
Truck	0.01
Motorcycle	0.00
Traffic Light	0.00
Bus	0.00
Train	0.00

Figure 11 – Example of a semantic segmentation performed in some images, their results, as well their classifications and respective percentage score per segment/label. Extracted from (NAGATA et al., 2020)

2.5 Genetic Algorithm

Introduced by John Holland in 1960, a genetic algorithm (GA) is a heuristic inspired by Charles Darwin's theory of natural evolution. This algorithm reflect the process of natural selection where the fittest individuals are selected for reproduction in order to produce offspring of the next generation. Genetic Algorithms are commonly used to generate high-quality solutions to optimization and search problems by relying on bio-inspired operators such as mutation, crossover and selection (MITCHELL, 1996).

A genetic algorithm starts with an initial population of individuals, where each one is assumed to be a solution to the problem to be solved. Each possible solution is represented as a chromosome $g(i)_j$, where $i = 1, \ldots, N$ is the i-th gene of an individual j, $j = 1, \ldots, M$, of a population of M individuals.

Then, a fitness function f is applied to determine how an individual j fits in the relation. Each individual will have a fitness score $f(j)$ in the end and this score will be used to determine the probability of that individual being selected for reproduction, where individuals containing high fitness have more chance to be selected. At the selection stage, the fittest individuals pass their genes (parameters) to the next generation. Two pairs of parent-individuals, A and B, are selected based on their fitness scores, $f(A)$ and $f(B)$, and have their parameters combined through a process called Crossover. This process is one of the most significant phase in a genetic algorithm. A crossover point p, $1 < p < N$, is randomly picked from within the parameters for each pair of parents to be mated. The genes of the parents, A and B, are combined in order to produce two new individuals, C and D, called offspring:

$$g_C = [g(1)_A, \ldots, g(p)_A, g(p+1)_B, \ldots, g(N)_B] \quad (1)$$

$$g_D = [g(1)_B, \ldots, g(p)_B, g(p+1)_A, \ldots, g(N)_A] \quad (2)$$

In certain individuals of offspring, some of their genes can be subjected to a mutation with a low random probability, r. This implies that some of the genes in the chromosome $g(i)_j$ that represents the individual can have its value modified. A mutation occurs to maintain diversity within the population and to prevent premature convergence of the population. The mutation depends on how the gene is used to represent the data. For binary genes, selected genes may have their bits flipped. For non-binary genes, for example, a unit Gaussian distributed random value can be added to the gene.

Finally, the algorithm terminates if the population converges to a solution, i.e., the population does not produce offspring significantly different from the previous generation. When it occurs, it can be said that the genetic algorithm has provided a set of solutions to the problem (EIBEN; SMITH, 2003).

In our work, we propose to use a genetic algorithm (GA) to estimate the best parameters of a kernel mask used to segment crop lines. We opted for this algorithm as it presents a large range of application in several areas. For example, GA is used to

select optimal parameter values for image defogging algorithms, a technique commonly used to correct image degradation produced by many outdoor working systems (GUO; PENG; TANG, 2016). A modified genetic algorithm (HEMANTH; ANITHA, 2019) is proposed to minimize the random nature of conventional GA. The authors proposed this modification aiming to improve medical image classification, more specifically, the classification of abnormal brain images from four different classes. This modification presented promising results and achieve 98% accuracy in the given problem. In (GHOSH et al., 2016), GA is used for prostate automatic segmentation on pelvic images. The authors propose a framework where GA evaluates candidate contours by combining representations of learned information (e.g., known shapes and local properties). Visual analysis of the three dimensional segmentation indicates that GA is a feasible approach for pelvic CT and MRI image segmentation.

2.6 Hough Transform

The Hough Transform (HT) was introduced and patented by P.V.C. Hough in 1962 (HOUGH, 1962). The primary original application of this technique was to detect lines and arcs in the area of physics, more specifically in particle detectors photographs. However, along the years various approaches have been used not only to improve this technique, but also to propose new applications for it (BELTRAMETTI; ROBBIANO, 2012). Despite its initial use, in the last years, this technique has been extended to identifying positions of arbitrary shapes being mostly used for DIP and image analysis, especially for the detection of simple curves such as lines, circles (BELTRAMETTI; ROBBIANO, 2012) and ellipses (KHADANGA; JAIN, 2020).

As said, this technique is capable of detecting arbitrary images shapes. However, the parametrized description of the shape in question must be given (BELTRAMETTI; ROBBIANO, 2012). Despite the great results in performing linear detection, it has been stated that this technique requires a huge memory use and it has a considerable computational complexity (EL HAJJOUJI et al., 2020).

In the last years, in particular in problems of recognition of special shapes in medical and astronomical images, much effort has been made to apply the above described procedure to the detection of more complicated objects, in particular special algebraic plane and space curves (BELTRAMETTI; ROBBIANO, 2012).

HT is a technique that aims feature extraction (DUDA; HART, 1972) and the mechanism behind this method uses a “voting” process to find imperfect objects contours encountered in the image that represent the class of shapes described in the parametrized description (KHADANGA; JAIN, 2020). The voting scheme is carried out to obtain the values in the parameter space that represent the line with more probability to exist. Concurrently, when the shape aimed to be detected is a straight line, the process, also

known as Standard Hough Transform (SHT), can be expressed as the Cartesian Equation 3 (VARUN et al., 2015).

$$y = mx + c \tag{3}$$

where, for each point (x, y) in the image space, a potential line passing through it has m as its slope, while c the y-intercept of the line.

For a more robust computation process, Duda and Hart (DUDA; HART, 1972) proposed the use of a different parameter space ρ-θ (rho-theta) aiming a better line detection using HT, as shown in Equation 4:

$$y = \left(-\frac{\cos\theta}{\sin\theta}\right)x + \left(\frac{\rho}{\sin\theta}\right) \tag{4}$$

where θ is the angle subtended by the line in relation to one of its axis (usually x-axis) and ρ is the perpendicular distance of the line from the origin, as shown in Figure 12a. Hence, Equation 4 can be transcribed as shown in Equation 5 (VARUN et al., 2015).

$$\rho = x\cos\theta + y\sin\theta \tag{5}$$

The general idea of HT can be illustrated by considering a point (x, y) in the spatial domain, obtained from the edge detection process. Every edge detected point is considered to be intersected by an infinite number of lines that subtend different angles in relation to the x-axis (usually) and vary in their algebraic distance from the origin (Figure 12b). Each of the points votes for a number of values (ρ, θ) for all the lines passing through it. The line receiving the predominance of the votes is the line of interest and it then characterizes a line in an edge detected image (VARUN et al., 2015).

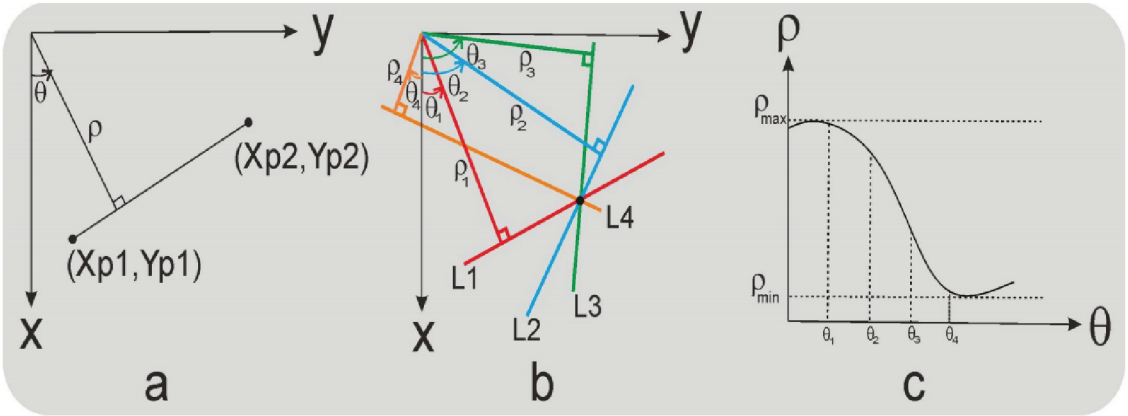


Figure 12 – Basic Principle of HT. (a) shows representation on equation of straight line, (b) illustrates intersection of many lines to a point, (c) transformation of point in image space to polar space, extracted from (VARUN et al., 2015).

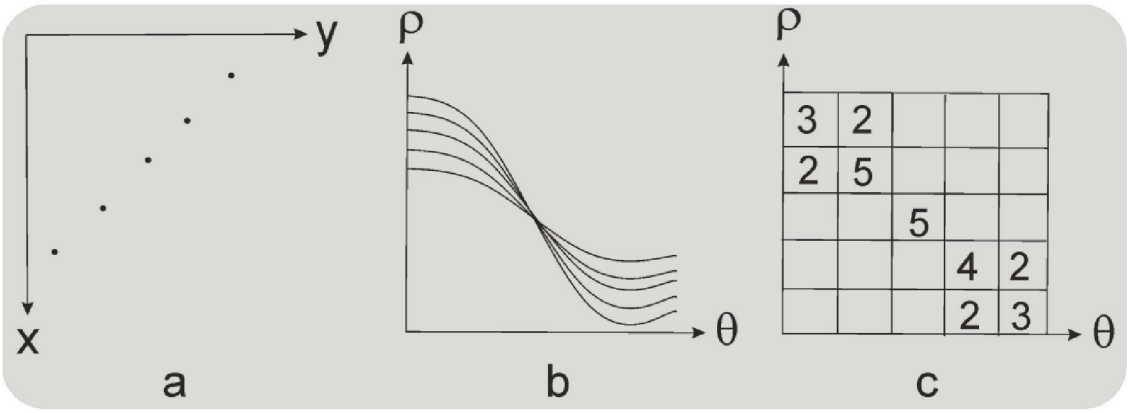


Figure 13 – Working of HT as feature extractor. (a) shows a set of points in image space, (b) illustrates convergence of points in image space to sinusoidal waves in polar space, (c) represents the accumulator space, extracted from (VARUN et al., 2015).

Similar to spatial domain, the polar domain is also made discrete in terms of (ρ, θ). For a particular value of x and y, i.e. a specific point in spatial domain, when the values of θ is varied, the ρ values change in a sinusoidal fashion (Figure 12c). Thus, for every point in the edge detected image, it gives a sinusoidal wave in the accumulator space and there exists a particular region which begins to accumulate more number of votes, resulting in a line with many feature points as shown in Figure 13. To detect these feature vectors, the maxima of the accumulator space is considered to compute the feature parameter extracted (VARUN et al., 2015).

2.7 Radon Transform

The Radon transform is an important mathematical tool in the area of image analysis (HASEGAWA; TABBONE, 2016; HASEGAWA; TABBONE, 2014). It was proposed by Johann Radon in 1917, and it represents projection of a function obtained in a specific orientation, being considered the basis of classical tomography (BARRETT; SWINDELL, 1996).

Given $f(x, y)$, $p_{\,phi}(x')$, we define the Radon transform as the integral on the line that is parallel to the y' axis at distance x' the origin. It is possible to get (x', y') coordinates by rotating (x, y) coordinates with an angle ϕ:

$$x' = x\cos\phi + y\sin\phi \tag{6}$$

$$y' = -x\sin\phi + y\cos\phi \tag{7}$$

e

$$x = x' \cos\phi - y' \sin\phi \tag{8}$$

$$y = x' \sin\phi + y' \cos\phi \tag{9}$$

Denoted by the operator $\Re$, the Radon transform is mathematically defined as:

$$p_\phi(x') = \Re[f(x,y)] \tag{10}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(x\cos\phi + y\sin\phi - x')dxdy \tag{11}$$

$$= \int_{-\inf}^{\inf} f(x'\cos\phi - y'\sin\phi, x'\sin\phi + y'\cos\phi)dy' \tag{12}$$

Since the one-dimensional projection of the $f(x,y)$ function at the ϕ angle is defined as $p_\phi(x')$, the Radon transform calculates the integral of a two-dimensional image on the axis y'.

Chapter 3

Related Work

As already mentioned in this dissertation, the popularization of UAVs has been enabling authors to address many existing problems in agriculture. One of these problems is to accurately identify existing crop lines in a region and, consequently, its geolocation, arrangement, as well as how failures and gaps are distributed in the field. Literature presents some existing approaches to address this problem, many of them are commonly based on the Hough transform (HT) (HOUGH, 1962), which is widely used in problems involving the detection of known objects, such as straight lines and circles (ILLINGWORTH; KITTLER, 1988; HASSANEIN et al., 2015). Nonetheless, other techniques in the state of the art have also been used to study problems similar to the one exposed in this work. Hence, in the remainder of this chapter, we will discuss some of classical approaches that have been taken mostly based on HT technique as well as some new and very interesting approaches applied to similar problems that, by the way, also use CNN as our work.

3.1 Hough Transform

Considering the imaging processing techniques available, the Hough transform (HOUGH, 1962) figures as an initial clear choice to detect crop lines. As presented in (ILLINGWORTH; KITTLER, 1988) this technique is widely used to identify fixed parameterized shapes formed by points on images. The basic Hough transform works well for straight lines and circles, but it can also be used to find arbitrary shapes (BALLARD, 1981). However it has the disadvantage of requiring that the object shape to be known in advance, which limits its application. Nonetheless, a number of solutions were proposed using the Hough transform as basis.

Some authors along the years have presented adaptations of the Hough transform to deal with images captured by cameras attached to tractors, such as (LEEMANS; DESTAIN, 2006). In this work a row localization method was proposed using an adaptation of the HT. This approach works on video sequences captured on ground level by

a camera mounted on a tractor. The method combines a series of techniques in order to reveal the rows or to eliminate unevenness such as shadows, since the HT requires bright lines on a dark background. The method is specifically tuned to deal with early sowing and it was tested only with chicory rows at very low altitude.

Another attempt to work with images provided by cameras attached to tractors can be found in (SØGAARD; OLSEN, 2003). In these cases, also, the perspective of the images captured is totally different and much closer to the ground when compared to the UAV imagery. With this image acquisition method, each image represents a closer crop of the field, making the crop-row segment present in each picture approximated to straight lines. This allows the use of HT according to the authors. However, due to the very low altitude of the images, the volume of image crops necessary to cover the whole field is obviously much higher. Figure 14 shows the first attempts of Søgaard and Olsen (SØGAARD; OLSEN, 2003) of capturing crop-rows by using a vehicle pulled by a human being having attached to it a camera, a computer and a batery for power supply. In Figure 15 we can see their later experimental set-up as an example of a tractor carrying the camera and showing the camera. While in Figures 16, instead of images of the vehicles and its capturing equipment, the actual crop images captured by the cameras are shown. The first one was used for Søgaard and Olsen's work (SØGAARD; OLSEN, 2003). The second, for the work of Leemans and Destain (LEEMANS; DESTAIN, 2006).

Figure 14 – A hand-operated vehicle carrying the camera, computer and a battery for power supply, extracted from (SØGAARD; OLSEN, 2003).

Also, in (LEEMANS; DESTAIN, 2006) the authors evaluate some closer crops of the gathered images and how different elements could be observed during the harvesting and under different lighting conditions. One important aspect to observe is that the tractor

Figure 15 – A tractor carrying a camera to capture crop-row images. The camera can be seen in the upper part of the image pointing forward-downward to the left of the tractor's left rear wheel, extracted from (SØGAARD; OLSEN, 2003).

itself carrying the camera can block the sun-light and produce shadows in the images that may affect the analysis.

This other work from Ji and Qi (JI; QI, 2011) also presented a crop line detection algorithm based on a variation of HT technique - Random Hough Transform (RHT). However, using manual cameras to produce reference images instead of wheeled vehicles or tractors (Figure 17). Nonetheless, the classic HT model also has some performance limitations mainly due to the real-time execution context. The project uses the I3 color feature to generate a gray image which presents a bimodal histogram (crop and soil), which is segmented using a threshold value. Then, they use an algorithm based on the Gradient-Based RHT on the resulting image to extract the center of the crops lines. The experiments presented by the authors show that the algorithm was efficient in regions of shadows and irregular noises; however, it showed limitations while attempting to reconstruct the lines in cases of crop gaps. The authors, don't specify how the threshold is computed or if it is a user-defined value and their approach works only on ground images taken by a hand camera at very low altitude ($\approx$ 1m), where the lines can be considered locally straight.

In 2012, BELTRAMETTI and ROBBIANO (BELTRAMETTI; ROBBIANO, 2012) proposed an algebraic approach to generalize the notion of the classical HT technique

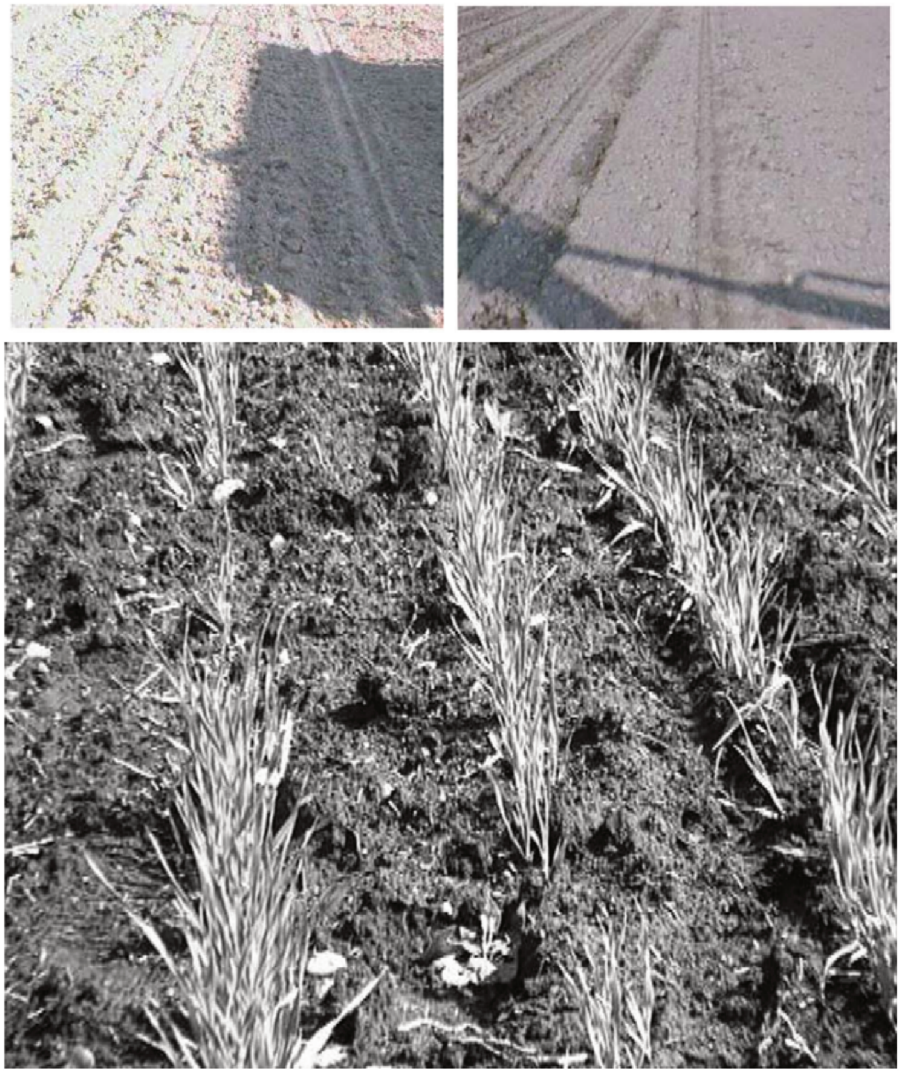


Figure 16 – Examples of images captured by cameras attached to tractors. The first two are raw images used in (LEEMANS; DESTAIN, 2006). The last one is an image captured by the camera shown in Figure 15 and changed to its grey-scale representation for typographical reasons, extracted from (SØGAARD; OLSEN, 2003).

using reduced Gröbner (KREUZER; ROBBIANO, 2000) bases of flat families of affine schemes in order to make it able to detect not only straight lines and circles, but also other structures. They introduced and developed the theory called Hough regularity. A great approach to diversify the use of HT despite it still being a complex computational technique to be applied in large fields.

It is worth to highlight that sometimes crops follow the terrain, are hindered by obstacles or any other arbitrary unknown geographical feature. To solve this problem, Soares et al. (SOARES; ABDALA; ESCARPINATI, 2018) used a tiling scheme to make feasible to apply the Hough transform in images obtained using an UAV. This enabled them to obtain an acceptably approximation of the rows inside each tile to straight lines. Results

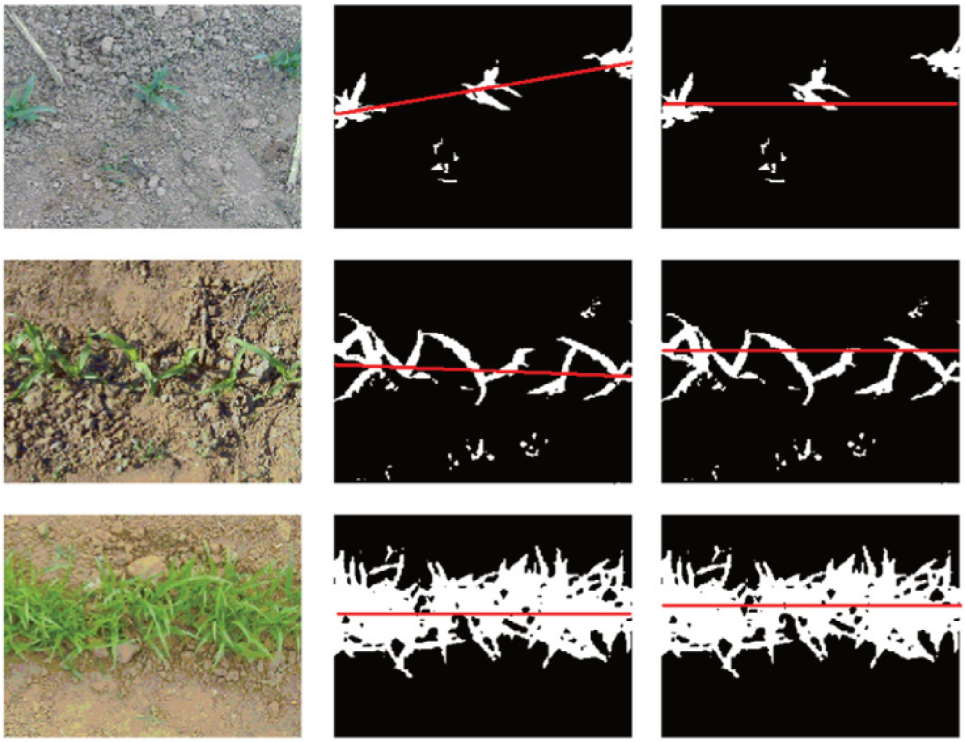


Figure 17 – Examples of crop images used for RHT line detection technique. Extracted from (JI; QI, 2011).

indicate that this approach successfully approximates real plantation rows. However, it created an undesirable aliasing effect in the lines, besides it generates a greater number of false positives lines on the final results.

3.2 Otsu Method

A framework to guide a mobile agricultural vehicle is proposed in (GUERRERO; RUZ; PAJARES, 2017). This framework uses a Greenness Index Combination technique to segment crop rows in images under perspective projection. From the segmented images, the framework computes weeds density and uses this information to control the vehicle, avoiding that it passes over the crop. Besides, the authors aim to achieve a correct orientation of the lines, to accomplish a precise guidance during path following as well as to determine the weed density and overlapping. Figure 18 shows the framework architecture proposed by the authors. Also, an example of an image processed by this framework is shown in Figure 19.

Similarly, the work in (GARCÍA-SANTILLÁN et al., 2018) also deals with images captured by a camera mounted on a moving tractor. The methodology uses the Excess Green (ExG) index for greenness identification. The authors considered a double thresholding approach based on Otsu's method for segmentation: the first threshold discriminates green plants and soil, while the second discriminates between crops and weeds.

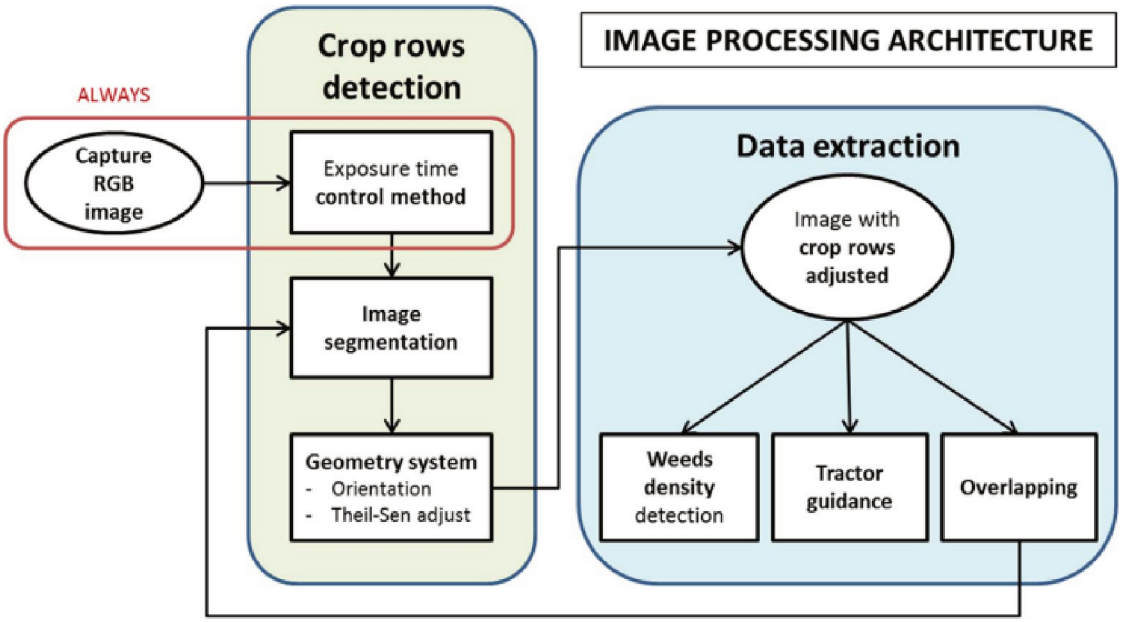


Figure 18 – Framework architecture from (GUERRERO; RUZ; PAJARES, 2017).

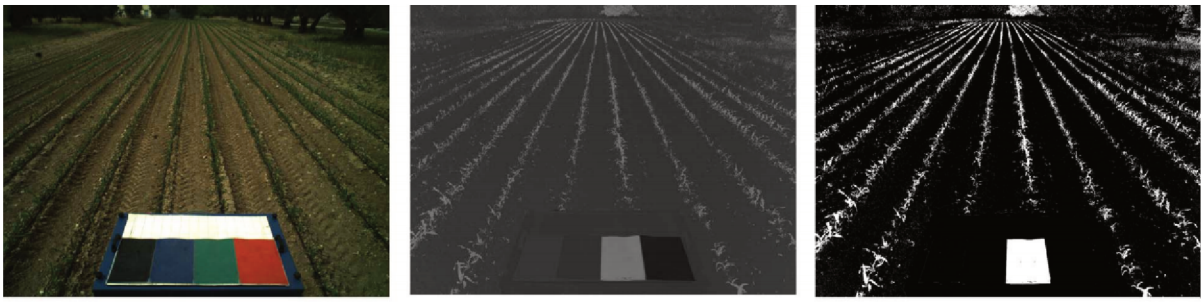


Figure 19 – Example of an image processed by the framework in Figure 18: (a) Original image; (b) Greenness Index Combination index from (a); (c) Binary image after Otsu thresholding. Extracted from (GUERRERO; RUZ; PAJARES, 2017).

Morphological operations are used to extract candidate points for the crop rows. The proposed image processing method architecture is shown in Figure 20. The authors report good results to detect curved and straight crop rows on images of maize fields obtained at very low altitude.

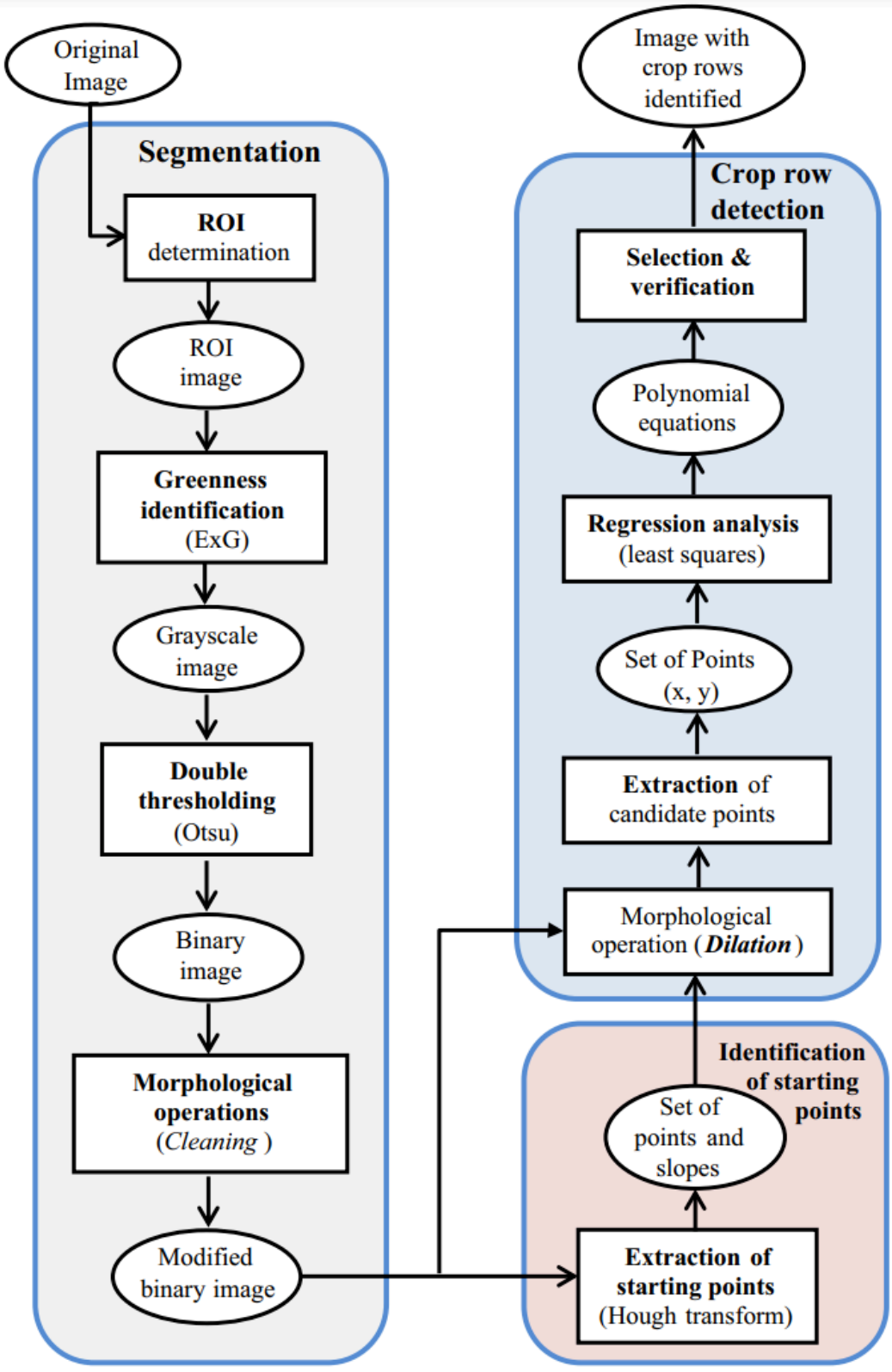


Figure 20 – Proposed image processing method architecture. Extracted from (GARCÍA-SANTILLÁN et al., 2018).

3.3 Convolutional Neural Networks

More recently, some authors have proposed deep learning models applied to crop analysis. The aim of the work presented by Pang et al. (PANG et al., 2020) is an example. In their paper, they explore one of the most common ways farmers assess plant growth conditions and management practices through seasons: the stand counting process. An alternative and more conventional method for early-season stand count is by manual human inspection. However, clearly it is laborious, very time-consuming and limited, being even unfeasible depending on the extension of the area. They also highlight the use of UAV imagery due to its low-altitude and high spatial resolution imagery to assist decision making. In their project, they proposed a system capable of using geometric descriptor information and deep neural networks to determine early-season average maize stands count. Even thought, the authors used a deep neural network, their approach was focused on the maize stand count itself and not on the structure of the crop row, either in finding the correct line to isolated plants nor its reconstruction. Figure 21 shows early-season maize stand count determination system proposed by the authors.

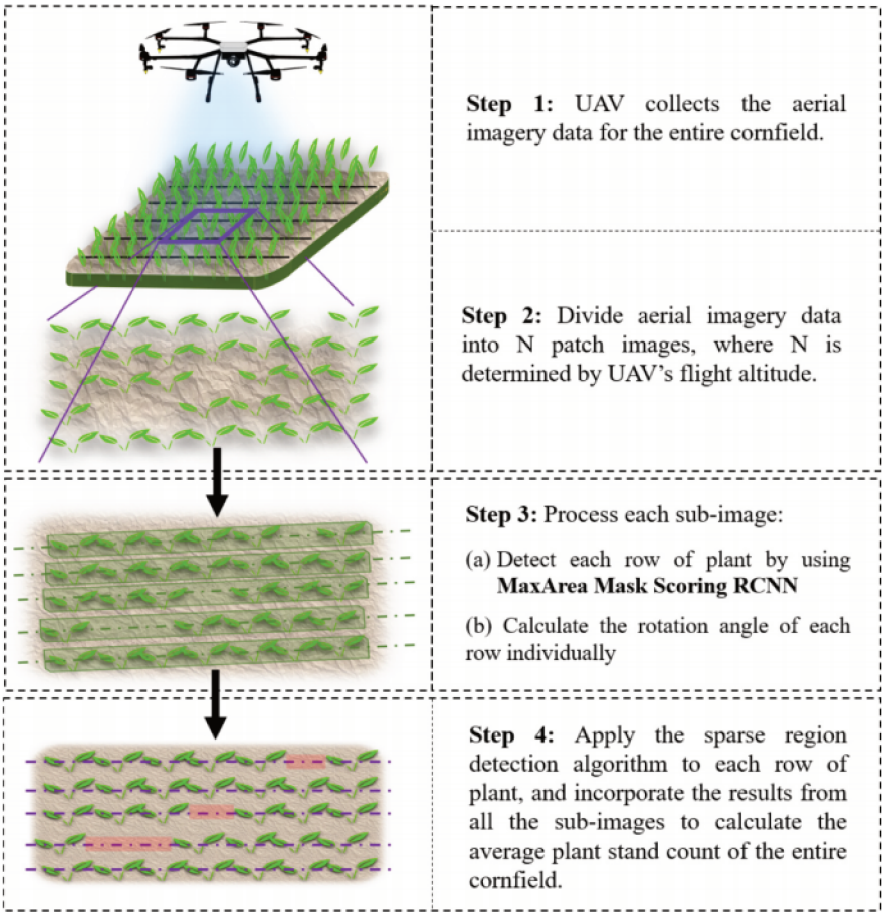


Figure 21 – Early-season maize stand count determination system proposed by Pang et al.. Extracted from (PANG et al., 2020).

In (BAH; HAFIANE; CANALS, 2020), the proposed approach uses two convolutional neural network (CNN). First, a S-SegNet model is trained to segment the Red, Green, Blue (RGB) images (Figure 22), thus resulting in a binary image with the most probable location of the crop lines. In the sequence, CNN based Hough transform (HoughCNet) is used to extract the pixels that form the longest lines in the image (Figure 23), which the authors consider as the best candidates to crop lines.

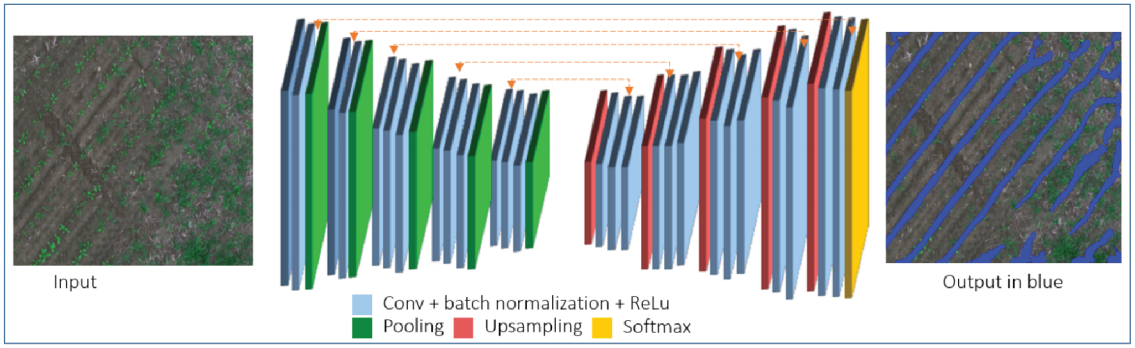


Figure 22 – Encoder-decoder architecture of SegNet. Extracted from (BAH; HAFIANE; CANALS, 2020).

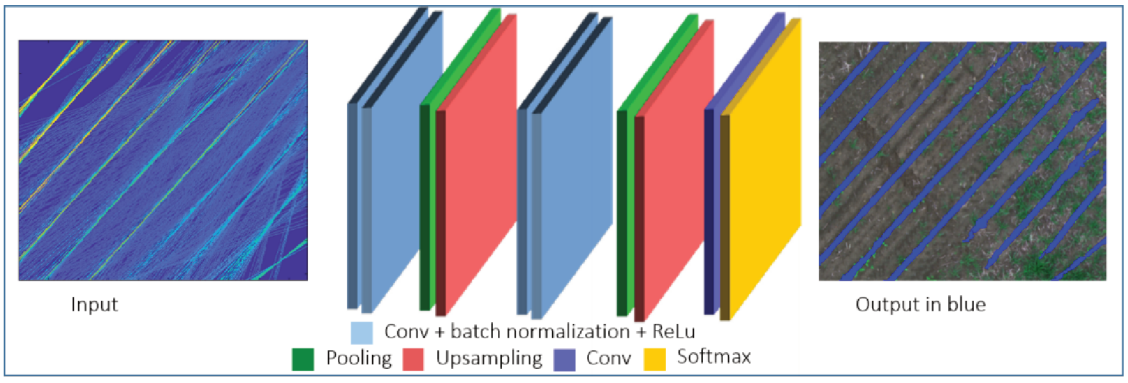


Figure 23 – Architecture of the HoughCNet network. Extracted from (BAH; HAFIANE; CANALS, 2020).

3.4 Other Techniques

Some very recent approaches have also applied modern computer vision into Precision Agriculture. In (OSPINA; NOGUCHI, 2019), for instance, the authors present the use of remote sensing for crop disease detection and management. The article provides an interesting overview of remote sensing and PA technologies that have been used for detection and management in the crop context. The paper illustrates how telephoto imagery combined to wide angle image technology have been used for detecting and mapping cotton

root rot that, despite being related to another culture crop, it gives us a great overview of the state of the art in terms of modern technology and techniques applied in the field and alternative approaches.

The work in (VIDOVIC; CUPEC; HOCENSKI, 2016) also uses ExG index to segment images under perspective projection. It combines image evidence and prior knowledge of the geometric structure of the crop using a dynamic programming technique. This is performed to detect regular patterns related to the appearance of crop rows, both straight and curved. This method is used as base algorithm in (BASSO; FREITAS, 2020), where an entire guiding system for spraying UAVs is proposed. The idea is to identify the crop rows during the UAV flight and to use this information to generate the driving parameters sent to the flight controller. The authors claim that their approach is able to deal with curved crop rows by dividing the curves into segments of straight line.

In (MONTALVO et al., 2012) yet another method of crop row identification is presented. It was devised to work on crops with high incidence of weeds and with camera mounted on ground vehicles. According to the authors the image processing of this work consists of three main stages: image segmentation, double thresholding based on the Otsu's method, and crop row detection. The image segmentation is based on the application of a VI, the double thresholding handles the separation between weeds and crops and the crop row detection applies least squares linear regression for line adjustment. Also, in (MONTALVO et al., 2013), the authors explore another technique based on image segmentation procedures that works independently of the loss of greenness. First, they perform a combination of vegetation indices and apply a first Otsu thresholding. Then, they select black pixels and apply a second Otsu thresholding. Lastly, the histogram obtained from pixels belonging to the background and masked plants is thresholded by applying a last stage of Otsu. However, the main focus of these works falls on the image processing task, since the crop row identification became trivial due the fact that only straight lines are expected.

A different approach (SOUZA; ESCARPINATI; ABDALA, 2018) was used for images of sugar cane and coffee taken on low altitude (≈100m). They assumed the images are well segmented and, from a cloud of points representing the plants in a field they subdivided such points into lines representing the true plantation rows. The proposed algorithm is a two-fold process. First the points were subdivided into preliminary lines by a procedure inspired in the formulation of hierarchical clustering. Afterward the lines are pruned to correct for imprecisions introduced by field specificities and the image preprocessing. They achieved good results, but the total dependence of the segmentation process represents a real problem for the algorithm.

3.5 Final Considerations

In general, most of the works in the current literature still focus on straight lines based on the HT technique (HASSANEIN et al., 2015). One of the problems with the use of the Hough transform is that the shape of the detected object must be previously known and described, which limits its application. In general, the crop lines can have irregular shapes as they follow the terrain, thus creating curves and non-uniform shapes. To circumvent this deficiency, other possibilities have been explored. However, most of these approaches either require processing of a crop line at a time, or focusing on plant counting, or even focus on different crop cultures. In sugarcane crops, for example, its is common that plants of different lines intersect to each-other, which impairs the detection and correct reconstruction of the line.

Chapter 4

Methodology

4.1 Datasets

To evaluate the proposed methodology we used four test images of different sizes. These images are mosaics of aerial images that represent areas of sugar cane cultivation and that contain planting lines of different ages and widths. These images were acquired using an eBee SenseFly mapping drone. We used a senseFly S.O.D.A. camera of 1in Sensor, 5472×3648 pixel resolution and RGB lens F/2.8-11, 10.6 mm. Each pixel in the image represents $5cm$ of ground (Ground sample distance - GSD - of 0.053 meters).

Figure 24 shows a preview of the four test images used in the experiments (named Dataset A, B, C, and D, respectively). It is important to notice that each image has the crop lines of its entire region segmented by an expert, as illustrated in Figure 25. Note that the rows can vary in width depending on the age of the crop and the level of success of the planting process.

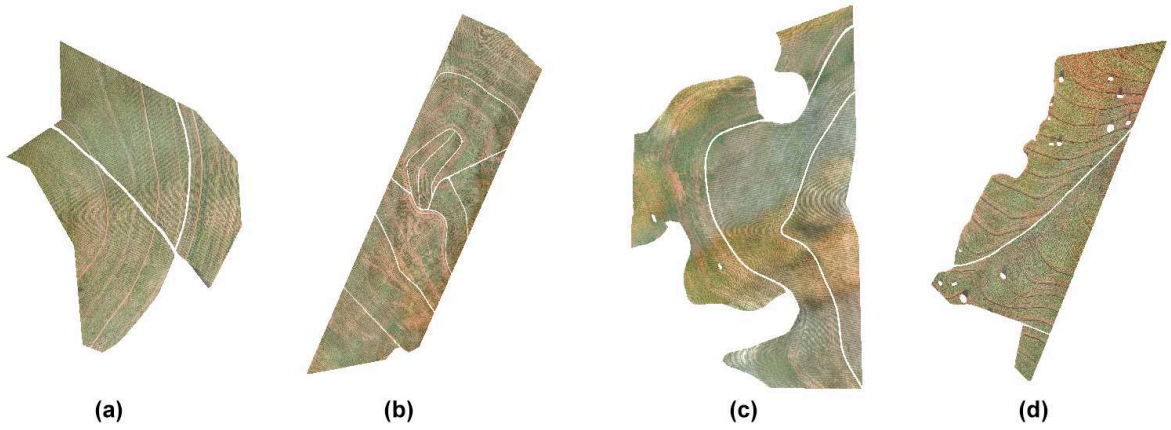


Figure 24 – Test images used to evaluate our approach and their respective sizes: (a) 11180×8449; (b) 19833×30255; (c) 17497×10771; (d) 16677×24181.

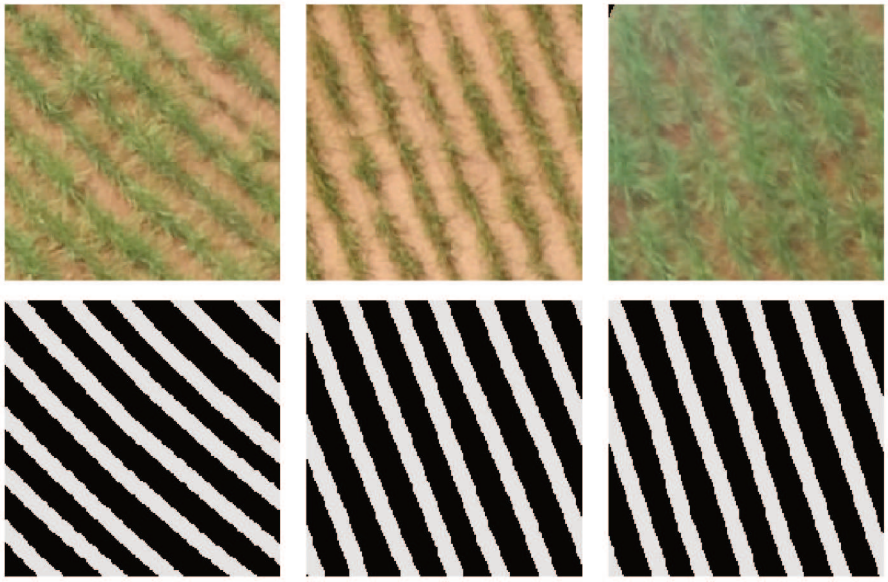


Figure 25 – Examples of crop lines and the segmentation provided by an expert.

Additionally, our datasets include both plant cane (first cut) and crop in the ratoon phase (second cut and on). The main difference between those two phases is the fact that in the last one there are dry leaves and ratoon between the crop rows originated from the last cut. Figure 26a shows clearly this phase, where the presence of the ratoon can be seen in contrast with the reddish soil in the right part of the image. After the cuts, the ratoon is left there as it would be very costy to remove it and also it will decompose gradually and be a natural fertilizer for the own crop benefiting the next cuts. However, depending on the stage of the after-cut, these ratoon can be confused with the crop-rows themselves, confounding the contrast between them and the soil and interfering in the computational process. In contrast, an example of sugar cane crop in the plant phase is shown in Figure 26b.

Figure 26 – Example of sugar cane crops in different phases: (a) Ratoon phase; (b) Plant Phase.

4.1.1 Evaluation metrics

For the evaluation step, we compared the binary images resulting from our segmentation process with the crop line markings provided by an expert. For this comparison we used the Dice Similarity Coefficient (DSC), a common index of segmentation accuracy that measures how similar two binary images are.

Given two images, A and B, Dice coefficient measures the intersection of the objects present in the images, as described in Equation 13. The result from this equation is the similarity D, $0 \leq D \leq 1$, where the more the value D is close to 1, the more similar the images A and B are (González Sánchez et al., 2020).

$$D = 2\frac{|A \cap B|}{|A| + |B|} \tag{13}$$

Also, for the SSN approach, the Jaccard Similarity Coefficient (JSC) (also known as Intersection over Union) was used as a loss function in the training process. This measure is highly recommended for segmentation problems where there are unbalanced classes. The Jaccard Loss between two images A and B is defined as:

$$J(A, B) = 1 - \frac{|A \cap B|}{|A \cup B|} = 1 - \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{14}$$

In addition, DSC is quite similar to JSC. In fact, it is possible to make a conversion from DSC to JSC as described in Equation 15 and vice versa (Equation 16).

$$J = \frac{D}{2 - D} \tag{15}$$

$$D = \frac{2J}{J + 1} \tag{16}$$

4.2 Segmentation using Genetic Algorithm

Crop lines may vary in width depending of the age of the crop and how successful the planting process was. Nevertheless, crop lines frequently show a greenish appearance in contrast to the reddish presence of the soil.

Our first approach was creating and experimenting a method based on GA. Instead of using color segmentation methods and different color space, such as Hue, Saturation, Value (HSV), our first hypothesis was that it is possible to optimize a $3 \times 3 \times 3$ kernel mask so that its convolution combines the local characteristics of the RGB channels of the image into a single gray scale image S, which could latter be segmented using a simple and automatic threshold selection method, such as Otsu. To accomplish that we used a genetic algorithm to optimize the 27 floating point values for the $3 \times 3 \times 3$ kernel mask. The genetic algorithm ran for 2700 generations, using a population size of 200 individuals. We used mutation rate of 0.05 and crossover rate of 0.8.

To train our genetic algorithm we used 35 training images of sugarcane crops with sizes ranging from 485×450 to 1136×1126 pixels. These images contain crop lines of different ages and width extracted from the 4 test maps (Figure 24). We used the computed kernel to filter each training image, thus resulting in an image S, which was then segmented using Otsu. We compared our resulting binary images with the segmented crop lines provided by an expert with DSC. As for the fitness function, the GA aimed to search for the kernel mask that maximized the average Dice coefficient of the images.

In order to evaluate the obtained kernel mask, we used 4 test images of different pixels size. These images are mosaics of aerial images that represents sugarcane crop areas containing crop lines of different ages and width. Figure 24 shows the four test images used in the experiments. It is important to emphasize that each image has their own crop lines segmented by an expert.

Due to external factors that may affect the plantation (such as sowing failures, weed plants and the presence of plants outside the crop line), the segmentation step alone may not be enough to accurately detect the crop lines. In order to improve our line segmentation we used Radon transform described in Section 4.4 to reconstruct the expert's line from our segmentation result. Figure 27 displays the flow chart of this entire first approach.

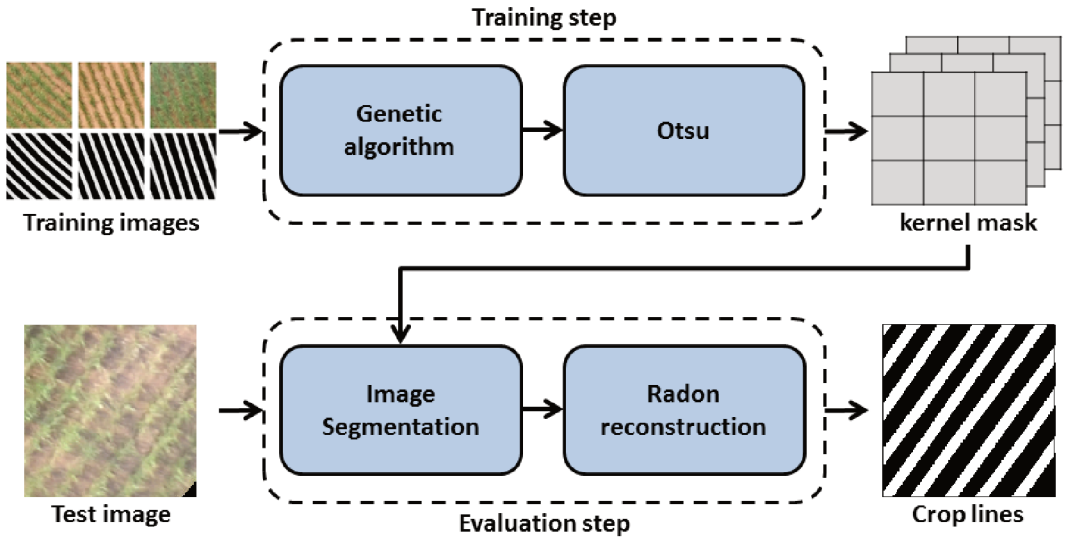


Figure 27 – Flow chart of the first approach based on Genetic Algorithm and Radon transform.

4.3 Semantic Segmentation Networks

Due to the stunning results in the recent literature and the various levels of abstraction that semantic segmentation provides, we decided to study and proposed a new approach based on this method. Semantic segmentation is one of the indispensable tasks in machine

learning and computer vision and it has been used with success in many applications, including autonomous driving. Its relevance has been increasingly growing in the last years due the resumption of convolution neural networks and its fast development, producing the named Semantic Segmentation Networks (SSN). As better stated in Section 2.4.2, these networks aim to assign semantic labels accurately to each pixel of an image. Thus, in the proposed approach, the binarization is performed by using a network trained with a dataset referenced by two classes (crop rows and background). An unclassified image is introduced into the CNN in order to generate an output binarized image, where each of the two color values represents the classification of pixels based on these two classes.

In addition, it is important to state that while the Genetic Algorithm technique was able to work over only one type of feature: reddish and greenish color tones to produce a kernel filter, Semantic Segmentation manages to extract several other different levels of abstraction, each of these levels focusing on a different type of feature, such as border, texture, etc. This is a very important aspect as depending on the stage of the after-cut, dry leaves and ratoon is present in the soil between the crop-rows, confounding the contrast between plants and the soil interfering in the computational analysis process. Thus, as this color contrast may be compromised, a GA based method, in this case, has a disadvantage and that is one of the reasons why we decided to go for the SSN approach, even though the two methods being trained with both cane plant and cane ratoon datasets.

Yet, as previously stated, Otsu global binarization method does not perform well to segment crop lines as they present different local features due to the age and width of the crop line. Also, the Otsu method picking a local threshold to perform binarization is not a feasible option. As the CNN approach does not depend on Otsu as the GA based approach did, this is another reason why we chose to follow the Semantic Segmentation path in our research.

For this approach, we used the same datasets previewed in Figure 24. For each image, we cropped the mosaic into pieces of 256×256 pixels size, with a stride of 256 pixels. If the cropped area does not contain at least 80% of useful information (i.e., pixels with values other than zero), the sample is discarded. After cropping, datasets A, B, C, and D contained a total of, respectively, 678, 3291, 1550 and 2162 images.

In this work, we evaluated U-net, LinkNet, and PSPNet semantic neural networks for the segmentation in aerial crop images. We replaced the encoder of each network by the VGG16 pre-trained with ImageNet weights (SIMONYAN; ZISSERMAN, 2014) due to its high performance in feature extraction in precision agriculture applications (FAWAKHERJI et al., 2019). The basics of these networks was explored in section 2.4.2 and the use of this networks in our new approach is described in the next section (4.3.1).

4.3.1 Proposed Approach

Our approach to perform automatic segmentation of crop lines consists of two steps. First, we use a Convolutional Neural Network to segment the image into crop lines (region of interest) and soil (background). To accomplish that we evaluated three segmentation networks: Unet, LinkNet and PSPNet. For all of them we used VGG16 structure (pre-trained with the ImageNet) to extract the feature maps. Figure 28 shows the architecture of the segmentation networks used. Each network was trained with all the images extracted from the dataset. Each image has a corresponding binarized image provided by an expert extracted from the binarized mosaic using the same pixel coordinates, so that the networks could learn which part of the image corresponded to the first class (crop-row) and which part corresponded to the other class that represents the background (soil, weed or ratton depending on the phase of the crop, region and image).

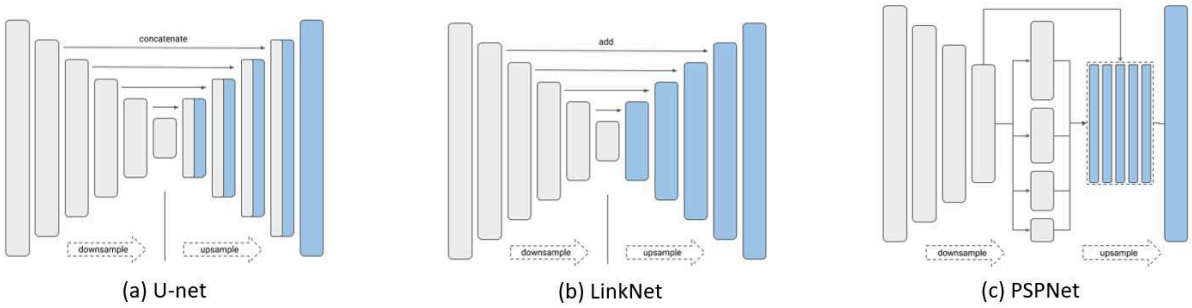


Figure 28 – Architectures used for semantic segmentation. Adapted from (YAKUBOVSKIY, 2019).

In an ideal crop, only the segmentation step should be sufficient to obtain the crop lines accurately. However, in the vast majority of cases, there are external factors that may affect the results. Among these factors are sowing failures (i.e., absence of plants in a section of the line), weed plants (which in the segmentation process will be treated as plants) and plants that are outside the crop row. An example of these problems are shown in Figure 29.

Thus, to improve the segmentation obtained, we used a refinement step, which aims to reconstruct and to improve the previous segmented lines by making them more uniform or linking line fragments and loose plants that belong to the same crop line. For this new approach we also used the Radon transform (Section 4.4) to reconstruct the line obtained during the segmentation step to coincide with the one marked by the expert, as shown in Figure 30.

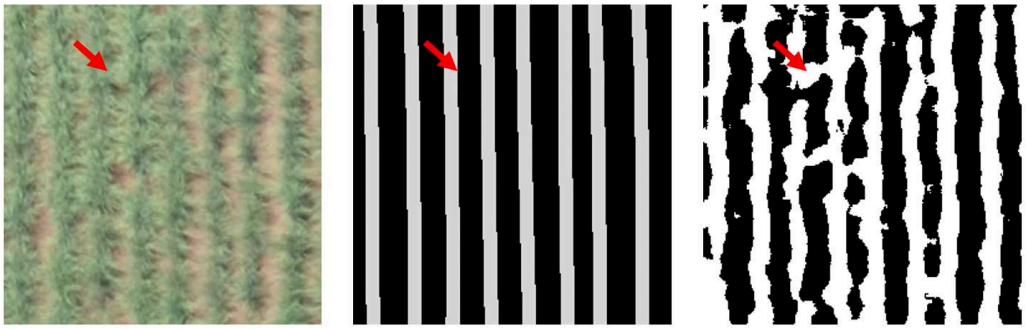


Figure 29 – Example of problems encountered after the segmentation step: (a) Original image; (b) Planting lines provided by an expert; (c) Image after segmentation.

4.4 Line Reconstruction and Refinement

After the binarization process, we needed to reconstruct and improve the already segmented lines. Our aim was to reconnect fragments of crop lines and loose plants that were supposed to belong together and have been separated by failures in the seeding process or in the segmentation step. Also, we aimed to make the crop lines more uniform and smooth, cutting false connection between parallel crop rows.

For this step, we used the Radon transform method. First, we compute the Radon transform from a given input image in order to obtain the orientation angle of the image (Figure 30b). Using this orientation angle it is possible to analyse how the pixels which compose the detected areas are distributed along the image and to detected the center of each crop line in the image (Figure 30c). Then, by using the orientation angle of the image and the center of each crop line we were able to reconstruct the crop lines with the same constant width as used by the expert, as shown in Figure 30d.

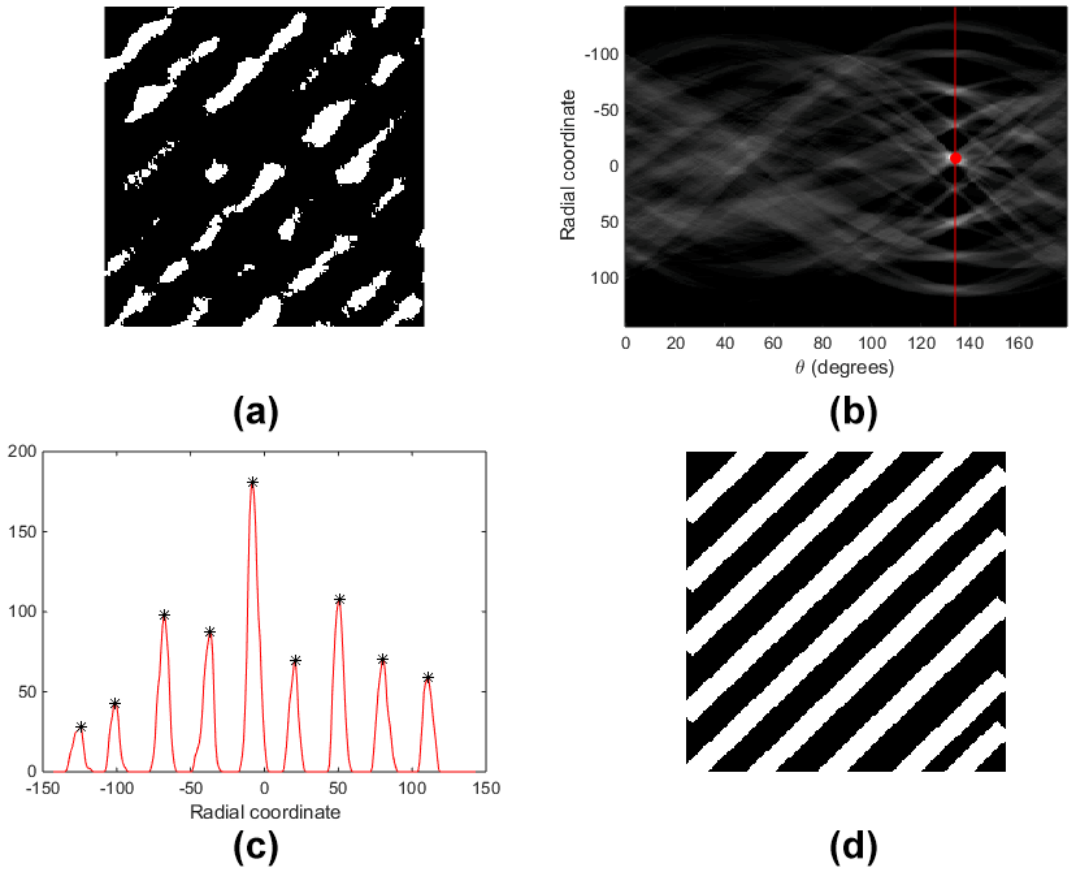


Figure 30 – Proposed scheme for crop line reconstruction using Radon transform: (a) Input image; (b) Matrix obtained with the Radon transform. The red dot represents the location of the maximum point and the orientation angle of the input image; (c) Radon transform obtained for the image orientation angle (red line in (b)). Each peak of the curve corresponds to the center of a line in the input image; (d) Reconstruction of the lines using the orientation angle and the peaks of the Radon transform for that angle.

Chapter 5

Experimental Results

As previously mentioned, in the first stages of this research we experimented with a Genetic Algorithm based approach. Then, due to some reasons including the relevance of Semantic Segmentation in the recent literature, its levels of abstraction, and the non-feasible results of Otsu associated with GA, we decided to study and proposed a new approach based on CNN to accomplish the binarizartion. In this chapter, we better describe experimental results for each of the two approaches including the crop line refinement and reconstruction performed in both approaches by the Radon transform. In the end, we synthesize and discuss the results and differences of our test approaches.

5.1 Segmentation using Generic Algorithm

Since Genetic Algorithm is stochastic, results may be different at independent runs. To address this important issue, we first investigated the convergence of the genetic algorithm through a cross-validation scheme. To accomplish that, we split all 35 training images into 5-folds and we used the images in the training folds as input for the genetic algorithm to compute the kernel mask. As for the fitness function, the GA aimed to search for the kernel mask that maximized the average Dice coefficient of the images in the current training folds.

Figure 31 shows the average Dice coefficient obtained by each image using the 5 kernel mask computed by GA. We must emphasize that, instead of using the images in their respective test fold, we evaluated each kernel in all images in order to have a better estimation of the result. Different kernel mask results in similar segmentation in terms of Dice coefficient. However, it does not mean that the 5 kernel mask present the same set of values, but that they produce similar segmentation results.

Figure 32 presents some examples of segmented areas obtained at different sections of the evaluated crop maps. It is important to notice that the crop lines segmented by experts did not follow exactly the width of each crop line (Figure 32b). Instead, these images contain uniform lines marking the location of each crop line. Moreover, the expert

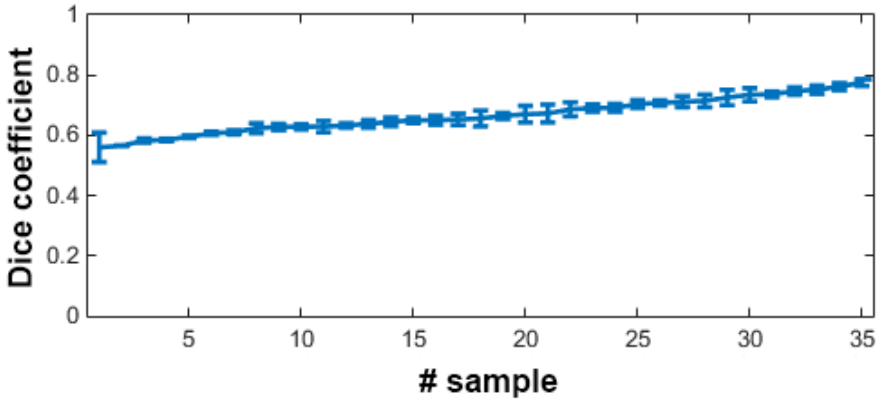


Figure 31 – Average Dice coefficient and standard deviation for different images for 5 different GA kernel masks.

also marked regions where the crop line should exist, even though there is not any plant there. Since these markings do not follow the natural width of the crop lines, an error is expected when comparing the segmentation provided by an expert with the results obtained by our approach.

We noticed that the application of the convolution kernel resulted in an image that is mostly black and white. Although other gray levels are present in the image, their frequency are not significant and could compromise the use of a Global Otsu threshold. This explain the poor segmentation obtained in Figure 32d. The same is true when using a manually defined threshold. Although faster than using the Otsu method, it is not possible to set a user defined global threshold that achieves good results for all evaluated images, as shown in Figure 33. As a result, depending on the threshold value used, some crop lines may be missed while we improve the detection of other crop lines (Figure 32c).

One must consider that global binarization may not be the best approach to segment crop line as they present different local features due to the age and width of the crop line. In order to investigate such matter we applied Otsu algorithm locally on the images. To accomplish that we used a square window of $W \times W$ pixels size and we moved this window along the image using different values of stride, S. We used "OR operator" to combine all local binarizations into a single one binary image. Figure 34 shows that local analysis improves the Dice coefficient obtained by Otsu binarizaton in all configurations evaluated. In general, small windows present better results as they enable Otsu method to capture the local features with more precision, thus providing a better binarization and a higher Dice coefficient, as shown in Figure 32e.

Although the use of local Otsu improves the detection of crop lines (Figure 32e), we noticed that crop lines which are parallels to each other are sometimes connected by regions incorrectly detected as crop lines. Moreover, the detected line presents an irregular width which compromises the comparison with the expert's segmentation.

Since the orientation angle of the image is a local feature, we evaluated this approach

using the same strategy used to compute local Otsu: a square window of $W \times W$ pixels size and a stride S. We used "OR operator" to combine all lines into a single image. Figure 35 shows that the line reconstruction approach enables us to improve the Dice coefficient and the best results are obtained when using a larger window W and a small overlap between windows. Most of the detection errors are a result of small variation of the image that are hard to predict and to detect. Some crop lines may present a small curvature or an abrupt change in direction, thus compromising the line reconstruction (Figure 36). There are cases where the line ends abruptly or lines with different orientations in the image. Nevertheless, Radon line reconstruction represents an improvement in the results obtained by the original segmentation.

Another important feature of this first approach is its ability to deal with curved crop lines and part of that is also used in our latest approach based on semantic segmentation. This approach uses a tiling scheme to compute the Radon transform to reconstruct the crop line. This operation allowed us to successfully approximate curved crop lines by a sequence of straight line segments, as shown in Figure 37. Unfortunately, as one can notice, this process is not perfect, which resulted in an undesirable aliasing effect. Moreover, crop lines with higher curvature tend to be poorly approximated in comparison to more straight lines. Notwithstanding, we must emphasize that both problems can be minimized by correctly choosing the box size and stride used to perform the tiling scheme.

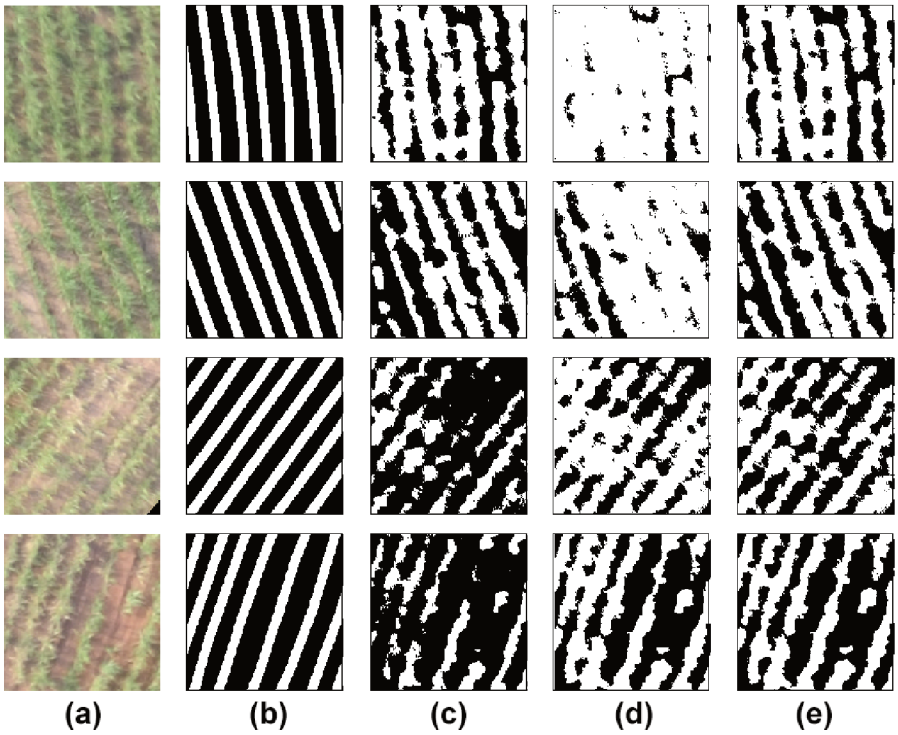


Figure 32 – Results for different sections of the map: (a) Original image; (b) Expert's segmentation; (c) Manual threshold ($t = 0.8$); (d) Global Otsu; (e) Local Otsu ($W = 50$ and $S = 25$).

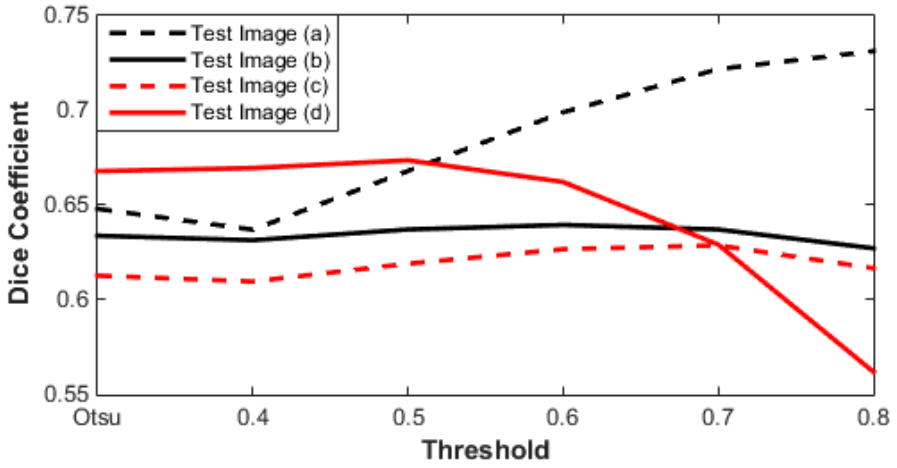


Figure 33 – Dice coefficient for various global threshold values.

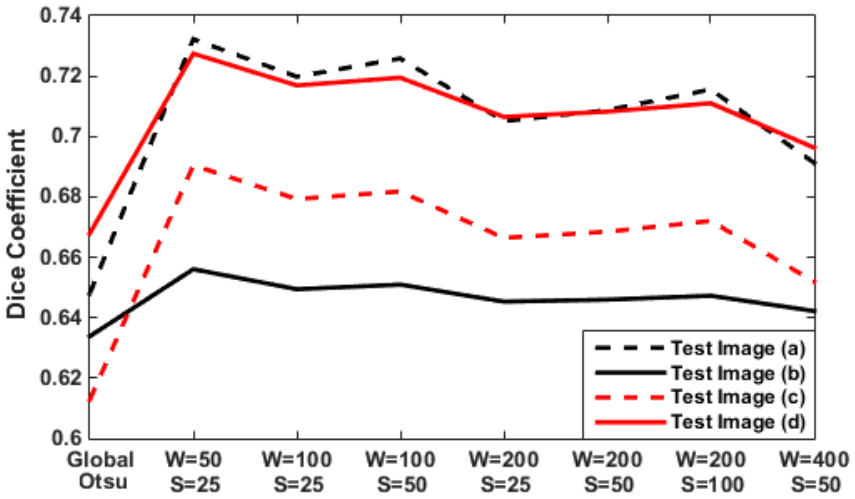


Figure 34 – Dice coefficient obtained using Global Otsu and Local Otsu for different combinations of Window W and Stride S.

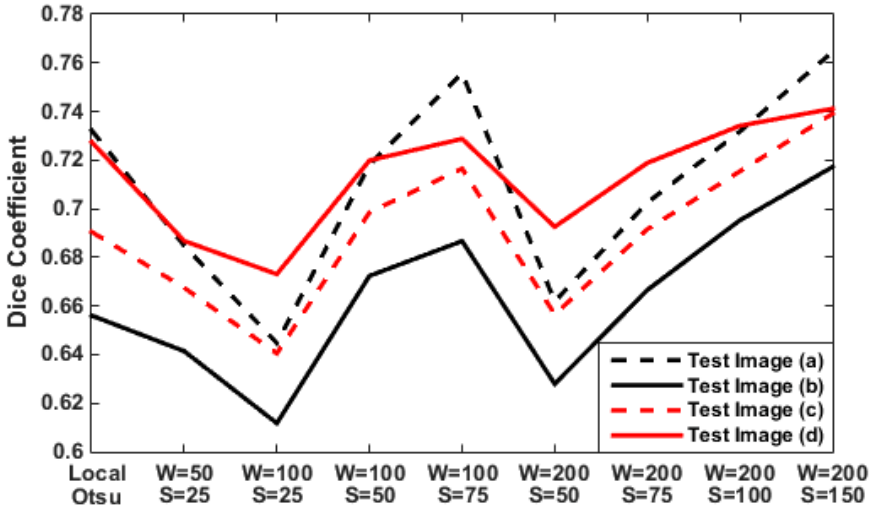


Figure 35 – Dice coefficient obtained for the line reconstruction for different combinations of Window W and Stride S.

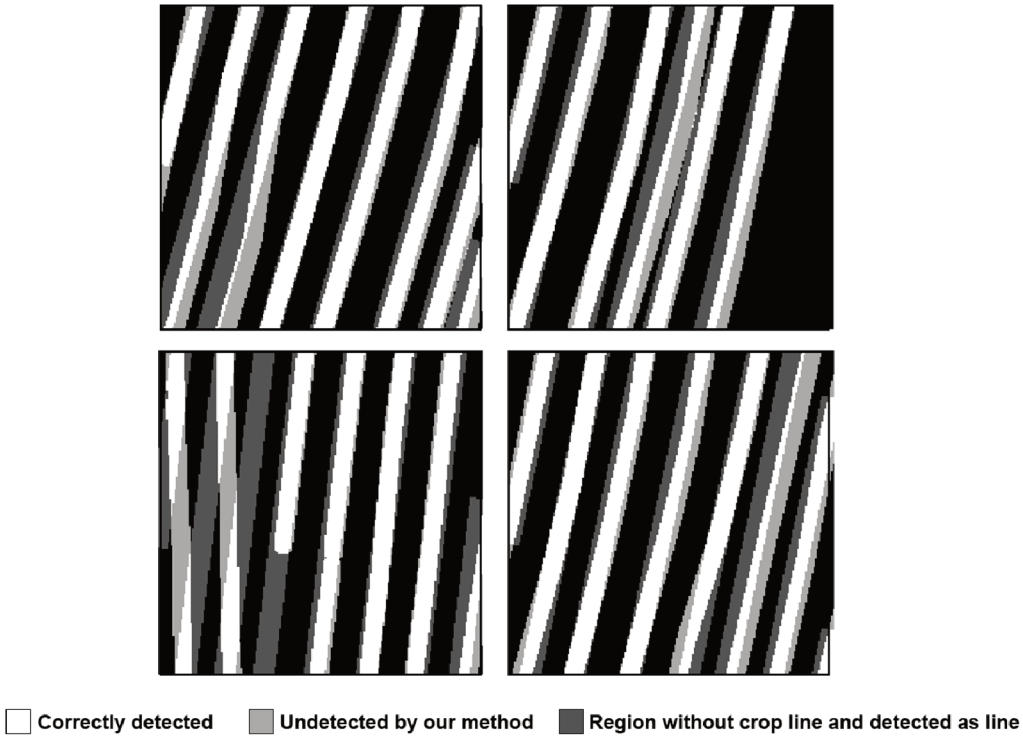


Figure 36 – Errors detected during the line reconstruction.

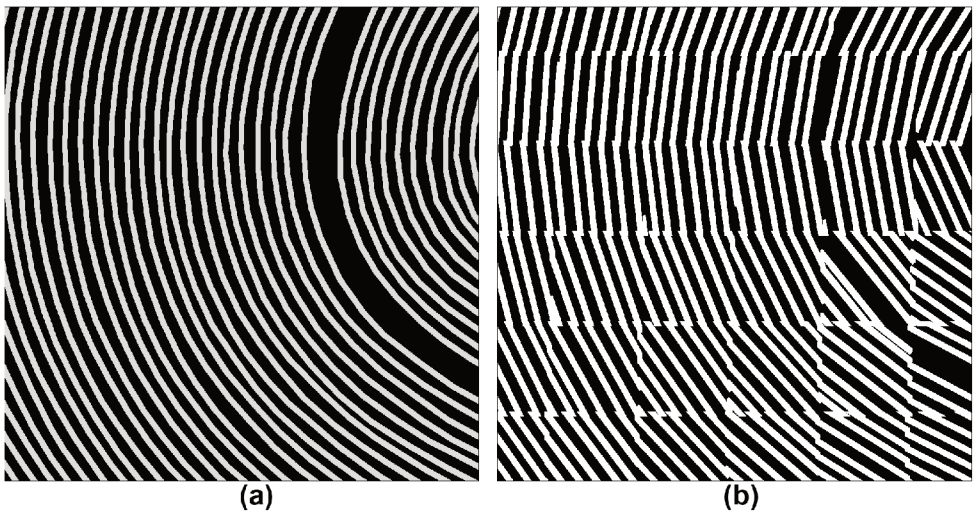


Figure 37 – Curved crop lines: (a) Expert's segmentation; (b) Line reconstruction using Radon transform in a tiling scheme.

5.2 Semantic Segmentation

Initially, we applied the semantic segmentation networks to the images obtained in each dataset. To accomplish that we adjusted the input size of the networks to $256 \times 256 \times 3$ (YAKUBOVSKIY, 2019) and each network has been trained for 50 epochs using the Adam optimizer with a 0.001 learning rate. As loss function, we used the Jaccard-loss (BEERS et al., 2019). We also used simple data augmentation methods, such as rotation (up to $\pm 180°$), translation (up to $\pm 20px$), scaling (up to $\pm 7\%$), and shearing, during the training step to prevent overfitting. To guarantee a correct validation process, a validation step was implemented using K-fold, with $K = 10$. For the training, we used only dataset A to train and validate the network. Figure 38 and Table 1 show the average results obtained by each network.

Segmentation Network	**Dice Coefficient**
VGG16 - LinkNet	0.90 ± 0.0062
VGG16 - PSPNet	0.88 ± 0.0075
VGG16 - Unet	0.87 ± 0.0113

Table 1 – Segmentation results obtained with the application of the segmentation networks in Dataset A.

Results show that LinkNet obtained the best result for the Dice coefficient when segmenting crop lines in the images. Besides, this network was also the one that showed the least variation among test folds. Due to its consistent result in the plant detection process, we chose LinkNet as the standard network for the segmentation stage. We restore the weights that generated the best result during the training of the LinkNet and we used them to evaluate other datasets (B, C, and D). Table 2 shows the results obtained for datasets B, C, and D.

Dataset	**Dice Coefficient**
A	0.90 ± 0.0062
B	0.80 ± 0.0702
C	0.84 ± 0.0724
D	0.86 ± 0.0588

Table 2 – Result obtained with the application of the LinkNet network trained in dataset A to segment other datasets.

We noticed a slight worsening of the LinkNet result when applied to other datasets. The decrease in the Dice coefficient depends on the dataset evaluated. However, it is possible to notice that the average Dice coefficient is superior to 0.80 in all datasets. A possible explanation for this behavior lies in the fact that crop lines can vary in width depending on the age of the crop and how the planting process was successful, as illustrated in Figure 25.

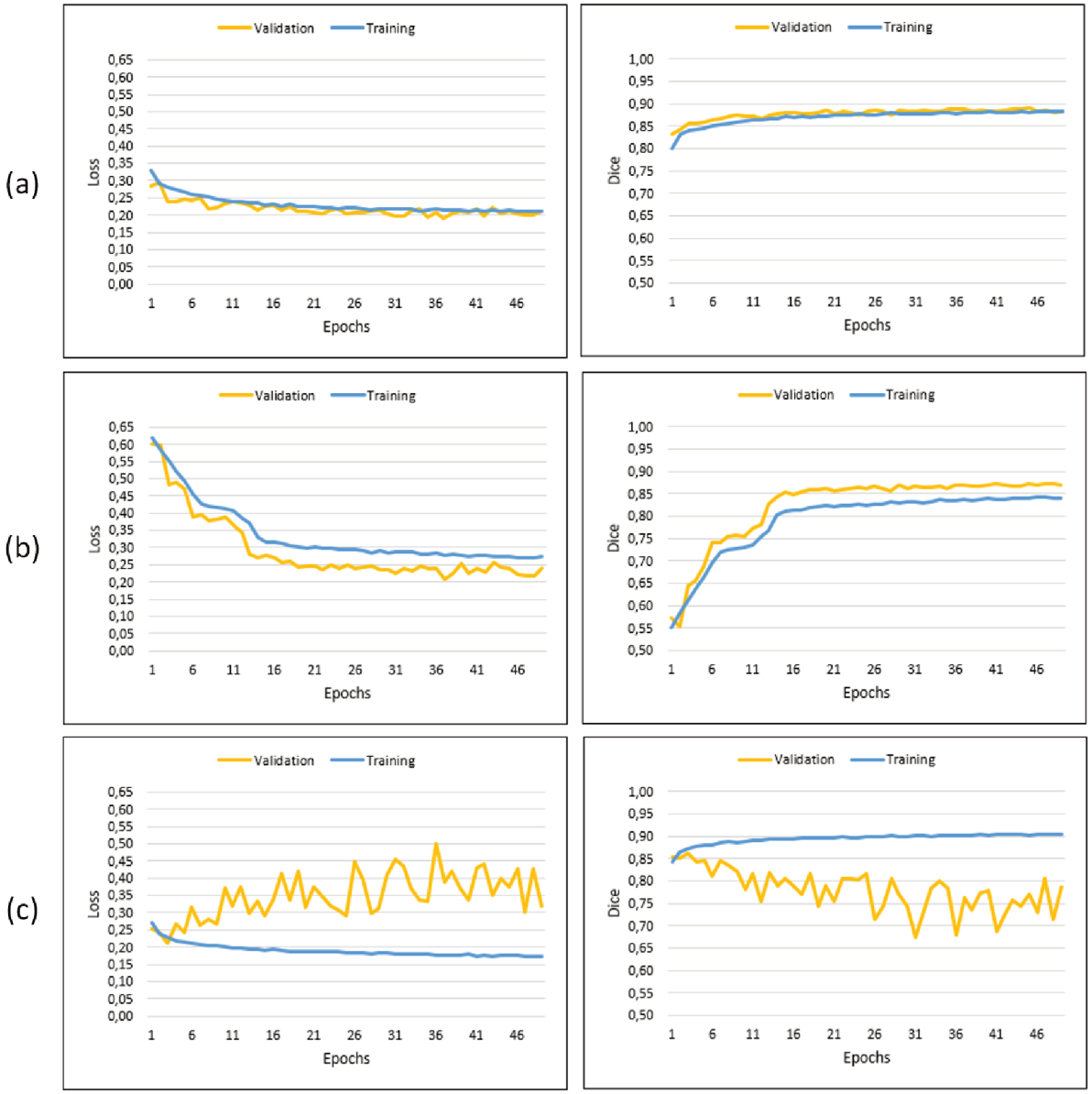


Figure 38 – Results obtained for each segmentation networks. Left column shows the loss function, while the right column shows the Dice Segmentation Coefficient: (a) LinkNet (b) PSPNet and (c) U-net.

After the segmentation, the line reconstruction process was carried out using the Radon transform for all datasets. The results are shown in Figure 39. Initially, we applied the line reconstruction to all images. Unfortunately, as the results show, this approach was not the most suitable. Some images already present a good level of segmentation so that the line reconstruction ends up worsening the result. Regardless of the dataset evaluated, line reconstruction decreased the average Dice coefficient. Due to this fact, we considered to limit the number of images where the reconstruction should be applied. We tested two strategies. First, we used the segmented area to evaluate if an image should be reconstructed or not. Empirically, we found that reconstruction is more suitable for

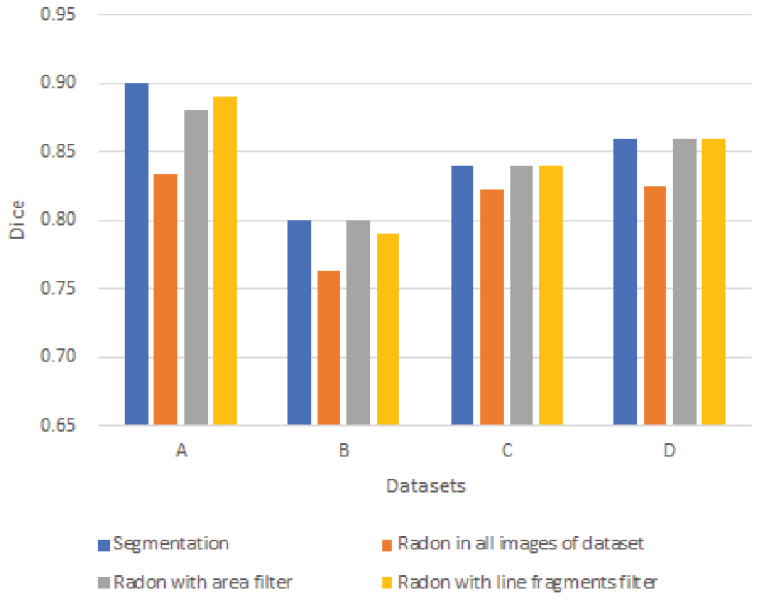


Figure 39 – Average Dice coefficient obtained for different selection approaches during the crop line reconstruction.

images that have up to 30% of segmented area. These images represent regions where the network was unable to correct segment the crop lines within, thus resulting in regions with only fragments of lines, as shown in Figure 40c. Notice that this strategy has an average result superior to the indiscriminate application of the proposed reconstruction step. Unfortunately, this selection criterion is not able to overcome the result obtained by the LinkNet network. Therefore, we evaluated a second strategy where the reconstruction is only performed if there is a minimum number of line fragments present in the image. Empirically, we obtained the best results when considering images that contain at least 20 line fragments. This strategy presents a result similar to the area selection criterion. However, regardless of the selection criteria used, the application of the reconstruction is not able to overcome the result of the LinkNet segmentation (Figure 39). While reconstruction can improve lines that were poorly detected during the segmentation step, the same reconstruction process is not able to adequately deal with various problems present in the image, such as the presence of a slight curvature in the crop rows, crop lines with different angles in the same image, and the fact that the reconstructed line is not fully aligned with the line markings of the expert.

Figures 40 and 41 show some cases of success and failure resulting from the line reconstruction process. When analyzing the segmentation obtained by LinkNet, it is possible to notice the existence of parallel crop lines connected by regions incorrectly detected as plant (Figures 40c and 41c). Also, some detected lines have irregular width, which compromises the comparison with the segmentation provided by the expert.

After the line reconstruction, it is possible to notice that most of the detection errors are a result of small variations in the image, which are difficult to predict and detect.

Figure 41b shows that some crop lines may show a small curvature or an abrupt change of direction, compromising its reconstruction. We also have cases where a planting line ends abruptly or lines with different orientations in the same image. However, the line reconstruction approach using the Radon transform shows an improvement in the results obtained by the original segmentation since it allows the connection in lines where a disconnection was created by a failure in the seeding stage or by an error in the segmentation process, as shown in Figure 40c-d.

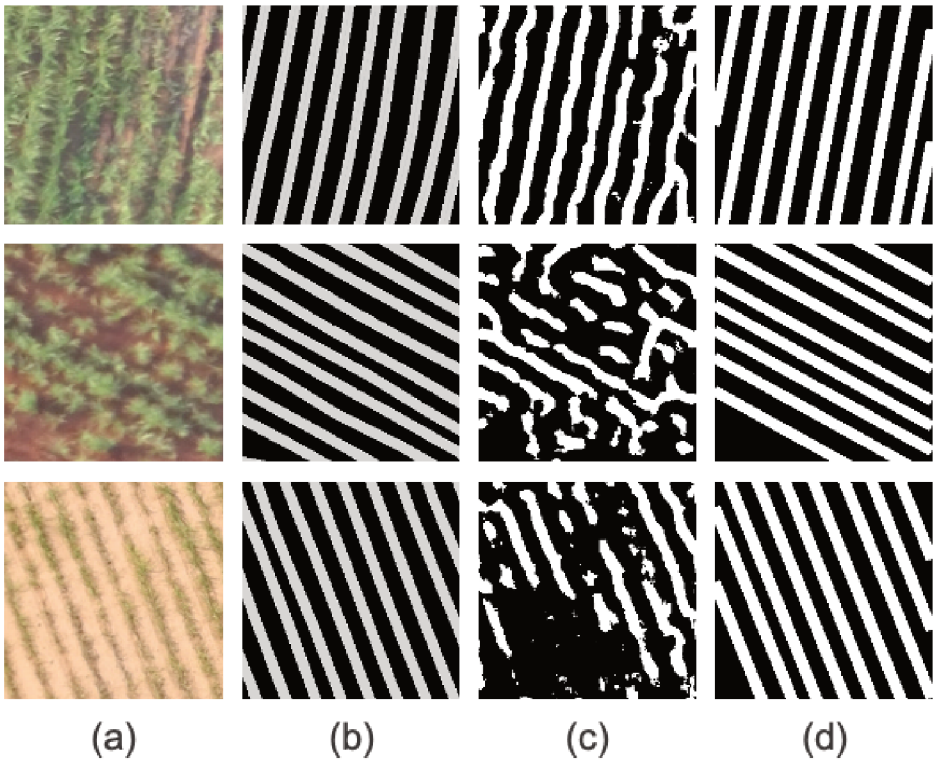


Figure 40 – Examples of images where there was an improvement in the Dice coefficients after line reconstruction using the Radon transform. (a) Original image; (b) Segmentation provided by the expert; (c) Segmentation obtained using LinkNet; (d) Line reconstructed.

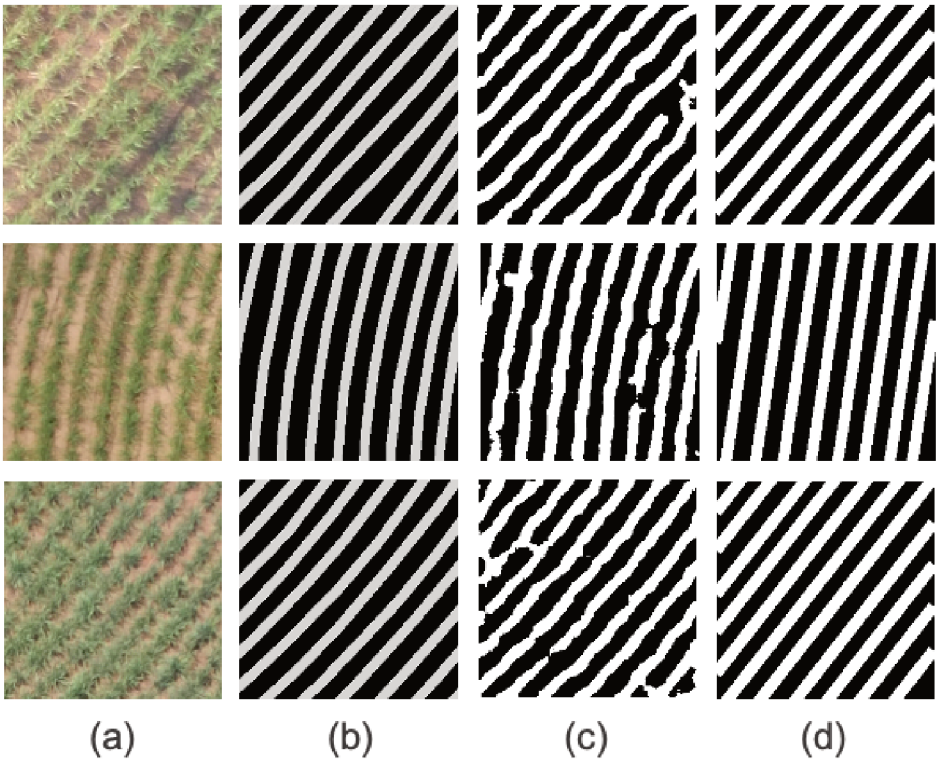


Figure 41 – Examples of images where there was a decrease in the Dice coefficients after line reconstruction using the Radon transform. (a) Original image; (b) Segmentation provided by the expert; (c) Segmentation obtained using LinkNet; (d) Line reconstructed.

5.3 Comparison of approaches

The first approach using Genetic Algorithm was experimented using 35 training images extracted from all four datasets (Figure 24), while in the Semantic Segmentation we used 678 images extracted from the entire dataset A (Figure 24a) in which we applied augmentation methods, such as rotations (up to $\pm 180°$), translations (up to $\pm 20px$), scaling (up to $\pm 7\%$), and shearing. It is important to reinforce that GA is stochastic as its results may be different at independent runs and it requires less training images, while Semantic Segmentation requires more images for the training process. The CNN based technique tend to be more effective with a large amount of images as, in theory, the more it trains and learns, the more it is able to predict, surely when overfitting and unbalanced classes are avoided. Plus, it is important to highlight that the method based on GA used only 27 parameters (3x3x3 floating point values for the kernel mask) to optimize the training, against the millions of parameters SSN used. Although the low number of parameters seems to be an advantage in terms of computational complexity, it may lead to a limitation regarding to how much it is able to learn and predict the correct classes.

In addition, the technique based on the LinkNet segmentation showed a much more

constant Dice coefficient when compared to Genetic Algorithm results. Certainly, this is due to the fact that in our approach GA was able to work over only one type of feature (reddish and greenish color tones to produce a kernel filter). Depending on the crop, dry leaves, weed and ratoon may be present between the crop-rows, compromising the reddish contrast between plants and the background. Thus, as the Semantic Segmentation based technique manages to extract several other different levels of abstraction it tends to be more capable of operating well in different stages of the crop regardless of color contrast.

Yet, the experiments with Genetic Algorithm method did not show good segmentation results according to the Dice evaluation metric when associated with a global Otsu technique. It is undeniable that a local Otsu analysis improves the Dice coefficient obtained. Small windows present better results as they enable Otsu method to better capture the features with more precision. However, the general results did not reach even 0.78, while LinkNet shows average Dice results from 0.80 to 0.86 in tested datasets that were not part of the training, reaching 0.90 when tested with the training dataset.

Chapter 6

Conclusion

In this work, we presented a methodology to segment crop lines from UAV images. First, we experimented an approach based on Genetic Algorithm associated with Otsu method to produce binarized images that were then reconstructed using Radon transform. Then, due to some reasons including the recent relevance of Semantic Segmentation in the literature, its levels of abstraction, and the non-feasible results of Otsu associated with GA, we studied and proposed a new approach based on SSN. This new approach uses a Convolutional Neural Network to perform the segmentation step. Among the networks evaluated, the one that stands is LinkNet presenting the best results to segment crop lines, obtaining a higher and much more consistent Dice coefficient for the datasets evaluated. Which is extremely positive despite the fact that this method requires a larger number of images for the training process. We also proposed a line reconstruction approach based on the Radon transform for this technique testing some variation filters. Although the crop row reconstruction sometimes producing a slight decrease in the Dice coefficient, it enables us to improve the segmentation results by connecting fragments of crop lines and by filling segmentation errors caused by missing plants, thus indicating that our approach is a feasible solution to segment crop lines in images.

6.1 Main Contributions

The main contributions of this work are: the development of a CNN training capable of classifying crop images in two classes (crop-rows and background) generating binarized images, as this efficient segmentation reduces drastically the cost of the post-processing step; and also the development of a framework that receives sugarcane crop images, binarize them using the CNN already trained and then perform a post-processing step based on Radon transform to achieve a refinement process reconstructing and enhancing the lines, making them more uniform and linking row fragments that are supposed to belong together at the same line/row.

In addition, the process of detecting the crop lines in an effective and concise way is

strongly for important for PA, specially planning the harvests, inputs usage, estimating of production, counting plants, early correction of sowing failures. In addition, the geolocation information of the detected crop rows lessens the waste of inputs, the harm to the environment and financial costs. Also, it allows autonomous machinery guidance through the crop,

Finally, sugarcane represents a great percentage of all crops worldwide and also it is a semi-perennial crop, which means that it can be harvested annually for several years without replanting. Thus, the correct and effective detection and frequent maintenance over the years can bring a huge economic impact for producers and, consequently, their countries.

6.2 Contributions in Bibliographic Production

- The following papers are strongly connected to the research presented in this dissertation. They were submitted for their respective journals and are currently in phase of analysis:
  - SILVA, R. R.; ESCARPINATI, M. C. and BACKES, A. R. Sugarcane Crop Line Detection From UAV Images Using Genetic Algorithm and Radon Transform. Submitted to Signal, Image and Video Processing manuscript.
  - SILVA, R. R.; DIAS JR., J. D.; ESCARPINATI, M. C. and BACKES, A. R. Detection of sugarcane crop line from UAV images using Semantic Segmentation and Radon Transform. Submitted to Computers and Electronics in Agriculture.
- The study of CNNs done for this research also allowed us to work in other problems related to computer vision. Hence, our following paper was recently published:
  - SILVA, R. R.; BRITO, L. F. A.; ALBERTINI, M. K.; NASCIMENTO, M. Z. and BACKES, A. R. Using CNNs for Quality Assessment of No-Reference and Full-Reference Compressed-Video Frames. In: XVI WORKSHOP DE VISÃO COMPUTACIONAL, 2020, Uberlândia. Anais do 16° Workshop de Visão Computacional, 2020.
- Furthermore, it is worth mentioning that this work is currently running for the *Prêmio Mercosul de Ciência e Tecnologia* (Mercosur Science and Technology Award):
  - SILVA, R. R.; DIAS JR., J. D.; ESCARPINATI, M. C. and BACKES, A. R. *Detecção de linha de plantio de cana de açúcar a partir de imagens de VANT usando Segmentação Semântica e Transformada de Radon*. Submitted to *Prêmio Mercosul de Ciência e Tecnologia – edição 2020*.

6.3 Future Work

The results obtained by this work demonstrate the good performance obtained by the proposed approach and motivate new lines of investigation, such as: evaluation of datasets of different cultures besides sugar cane; explore how mosaic alignment techniques interfere in the result; explore the use of other sensors in association with the images to produce better results; study new methods to enhance crop reconstruction of regions with highly-curved lines.

Bibliography

ANTHIMOPOULOS, M. et al. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. **IEEE Trans. Medical Imaging**, v. 35, n. 5, p. 1207–1216, 2016. Disponível em: <https://doi.org/10.1109/TMI.2016.2535865>.

BAH, M. D.; HAFIANE, A.; CANALS, R. Crownet: Deep network for crop row detection in uav images. **IEEE Access**, v. 8, p. 5189–5200, 2020. Disponível em: <https://doi.org/10.1109/ACCESS.2019.2960873>.

BALLARD, D. H. Generalizing the hough transform to detect arbitrary shapes. **Pattern recognition**, Elsevier, v. 13, n. 2, p. 111–122, 1981. Disponível em: <https://doi.org/10.1016/0031-3203(81)90009-1>.

BARRETT, H.; SWINDELL, W. **Radiological Imaging: The Theory of Image Formation, Detection, and Processing**. [S.l.]: Elsevier Science, 1996. ISBN 9780120796038.

BASSO, M.; FREITAS, E. P. de. A uav guidance system using crop row detection and line follower algorithms. **J. Intell. Robotic Syst**, v. 97, n. 3, p. 605–621, 2020. Disponível em: <https://doi.org/10.1007/s10846-019-01006-0>.

BEERS, F. van et al. Deep neural networks with intersection over union loss for binary image segmentation. In: **ICPRAM**. [S.l.: s.n.], 2019. p. 438–445.

BELTRAMETTI, M. C.; ROBBIANO, L. An algebraic approach to hough transforms. **Journal of Algebra**, v. 371, p. 669 – 681, 2012. ISSN 0021-8693. Disponível em: <http://www.sciencedirect.com/science/article/pii/S002186931200453X>.

BOSILJ, P.; DUCKETT, T.; CIELNIAK, G. Connected attribute morphology for unified vegetation segmentation and classification in precision agriculture. **Computers in Industry**, v. 98, p. 226 – 240, 2018. ISSN 0166-3615. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0166361517306139>.

BRAS, G. et al. Transfer learning method evaluation for automatic pediatric chest x-ray image segmentation. In: **2020 International Conference on Systems, Signals and Image Processing (IWSSIP)**. [s.n.], 2020. p. 128–133. Disponível em: <https://doi.org/10.1109/IWSSIP48289.2020.9145401>.

BRINKMAN, M. L. et al. Interregional assessment of socio-economic effects of sugarcane ethanol production in brazil. **Renewable and Sustainable Energy Reviews**, v. 88, p. 347 – 362, 2018. ISSN 1364-0321. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1364032118300340>.

CHAMOLA, V. et al. A comprehensive review of the covid-19 pandemic and the role of iot, drones, ai, blockchain, and 5g in managing its impact. **IEEE Access**, v. 8, p. 90225–90265, 2020. Disponível em: <https://doi.org/10.1109/ACCESS.2020.2992341>.

CHAN, Y. K.; KOO, V. C. An introduction to synthetic aperture radar (sar). **Progress In Electromagnetics Research B**, v. 2, p. 27–60, 2008. Disponível em: <https://doi.org/10.2528/PIERB07110101>.

CHAURASIA, A.; CULURCIELLO, E. Linknet: Exploiting encoder representations for efficient semantic segmentation. In: IEEE. **2017 IEEE Visual Communications and Image Processing (VCIP)**. 2017. p. 1–4. Disponível em: <https://doi.org/10.1109/VCIP.2017.8305148>.

CHENG, X. et al. Pest identification via deep residual learning in complex background. **Comput. Electron. Agric**, v. 141, p. 351–356, 2017. Disponível em: <https://doi.org/10.1016/j.compag.2017.08.005>.

CLEVERT, D.-A.; UNTERTHINER, T.; HOCHREITER, S. Fast and accurate deep network learning by exponential linear units (elus). In: BENGIO, Y.; LECUN, Y. (Ed.). **4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings**. [s.n.], 2016. Disponível em: <http://arxiv.org/abs/1511.07289>.

DEWA, C. K.; AFIAHAYATI. Suitable cnn weight initialization and activation function for javanese vowels classification. **Procedia Computer Science**, v. 144, p. 124 – 132, 2018. ISSN 1877-0509. INNS Conference on Big Data and Deep Learning. Disponível em: <http://www.sciencedirect.com/science/article/pii/S187705091832221X>.

DIAS, P. A.; TABB, A.; MEDEIROS, H. Apple flower detection using deep convolutional networks. **Comput. Ind**, v. 99, p. 17–28, 2018. Disponível em: <https://doi.org/10.1016/j.compind.2018.03.010>.

DOW, J. M.; NEILAN, R. E.; RIZOS, C. The international gnss service in a changing landscape of global navigation satellite systems. **Journal of Geodesy**, v. 83, n. 3, p. 191–198, 2009. Disponível em: <https://doi.org/10.1007/s00190-008-0300-3>.

DUDA, R. O.; HART, P. E. Use of the hough transformation to detect lines and curves in pictures. **Commun. ACM**, v. 15, n. 1, p. 11–15, 1972. Disponível em: <https://doi.org/10.1145/361237.361242>.

ECKLE, K.; SCHMIDT-HIEBER, J. A comparison of deep networks with relu activation function and linear spline-type methods. **Neural Networks**, v. 110, p. 232 – 242, 2019. ISSN 0893-6080. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0893608018303277>.

EIBEN, A. E.; SMITH, J. E. **Introduction to evolutionary computation**. 2. ed. [S.l.]: Springer-Verlag, 2003. (Natural computing series). ISBN 978-3-662-05094-1.

EL HAJJOUJI, I. et al. A novel fpga implementation of hough transform for straight lane detection. **Engineering Science and Technology, an International Journal**, v. 23, n. 2, p. 274 – 280, 2020. ISSN 2215-0986. Disponível em: <http://www.sciencedirect.com/science/article/pii/S2215098618314782>.

ELFWING, S.; UCHIBE, E.; DOYA, K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. **Neural Networks**, v. 107, p. 3 – 11, 2018. ISSN 0893-6080. Special issue on deep reinforcement learning. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0893608017302976>.

FAWAKHERJI, M. et al. Crop and weeds classification for precision agriculture using context-independent pixel-wise segmentation. In: IEEE. **2019 Third IEEE International Conference on Robotic Computing (IRC)**. 2019. p. 146–152. Disponível em: <https://doi.org/10.1109/IRC.2019.00029>.

FENG, D. et al. Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges. **IEEE Transactions on Intelligent Transportation Systems**, IEEE, 2020. Disponível em: <https://doi.org/10.1109/TITS.2020.2972974>.

FERENTINOS, K. P. Deep learning models for plant disease detection and diagnosis. **Comput. Electron. Agric**, v. 145, p. 311–318, 2018. Disponível em: <https://doi.org/10.1016/j.compag.2018.01.009>.

FERREIRA, A. dos S. et al. Weed detection in soybean crops using convnets. **Comput. Electron. Agric**, v. 143, p. 314–324, 2017. Disponível em: <https://doi.org/10.1016/j.compag.2017.10.027>.

FUENTES-PEÑAILILLO, F. et al. Using clustering algorithms to segment uav-based rgb images. In: IEEE. **2018 IEEE International Conference on Automation/XXIII Congress of the Chilean Association of Automatic Control (ICA-ACCA)**. 2018. p. 1–5. Disponível em: <https://doi.org/10.1109/ICA-ACCA.2018.8609822>.

GARCÍA-SANTILLÁN, I. et al. Curved and straight crop row detection by accumulation of green pixels from images in maize fields. **Precision Agriculture**, v. 19, n. 1, p. 18 – 41, 2018. Disponível em: <https://doi.org/10.1007/s11119-016-9494-1>.

GHOSH, P. et al. Incorporating priors for medical image segmentation using a genetic algorithm. **Neurocomputing**, v. 195, p. 181 – 194, 2016. ISSN 0925-2312. Learning for Medical Imaging. Disponível em: <https://doi.org/10.1016/j.neucom.2015.09.123>.

GLOROT, X.; BENGIO, Y. Understanding the difficulty of training deep feedforward neural networks. In: TEH, Y. W.; TITTERINGTON, D. M. (Ed.). **Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2010, Chia Laguna Resort, Sardinia, Italy, May 13-15, 2010**. JMLR.org, 2010. (JMLR Proceedings, v. 9), p. 249–256. Disponível em: <http://proceedings.mlr.press/v9/>.

GONZALEZ, R.; WOODS, R. **Processamento Digital De Imagens**. ADDISON WESLEY BRA, 2000. ISBN 9788576054016. Disponível em: <https://books.google.com.br/books?id=r5f0RgAACAAJ>.

González Sánchez, J. C. et al. Segmentation of bones in medical dual-energy computed tomography volumes using the 3d u-net. **Physica Medica**, v. 69, p. 241 – 247, 2020. ISSN 1120-1797. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1120179719305356>.

GUERRERO, J. M.; RUZ, J. J.; PAJARES, G. Crop rows and weeds detection in maize fields applying a computer vision system based on geometry. **Computers and Electronics in Agriculture**, v. 142, p. 461 – 472, 2017. ISSN 0168-1699. Disponível em: <https://doi.org/10.1016/j.compag.2017.09.028>.

GUO, F.; PENG, H.; TANG, J. Genetic algorithm-based parameter selection approach to single image defogging. **Information Processing Letters**, v. 116, n. 10, p. 595 – 602, 2016. ISSN 0020-0190. Disponível em: <https://doi.org/10.1016/j.ipl.2016.04.013>.

HASEGAWA, M.; TABBONE, S. Amplitude-only log radon transform for geometric invariant shape descriptor. **Pattern Recognition**, v. 47, n. 2, p. 643–658, 2014. Disponível em: <https://doi.org/10.1016/j.patcog.2013.07.024>.

HASEGAWA, M.; TABBONE, S. Histogram of radon transform with angle correlation matrix for distortion invariant shape descriptor. **Neurocomputing**, v. 173, p. 24–35, 2016. Disponível em: <https://doi.org/10.1016/j.neucom.2015.04.100>.

HASSANEIN, A. S. et al. A survey on hough transform, theory, techniques and applications. **CoRR**, abs/1502.02160, 2015. Disponível em: <http://arxiv.org/abs/1502.02160>.

HE, K. et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: **ICCV**. IEEE Computer Society, 2015. p. 1026–1034. ISBN 978-1-4673-8391-2. Disponível em: <http://doi.ieeecomputersociety.org/10.1109/ICCV.2015.123>.

HE, K. et al. Deep residual learning for image recognition. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [s.n.], 2016. p. 770–778. Disponível em: <https://doi.org/10.1109/CVPR.2016.90>.

HEMANTH, D. J.; ANITHA, J. Modified genetic algorithm approaches for classification of abnormal magnetic resonance brain tumour images. **Applied Soft Computing**, v. 75, p. 21 – 28, 2019. ISSN 1568-4946. Disponível em: <https://doi.org/10.1016/j.asoc.2018.10.054>.

HOUGH, P. V. **Method and means for recognizing complex patterns**. [S.l.]: Google Patents, 1962. US Patent 3,069,654.

ILIEV, A. I.; KYURKCHIEV, N.; MARKOV, S. On the approximation of the step function by some sigmoid functions. **Math. Comput. Simul**, v. 133, p. 223–234, 2017. Disponível em: <https://doi.org/10.1016/j.matcom.2015.11.005>.

ILLINGWORTH, J.; KITTLER, J. A survey of the hough transform. **Computer vision, graphics, and image processing**, Elsevier, v. 44, n. 1, p. 87–116, 1988. Disponível em: <https://doi.org/10.1016/S0734-189X(88)80033-1>.

JI, R.; QI, L. Crop-row detection algorithm based on random hough transformation. **Mathematical and Computer Modelling**, Elsevier, v. 54, n. 3-4, p. 1016–1020, 2011. Disponível em: <https://doi.org/10.1016/j.mcm.2010.11.030>.

JR, E. R. H.; DAUGHTRY, C. S. What good are unmanned aircraft systems for agricultural remote sensing and precision agriculture? **International journal of remote sensing**, Taylor & Francis, v. 39, n. 15-16, p. 5345–5376, 2018. Disponível em: <https://doi.org/10.1080/01431161.2017.1410300>.

KANG, L. et al. Convolutional neural networks for no-reference image quality assessment. In: **CVPR**. IEEE Computer Society, 2014. p. 1733–1740. ISBN 978-1-4799-5118-5. Disponível em: <http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=6909096>.

KHADANGA, G.; JAIN, K. Tree census using circular hough transform and grvi. **Procedia Computer Science**, v. 171, p. 389 – 394, 2020. ISSN 1877-0509. Third International Conference on Computing and Network Communications (CoCoNet'19). Disponível em: <http://www.sciencedirect.com/science/article/pii/S1877050920310061>.

KREUZER, M.; ROBBIANO, L. **Computational Commutative Algebra 1**. Springer, 2000. Disponível em: <https://doi.org/10.1007/978-3-540-70628-1>.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: **NIPS**. [S.l.: s.n.], 2012.

LAGACHERIE, P.; MCBRATNEY, A. Chapter 1 spatial soil information systems and spatial soil inference systems: Perspectives for digital soil mapping. In: LAGACHERIE, P.; MCBRATNEY, A.; VOLTZ, M. (Ed.). **Digital Soil Mapping**. Elsevier, 2006, (Developments in Soil Science, v. 31). p. 3 – 22. Disponível em: <http://www.sciencedirect.com/science/article/pii/S016624810631001X>.

LEEMANS, V.; DESTAIN, M.-F. Line cluster detection using a variant of the hough transform for culture row localisation. **Image and Vision Computing**, Elsevier, v. 24, n. 5, p. 541–550, 2006. Disponível em: <https://doi.org/10.1016/j.imavis.2006.02.004>.

LI, W. et al. A hybrid modelling approach to understanding adoption of precision agriculture technologies in chinese cropping systems. **Computers and Electronics in Agriculture**, v. 172, p. 105305, 2020. ISSN 0168-1699. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0168169919323191>.

LIMA, M. et al. Sugarcane: Brazilian public policies threaten the amazon and pantanal biomes. **Perspectives in Ecology and Conservation**, 2020. ISSN 2530-0644. Disponível em: <http://www.sciencedirect.com/science/article/pii/S2530064420300262>.

LIN, G.; SHEN, W. Research on convolutional neural network based on improved relu piecewise activation function. **Procedia Computer Science**, v. 131, p. 977 – 984, 2018. ISSN 1877-0509. Recent Advancement in Information and Communication Technology:. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1877050918306197>.

LINDBLOM, J. et al. Promoting sustainable intensification in precision agriculture : review of decision support systems development and strategies. Högskolan i Skövde, Institutionen för informationsteknologi; Högskolan i Skövde, Forskningscentrum för Informationsteknologi; Department of Urban and Rural development, Swedish University

of Agricultural Sciences, Skara, Sweden; National Competence Centre for Advisory Services, Swedish University of Agricultural Sciences, Skara, Sweden; Department of Soil and Environment, Swedish University of Agricultural Sciences, Skara, Sweden, 2017. ISSN 1385-2256. Disponível em: <http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-13271>.

LIU, Z. et al. A heterogeneous processor design for cnn-based ai applications on iot devices. **Procedia Computer Science**, v. 174, p. 2 – 8, 2020. ISSN 1877-0509. 2019 International Conference on Identification, Information and Knowledge in the Internet of Things. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1877050920315611>.

LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. **CoRR**, abs/1411.4038, 2014. Disponível em: <http://arxiv.org/abs/1411.4038>.

LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. In: **2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [s.n.], 2015. p. 3431–3440. Disponível em: <https://doi.org/10.1109/CVPR.2015.7298965>.

MCBRATNEY, A. et al. Future directions of precision agriculture. **Precision agriculture**, Springer, v. 6, n. 1, p. 7–23, 2005. Disponível em: <https://doi.org/10.1007/s11119-005-0681-8>.

MILELLA, A.; REINA, G.; NIELSEN, M. A multi-sensor robotic platform for ground mapping and estimation beyond the visible spectrum. **Precision Agriculture**, Springer, p. 1–22, 2018. Disponível em: <https://doi.org/10.1007/s11119-018-9605-2>.

MIRZAEI, F. M.; ROUMELIOTIS, S. I. A kalman filter-based algorithm for imu-camera calibration: Observability analysis and performance evaluation. **IEEE Transactions on Robotics**, v. 24, n. 5, p. 1143–1156, 2008. Disponível em: <https://doi.org/10.1109/TRO.2008.2004486>.

MITCHELL, M. **An Introduction to Genetic Algorithms**. [S.l.]: The MIT Press, 1996.

MONTALVO, M. et al. Automatic expert system for weeds/crops identification in images from maize fields. **Expert Systems with Applications**, Elsevier, v. 40, n. 1, p. 75–82, 2013. Disponível em: <https://doi.org/10.1016/j.eswa.2012.07.034>.

MONTALVO, M. et al. Automatic detection of crop rows in maize fields with high weeds pressure. **Expert Systems with Applications**, Elsevier, v. 39, n. 15, p. 11889–11897, 2012. Disponível em: <https://doi.org/10.1016/j.eswa.2012.02.117>.

NADIPALLY, M. Chapter 2 - optimization of methods for image-texture segmentation using ant colony optimization. In: HEMANTH, D. J.; GUPTA, D.; Emilia Balas, V. (Ed.). **Intelligent Data Analysis for Biomedical Applications**. Academic Press, 2019, (Intelligent Data-Centric Systems). p. 21 – 47. ISBN 978-0-12-815553-0. Disponível em: <http://www.sciencedirect.com/science/article/pii/B9780128155530000021>.

NAGATA, S. et al. Objective scoring of streetscape walkability related to leisure walking: Statistical modeling approach with semantic segmentation of google street

view images. **Health & Place**, v. 66, p. 102428, 2020. ISSN 1353-8292. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1353829220302720>.

NAIR, V.; HINTON, G. E. Rectified linear units improve restricted boltzmann machines. In: FÜRNKRANZ, J.; JOACHIMS, T. (Ed.). **Proceedings of the 27th International Conference on Machine Learning (ICML-10), June 21-24, 2010, Haifa, Israel**. [S.l.]: Omnipress, 2010. p. 807–814.

NEMOTO, T. et al. Simple low-cost approaches to semantic segmentation in radiation therapy planning for prostate cancer using deep learning with non-contrast planning ct images. **Physica Medica**, v. 78, p. 93 – 100, 2020. ISSN 1120-1797. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1120179720302209>.

OSPINA, R.; NOGUCHI, N. Simultaneous mapping and crop row detection by fusing data from wide angle and telephoto images. **Computers and Electronics in Agriculture**, v. 162, p. 602 – 612, 2019. ISSN 0168-1699. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0168169919301565>.

OTSU, N. A threshold selection method from gray-level histograms. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 9, n. 1, p. 62–66, 1979. Disponível em: <https://doi.org/10.1109/TSMC.1979.4310076>.

OUTAY, F.; MENGASH, H. A.; ADNAN, M. Applications of unmanned aerial vehicle (uav) in road safety, traffic and highway infrastructure management: Recent advances and challenges. **Transportation Research Part A: Policy and Practice**, v. 141, p. 116 – 129, 2020. ISSN 0965-8564. Disponível em: <http://www.sciencedirect.com/science/article/pii/S096585642030728X>.

PANG, Y. et al. Improved crop row detection with deep neural network for early-season maize stand count in uav imagery. **Computers and Electronics in Agriculture**, v. 178, p. 105766, 2020. ISSN 0168-1699. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0168169920311376>.

RADOGLOU-GRAMMATIKIS, P. et al. A compilation of uav applications for precision agriculture. **Computer Networks**, v. 172, p. 107148, 2020. ISSN 1389-1286. Disponível em: <http://www.sciencedirect.com/science/article/pii/S138912862030116X>.

REN, G. et al. Agricultural robotics research applicable to poultry production: A review. **Comput. Electron. Agric**, v. 169, p. 105216, 2020. Disponível em: <https://doi.org/10.1016/j.compag.2020.105216>.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. **International Conference on Medical image computing and computer-assisted intervention**. 2015. p. 234–241. Disponível em: <https://doi.org/10.1007/978-3-319-24574-4_28>.

SANG, H.; ZHOU, Q.; ZHAO, Y. Pcanet: Pyramid convolutional attention network for semantic segmentation. **Image and Vision Computing**, v. 103, p. 103997, 2020. ISSN 0262-8856. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0262885620301293>.

SENIOR, A. W.; LEI, X. Fine context, low-rank, softplus deep neural networks for mobile speech recognition. In: **ICASSP**. IEEE, 2014. p. 7644–7648. Disponível em: <https://doi.org/10.1109/ICASSP.2014.6855087>.

SEZGIN, M.; SANKUR, B. Survey over image thresholding techniques and quantitative performance evaluation. **Journal of Electronic Imaging**, v. 13, p. 13 – 13 – 20, 2004. Disponível em: <https://doi.org/10.1117/1.1631315>.

SILVA, G. R. et al. Definition of management zones through image processing for precision agriculture. In: IEEE. **2017 Workshop of Computer Vision (WVC)**. [S.l.], 2017. p. 150–154.

SIMONYAN, K.; ZISSERMAN, A. **Very Deep Convolutional Networks for Large-Scale Image Recognition**. 2014. Disponível em: <http://arxiv.org/abs/1409.1556>.

SOARES, G. A.; ABDALA, D. D.; ESCARPINATI, M. Plantation rows identification by means of image tiling and hough transform. In: **VISIGRAPP (4: VISAPP)**. [s.n.], 2018. p. 453–459. Disponível em: <https://doi.org/10.5220/0006657704530459>.

SØGAARD, H. T.; OLSEN, H. J. Determination of crop rows by image analysis without segmentation. **Computers and electronics in agriculture**, Elsevier, v. 38, n. 2, p. 141–158, 2003. Disponível em: <https://doi.org/10.1016/S0168-1699(02)00140-0>.

SOUZA, I. R.; ESCARPINATI, M. C.; ABDALA, D. D. A curve completion algorithm for agricultural planning. In: ACM. **Proceedings of the 33rd Annual ACM Symposium on Applied Computing**. 2018. p. 284–291. Disponível em: <https://doi.org/10.1145/3167132.3167158>.

STAN, T.; THOMPSON, Z. T.; VOORHEES, P. W. Optimizing convolutional neural networks to perform semantic segmentation on large materials imaging datasets: X-ray tomography and serial sectioning. **Materials Characterization**, Elsevier, v. 160, p. 110119, 2020. Disponível em: <https://doi.org/10.1016/j.matchar.2020.110119>.

SYLVESTER, G. **E-agriculture in action: Drones for agriculture**. Food and Agriculture Organization of the United Nations and International, 2018. ISBN 978-92-5-130246-0. Disponível em: <http://www.fao.org/documents/card/en/c/I8494EN/>.

UCHIMIYA, M.; SPAUNHORST, D. J. Influence of summer fallow on aromatic secondary products in sugarcane (saccharum spp. hybrids). **Journal of Agriculture and Food Research**, v. 2, p. 100064, 2020. ISSN 2666-1543. Disponível em: <http://www.sciencedirect.com/science/article/pii/S2666154320300454>.

VARUN, R. et al. Face recognition using hough transform based feature extraction. **Procedia Computer Science**, v. 46, p. 1491 – 1500, 2015. ISSN 1877-0509. Proceedings of the International Conference on Information and Communication Technologies, ICICT 2014, 3-5 December 2014 at Bolgatty Palace & Island Resort, Kochi, India. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1877050915001337>.

VERMA, O. P.; PARIHAR, A. S. An optimal fuzzy system for edge detection in color images using bacterial foraging algorithm. **IEEE Trans. Fuzzy Syst**, v. 25, n. 1, p. 114–127, 2017. Disponível em: <https://doi.org/10.1109/TFUZZ.2016.2551289>.

VIDOVIC, I.; CUPEC, R.; HOCENSKI, Z. Crop row detection by global energy minimization. **Pattern Recognit**, v. 55, p. 68–86, 2016. Disponível em: <https://doi.org/10.1016/j.patcog.2016.01.013>.

XU, B. et al. Empirical evaluation of rectified activations in convolutional network. **CoRR**, abs/1505.00853, 2015. Disponível em: <http://arxiv.org/abs/1505.00853>.

YAKUBOVSKIY, P. **Segmentation Models**. [S.l.]: GitHub, 2019. <https://github.com/qubvel/segmentation_models>.

ZADA, B.; ULLAH, R. Pashto isolated digits recognition using deep convolutional neural network. **Heliyon**, v. 6, n. 2, p. e03372, 2020. ISSN 2405-8440. Disponível em: <http://www.sciencedirect.com/science/article/pii/S2405844020302176>.

ZHAO, H. et al. Pyramid scene parsing network. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [s.n.], 2017. p. 2881–2890. Disponível em: <https://doi.org/10.1109/CVPR.2017.660>.