



## **University of Bradford eThesis**

This thesis is hosted in [Bradford Scholars](#) – The University of Bradford Open Access repository. Visit the repository for full metadata or to contact the repository team



© University of Bradford. This work is licenced for reuse under a [Creative Commons Licence](#).

TOWARDS REFINEMENT FOR MEASURING  
SUBCUTANEOUSLY TRANSPLANTED TUMOUR  
MODELS IN MICE

N. HUSSAIN

PHD

2021

Towards Refinement for Measuring Subcutaneously Transplanted Tumour  
Models in Mice

Nosheen HUSSAIN

Submitted for the Degree of  
Doctor of Philosophy

School of Pharmacy and Medical Sciences  
Faculty of Life Sciences  
Institute of Cancer Therapeutics  
University of Bradford

2021

## **Abstract**

Nosheen Hussain

Towards Refinement for Measuring Subcutaneously Transplanted Tumour Models in Mice

Keywords: Subcutaneous tumour models; in vivo models; preclinical pharmacology; tumour detection; image processing; machine learning; deep learning; VGG

Evaluation using mouse subcutaneous tumour models is a key process in cancer drug development. Tumour material is implanted subcutaneously and tumour growth measured using callipers. However this methodology can have poor reproducibility and accuracy due to observer variation. Furthermore the physical pressure of using callipers can distress the mouse and lead to tumour damage. Non-invasive digital tumour imaging would reduce handling stresses and allow volume determination without physical contact. This thesis focusses on capturing 2D digital images of subcutaneous tumours, then using image processing and machine learning methods to determine 3D volume. The biggest challenge faced was lack of differentiation between tumour and surrounding skin, rendering tumour boundary identification difficult. Whilst image processing methods such as colour segmentation and edge detection were unsuccessful, machine learning proved more successful. Three convolutional neural networks, VGG-Face, VGG-19 and VGG-16 models were evaluated, with VGG-Face producing the best results. Using the layer FC7 before RELU activation for extraction in the VGG-

Face model, a tumour recognition rate of 98.86% was achieved. This was increased to 100% through a semi-automatic step with detection repeated on cropped versions of negatively classified images. Finally, volume was determined through extracting image features using the VGG-Face model and conducting partial least squares regression (error of 0.1). This work has successfully demonstrated that with computational methods the volume of subcutaneous tumours can be evaluated through non-invasive digital imaging without need to have contact with the tumour itself, thus offering refinement benefits to the mice as well as eliminating observer bias.

## **Acknowledgements**

In the name of God, the Most Gracious, the Most Merciful

I owe everything to my beloved parents, Khariat Hussain and Rukhsana Ali, for their unwavering faith in me. Their love, selflessness and continuous encouragement moulded me into the person I am today. Thank you to my grandparents for being my biggest supporters. I dedicate this thesis to my beautiful Liyana and Eesa. I hope this thesis teaches you that things may not always go as planned but with strength and perseverance, you can achieve anything you set your mind to.

I offer my sincere thanks and gratitude to my role models, my supervisory team. Firstly, Dr Steve Shnyder played a significant role in my journey and was the backbone of my entire experience; Dr David Connah was instrumental in introducing me to the world of computer science, and instilling in me the faith and self-confidence to move forward. Finally, I thank Prof. Hassan Ugail, who constantly awed me with his invaluable knowledge and encouraged innovation and independent thinking.

An extended thanks to my colleagues from the Centre for Visual Computing for providing a strong support network, with particular mention to Zahra Syed and Ali Bukar.

Finally and most significantly, I would like to praise God, the Almighty, for bestowing upon me the opportunity and capability to successfully complete this thesis.

## Publications

Hussain, N., Connah, D., Ugail, H., Cooper, P. A., Falconer, R. A., Patterson, L. H., & Shnyder, S. D. (2016). The use of thermographic imaging to evaluate therapeutic response in human tumour xenograft models. *Scientific reports*, 6(1), 1-6.

Hussain, N., Cooper, P., Shnyder, S., Ugail, H., Bukar, A. M., & Connah, D. (2017, September). A non-invasive 2D digital imaging method for detection of surface lesions using machine learning. In *2017 International Conference on Cyberworlds (CW)* (pp. 166-169). IEEE.

Bukar, A. M., Ugail, H., & Hussain, N. (2017). On facial age progression based on modified active appearance models with face texture. In *Advances in computational intelligence systems* (pp. 465-479). Springer, Cham.

## Table of Contents

Abstract.....	i
Acknowledgements.....	iii
Publications.....	iv
Table of Contents.....	v
List of Figures.....	viii
List of Tables.....	xiii
List of Abbreviations.....	xiv
<b>1. Introduction .....</b>	<b>1</b>
1.1 The Challenges of Cancer Treatment.....	1
1.1.1 Surgery.....	1
1.1.2 Radiotherapy .....	2
1.1.3 Chemotherapy.....	3
1.1.4 Hormone Therapy.....	4
1.1.5 Immunotherapy.....	5
1.1.6 Biological (Targeted) Therapy.....	6
1.1.7 Heterogeneity and Drug Resistance .....	7
1.2 Preclinical Screening .....	9
1.2.1 Ethics of Animal Use .....	11
1.2.2 Preclinical Animal Models.....	14
1.3 Evaluation of Efficacy in Subcutaneous Tumour Transplantation Models	
19	
1.3.1 Calliper Measurements.....	19
1.3.2 Alternatives to Calliper Measurements .....	23



1.4	Digital Imaging .....	35
1.4.1	Image Processing Techniques that can be applied to find 3D measurement from 2D image .....	36
1.4.2	3D Measurement from 2D image .....	43
1.4.3	Machine Learning .....	49
1.5	Aim and Objectives .....	52
<b>2.</b>	<b>Materials and Methods .....</b>	<b>54</b>
2.1	Materials .....	55
2.1.1	Computer Software Programs .....	55
2.1.2	Equipment .....	55
2.1.3	Tumour Models .....	56
2.2	Methods .....	57
2.2.1	Tumour Xenografts .....	57
2.2.2	Virtual Tumour Models .....	59
2.2.3	Pseudo-tumour Models .....	62
2.2.4	Tumour Detection .....	71
2.2.5	Tumour Volume – 2D Tumour Masks .....	95
2.2.6	Tumour Volume - Shape from Shading .....	102
2.2.7	Tumour Volume – Machine Learning .....	105
<b>3.</b>	<b>Results .....</b>	<b>111</b>
3.1	Tumour Detection – Image Processing .....	111
3.1.1	Colour Segmentation .....	111
3.1.2	Edge Detection .....	119
3.1.3	Contours .....	123

3.1.4	Gradients.....	124
3.1.5	Threshold Image.....	125
3.2	Tumour Detection – Machine Learning.....	128
3.2.1	Cascade Classification .....	128
3.2.2	Convolutional Neural Networks.....	133
3.3	Tumour Volume – Image Processing.....	140
3.3.1	Using Linear Measurements to Determine Tumour Volume.....	140
3.3.2	Mask Volume Created With Length and Width .....	141
3.3.3	Length and Width comparison between mask and calliper .....	145
3.3.4	Mask Volume Represented by Mask Area.....	147
3.3.5	Pseudo-tumour Model Volume .....	149
3.4	Tumour Volume – Shape from Shading.....	151
3.4.1	Depth and Pixel Intensity .....	151
3.4.2	Virtual tumours and different lighting angles .....	152
3.5	Tumour Volume – Machine Learning.....	155
<b>4.</b>	<b>Discussion .....</b>	<b>158</b>
4.1	Conclusion & next steps .....	180
<b>5.</b>	<b>References .....</b>	<b>183</b>

## List of Figures

<b>Figure 1.1:</b> Vernier calliper .....	20
<b>Figure 1.2:</b> Interobserver Bland–Altman analysis for calliper measurements. .	22
<b>Figure 1.3:</b> Linear regression with CT and calliper tumour volume. ....	25
<b>Figure 1.4:</b> Findings from Ayers et al showing deviation in subcutaneous xenograft volume determined via ultrasound and callipers from the true volume. ....	29
<b>Figure 1.5:</b> Thermographic imaging of tumours .....	31
<b>Figure 1.6:</b> Types of step edges.....	39
<b>Figure 1.7:</b> 2D and 3D axes .....	43
<b>Figure 1.8:</b> Diagram illustrating the inverse square law regarding light intensity. ....	46
<b>Figure 1.9:</b> Diagram to show the angle of incidence .....	47
<b>Figure 2.1:</b> Flowchart showing the techniques used in determining tumour detection and tumour volume and the materials used to achieve this.....	54
<b>Figure 2.2:</b> H460 s.c xenograft growth curve .....	59
<b>Figure 2.3:</b> Virtual Tumour Models.....	60
<b>Figure 2.4:</b> Representative virtual model showing lighting angles along the X axis .....	61
<b>Figure 2.5:</b> Pseudo-tumour models used for volume measurements.....	66
<b>Figure 2.6:</b> Pseudo-Tumour models used in the image database.....	68
<b>Figure 2.7:</b> Image of 3D FARO QUANTUM ARM scanner.....	69
<b>Figure 2.8:</b> Scanning Process.....	70
<b>Figure 2.9:</b> An illustration of a pseudo-tumour model and its corresponding 3D scan model, before and after the hole-filling process.....	71

<b>Figure 2.10:</b> Representative primary mouse image used for image processing and detection .....	72
<b>Figure 2.11:</b> Flowchart demonstrating the sequence of image processing methodologies applied to the original image to isolate the mouse from the rest of the image and identify the tumour boundary.....	73
<b>Figure 2.12:</b> Algorithm for RGB colour segmentation.....	75
<b>Figure 2.13:</b> Algorithm for HSV colour segmentation.....	77
<b>Figure 2.14:</b> Pre-processing cascade classification – Prepare images .....	83
<b>Figure 2.15:</b> Sample of positive images .....	84
<b>Figure 2.16:</b> Sample of negative images.....	85
<b>Figure 2.17:</b> Cascade classification - classification .....	86
<b>Figure 2.18:</b> Sample of images used for cascade classification.....	88
<b>Figure 2.19:</b> Pre-processing CNN – create image database.....	90
<b>Figure 2.20:</b> CNN – extract features .....	91
<b>Figure 2.21:</b> CNN - classification.....	93
<b>Figure 2.22:</b> Original and cropped images used for detection by CNN .....	94
<b>Figure 2.23:</b> A flowchart showing the different methodologies used to decipher tumour volume from s.c. xenograft images. ....	95
<b>Figure 2.24:</b> Digital mask of xenograft tumour .....	97
<b>Figure 2.25:</b> Mask volume using the modified ellipsoid formula.....	98
<b>Figure 2.26:</b> Algorithm: Mask volume using only length and width in modified ellipsoid formula.....	99
<b>Figure 2.27:</b> Algorithm- mask area.....	100
<b>Figure 2.28:</b> Algorithm 1 for machine learning to find volume showing how to create the image database .....	107

<b>Figure 2.29:</b> Algorithm 2 to find volume with machine learning is to extract features to find volume.....	108
<b>Figure 3.1:</b> Original image split into red, green and blue channels, displayed at different pixel threshold levels to isolate the mouse from the remainder of the image.....	112
<b>Figure 3.2:</b> Applying the blue binary image mouse mask to the original image to isolate the mouse.....	114
<b>Figure 3.3:</b> Blue mask thresholds between 85 and 110 applied to the original image to remove the blue glove from the image .....	114
<b>Figure 3.4:</b> HSV colour channels and corresponding pixel intensity histograms .....	115
<b>Figure 3.5:</b> The hue mouse mask was applied to the original image resulting in the mouse being isolated from the image. ....	116
<b>Figure 3.6:</b> Sample of images with hue mask applied showing the inconsistency in isolating the mouse outline for each image .....	117
<b>Figure 3.7:</b> Effects of applying erosion on the hue mask and isolated mouse image .....	118
<b>Figure 3.8:</b> Right hand side image shows result of edge detection on eroded mouse mask image on left hand side.....	119
<b>Figure 3.9:</b> Edge detection applied on the original mouse image to identify the tumour boundary.....	120
<b>Figure 3.10:</b> Edge detection with a threshold applied on original mouse image to identify the tumour boundary.....	121
<b>Figure 3.11:</b> Edge detection on the mouse flank showing the tumour and its surroundings was unsuccessful in detecting the tumour boundary .....	122
<b>Figure 3.12:</b> Noise reduction on image of tumour .....	123

<b>Figure 3.13:</b> Xenograft images with their contour maps.....	124
<b>Figure 3.14:</b> Gradient magnitude and gradient direction applied to the sample image.....	125
<b>Figure 3.15:</b> Original image with corresponding threshold image .....	126
<b>Figure 3.16:</b> Threshold image with image processing techniques.....	127
<b>Figure 3.17:</b> Cascade classification detection results to locate the tumour in the sample image.....	130
<b>Figure 3.18:</b> Version 2 of cascade classification testing to identify the tumour on the cropped image through enclosing the tumour in a yellow box.....	131
<b>Figure 3.19:</b> Version 3 of cascade classification testing to identify the tumour on the cropped image through enclosing the tumour in a yellow box.....	132
<b>Figure 3.20:</b> Version 4 of cascade classification testing to identify the tumour on the cropped image through enclosing the tumour in a yellow box.....	133
<b>Figure 3.22:</b> Detection results using classifier trained with VGG-Face with layer FC7-before-relu extracted.....	136
<b>Figure 3.23:</b> Crop image and repeat detection for 100% classification .....	137
<b>Figure 3.24:</b> Relationship between calliper and mask volumes found using the modified ellipsoid formula. ....	141
<b>Figure 3.25:</b> Mask volume created with length Vs original mask volume and calliper volume .....	143
<b>Figure 3.26:</b> Mask volume created with width Vs original mask volume and calliper volume .....	144
<b>Figure 3.27:</b> Comparing length and width for mask and calliper measurements .....	146
<b>Figure 3.28:</b> Mask area compared to mask and calliper volume .....	148

<b>Figure 3.29:</b> Pseudo tumour volumes from 2D masks, 3D scanning and callipers .....	150
<b>Figure 3.29:</b> Depth vs pixel intensity .....	152
<b>Figure 3.31:</b> Pixel intensity vs. volume .....	153
<b>Figure 3.33:</b> Error found through multiple regression for model shape vs. volume .....	155
<b>Figure 3.34:</b> Datasets used to train classifiers and their corresponding volumes .....	156
<b>Figure 3.35:</b> Real Vs estimated tumour volume .....	157

## List of Tables

<b>Table 1.1:</b> Formulas used to calculate tumour volume.....	21
<b>Table 1.2:</b> Non-invasive imaging methods .....	34
<b>Table 1.3:</b> Showing the commonly used colour spaces RGB, YUV, CIE and HSV with their colour definitions describing how they define and segment colour ....	37
<b>Table 1.4:</b> Edge detectors .....	41
<b>Table 1.5:</b> Depth cues in single and multiple images .....	44
<b>Table 2.1:</b> Summary of the computer software programs used in this project. .	55
<b>Table 2.2:</b> Description of equipment used in the project.....	56
<b>Table 2.3:</b> Description of the different tumour types used in the project.....	57
<b>Table 2.4:</b> Preliminary testing for artificial model scanning.....	64
<b>Table 3.1:</b> Experiments ran using cascade classification for tumour detection .....	129
<b>Table 3.2:</b> Convolutional neural networks tested in this project for tumour detection .....	134
<b>Table 3.3:</b> Datasets used to train convolutional neural network .....	134
<b>Table 3.4:</b> Using CNNs to detect tumours .....	135
<b>Table 3.5:</b> Tumour detection summary table .....	139



## List of Abbreviations

2D	Two-dimensional
3D	Three-dimensional
ACO	Ant Colony Optimisation
AI	Aromatase inhibitors
ATP	Adenosine triphosphate
BLI	Bioluminescence imaging
CIE	Comission Internationale de l'Eclairage (International Commission on Illumination)
CLL	Chronic lymphocytic leukaemia
CML	Chronic myelogenous leukaemia
CMY	Cyan, magenta, yellow
CNN	Convolutional neural network
CONV	Convolutional
DNA	Deoxyribonucleic acid
EGFR	Epidermal growth factor receptor
FC	Fully connected
FDG	[18F]-fluorodeoxyglucose
FLI	Fluorescence imaging
FP	Fluorescence proteins
GEM	Genetically engineered models
HER-2	Human epidermal growth factor receptor 2
HoG	Histogram of oriented gradient
HSV	Hue, saturation, value
ILSVRC	ImageNet Large Scale Visual Recognition Challenge

KS	Kolmogorov Smirnov
L	Maximum length
LBP	Local binary patterns
MAE	Mean absolute error
MRI	Magnetic resonance imagine
MicroCT	Micro computed tomography
PCA	Principal component analysis
PDX	Patient derived xenografts
PET	Positron emission tomography
PLS	Partial least squares
PSF	Point spread function
RELU	Rectified linear unit
RGB	Red, green, blue
SARM	Selective androgen receptor modulator
SERM	Selective oestrogen receptor modulator
S.c.	Subcutaneous
SVM	Support vector machine
UV	Ultraviolet
VGG	Oxford Visual Geometry Group
W	Perpendicular width
YCbCr	Luminance, chroma blue, chroma red
YCgCr	Luminance, chrominance components
YDbDr	Luminance, chrominance components
YIQ	Luma, in-phase, quadrature
YPbPr	Green (Y), Blue (Pb), Red (Pr)
YUV	Luminance, bandwidth, chrominance

## 1. Introduction

### 1.1 The Challenges of Cancer Treatment

In the UK, 1 in 2 people born after 1960 will be diagnosed with cancer during their lifetime[1][2]. Cancer is an aggressive neoplastic disease characterised by uncontrollable cell division, resulting from mutations in the cell genome. Mutations occur spontaneously or through stimulation from environmental factors, such as UV radiation[3]. They trigger increased expression of oncogenes, cancer causing genes, and deactivation of tumour suppressor genes. Multiple cell pathways are manipulated resulting in heterogeneity in the disease and therefore, cancer is extremely challenging to control. With a diagnosis of cancer every 2 minutes and a cancer-related death every 4 minutes in the UK[1][2], there is a pressing need for new therapies[4]. Treatment is challenging due to the heterogeneous nature of cancer coupled with its ability to metastasise. With over 200 types of cancer, patients are commonly given a cocktail of therapies to ensure maximum effectiveness.

#### 1.1.1 Surgery

When a cancerous mass is localised and caught before metastasis, the best option for treatment is surgical removal, known as resection. Other therapies can kill a proportion of the cancerous cells but surgery guarantees the death of all removed cells[5]. For example, 90-92% of patients with colon cancer[6] and 84% of patients with rectal cancer[7] are treated with surgery. However, cancers eventually metastasize leading to a poor prognosis. When colorectal cancer

metastasizes to the liver, surgery is used to remove both the primary and secondary tumour resulting in 30-40% improvement for 5-year survival rates[8]. Surgery can be used in combination with other therapies, such as radiotherapy, to shrink the size of the tumour thus making it easier to remove and to ensure the majority of cancer cells are destroyed[5]. As with other treatments, new improvements to surgery are constantly being introduced, such as fluorescence-guided surgery which is a relatively new approach currently in clinical trials. This allows visualisation of the tumour through targeting the tumour with fluorescently labelled molecules[9]. In some cases, cancerous masses can grow in difficult to reach places thus rendering surgery obsolete as a treatment option and therefore, alternative treatment options must be explored.

### **1.1.2 Radiotherapy**

Radiotherapy gained popularity in the 1920's and is used for treatment in at least 45% of new cancer cases[10]. Radiation, such as X-rays, are used to damage the DNA of cancer cells thus killing the cell. Early head and neck tumours and early Hodgkin's disease look to radiotherapy as a curative treatment[5]. Radiotherapy is often given in conjunction with surgery, being utilised both before and after surgery. Preoperative radiotherapy is used to control the disease ensuring that it does not spread further before surgery; this has proven to be successful for rectal cancer. Treating patients with only local surgical resection resulted in a 5.3% occurrence at two years. This decreased to 2.4% when combined with preoperative radiotherapy[11]. Furthermore, radiotherapy is used postoperatively to decrease the chances of recurrence and ensure that all cancer cells have been killed. For example, there was a 9-10% improvement in the

survival of breast cancer patients after 10 years when postoperative radiation was applied[12].

Another methodology is intraoperative radiotherapy, where radiation is applied to exposed cancer tissue during surgery. This method provides a boost for postoperative treatment permitting reduced exposure to radiation therapy resulting in fewer side effects. 21Gy given in intraoperative radiation is equivalent to 60Gy given postoperatively[13]. Recurrent adenocarcinoma of the prostate, after treatment with radiation therapy, was found in 75% of patients, measured by the elevated levels of a prostate-specific antigen found through biopsy[14]. Cryotherapy, a treatment injecting cold gasses directly into the prostate to kill cancer cells, is often used as a salvage therapy for failed radiotherapy treatment of the prostate. A study by Bahn et al conducted a 7 year follow up after cryotherapy treatment and reported improved survival rates[14]. Radiotherapy can induce an immune response as well as altering the tumour microenvironment, and thus combination therapy of radiotherapy with immunotherapy is an area of great interest [15].

### **1.1.3 Chemotherapy**

Anticancer agents were initially developed in the 1940s. Nitrogen mustards were alkylating agents causing cross linking in DNA molecules which resulted in cell death[16]. Antifolate drugs blocked the functions of folate requiring enzymes thus causing regression in tumours[16]. Since then, a variety of cytotoxic DNA damaging drugs designed to kill cancer cells by obstructing cell replication have been developed. For example, Doxorubicin is an anthracycline chemotherapy

drug and works by intercalating DNA strands, directly obstructing DNA replication[17]. Paclitaxel is also a chemotherapy drug but works by interfering with microtubules during the cell division process[18]. In a study for metastatic breast cancer, 36% of patients treated with doxorubicin had a response and had median survival of 18.9 months and 34% of patients treated with paclitaxel had a response and had a median survival of 22.2 months. Interestingly, when patients were treated with both doxorubicin and paclitaxel, a response was seen in 46% of patients[19].

The disadvantages of chemotherapy are the toxic side effects. Even though the agents have highly specific interactions with their target cells they are not selective for tumour cells. Therefore, normal cells with high proliferation rates are also targeted resulting in a narrow therapeutic window[20]. Tumours are excellent in adapting to their environment and are quickly able to develop resistance against the treatment given via the activation and overregulation of counteracting pathways, this is aided by their heterogeneity[3]. An increase in the understanding of the mechanisms of cancer enables a more targeted approach to drug therapies. An example of this is hormone therapy.

#### **1.1.4 Hormone Therapy**

Certain cancers are driven via hormones and these hormones can be targeted. Breast and prostate cancer are largely driven by hormones oestrogen and testosterone, respectively. Oestrogen plays an important role in breast cancer by aiding proliferation and stimulating the breast cancer growth. 50-80% of breast tumours have an oestrogen receptor[21]. Tamoxifen is a selective oestrogen

receptor modulator (SERM) which is an antagonist to the oestrogen receptor thus stopping its effects. An 11% improvement in the 10 year survival rate was found in patients taking tamoxifen[22]. However, tamoxifen led to an increased risk of other cancers such as uterine cancer and so was succeeded by the more successful aromatase inhibitors (AIs). Aromatase is an enzyme which is involved in oestrogen production thus AIs cause oestrogen depletion without interacting with oestrogen receptors[23].

Testosterone is an androgen hormone which contributes to stimulating prostate cancer growth. Bicalutamide is a selective androgen receptor modulator (SARM) used to treat prostate cancer through blocking testosterone by binding to the androgen receptor. Disease progression was significantly reduced with a reduction of 42%[24]. Not all cancers are driven by hormones; an alternative approach is to focus on the body's own immune system.

### **1.1.5 Immunotherapy**

The immune system is configured to ensure that all foreign substances, known as antigens, are eliminated. Cancer cells are either too powerful for the immune system or undetectable as foreign as they originated from normal cells. Immunotherapy involves using the body's natural defence system to fight the cancer. Antibodies are proteins created by plasma cells which can attach to antigens and mark them to be destroyed by other components of the immune system[25]. The drug Alemtuzumab is a monoclonal antibody used to treat chronic lymphocytic leukaemia (CLL) in which the white blood cells, mature B lymphocytes, are abnormal. The B lymphocytes have a CD52 protein on the

surface and Alemtuzumab attach to the CD52 protein thus marking them for destruction, ensuring that the abnormal cells are destroyed[26].

One of the fastest growing areas of cancer treatment research at the moment is in the development of immune-checkpoint inhibitors (ICI), as highlighted by the award of the 2018 Nobel Prize for Medicine and Physiology to James Allison and Tasuku Honjo. Although monoclonal antibodies have been approved for immune checkpoint proteins PD-1/PD-L1 and CTLA4, such as Nivolumab, Avelumab and Ipilimumab, there are issues with off-target toxicities with these agents, and further work is necessary to develop safer and more effective ICI's, looking at targeting other immune checkpoint proteins [27].

### **1.1.6 Biological (Targeted) Therapy**

An increase in knowledge of the biology of cancer and a need to avoid the toxic side effects has led to an up rise in specific target driven agents. These are aimed at targeting a specific physiological process or function in tumours which are essential for the survival of the cancer but not essential in normal cells[28].

Chronic myelogenous leukaemia (CML) is a cancer in which there is an increased unregulated production of mature granulocytes in the bone marrow. CML has a characteristic abnormal chromosome in over 90% of its patients. Fusion of chromosome 9 and 22 known as the Philadelphia chromosome results in the synthesis of the oncoprotein P210bcr/abl, a tyrosine kinase[29]. This deregulates tyrosine kinase activity resulting in inhibiting apoptosis and increasing the potential for the mutant bcr/abl cells to proliferate[30]. Imatinib is an oral targeted



therapy made to inhibit the bcr/abl tyrosine kinase and a study was conducted comparing this to interferon-alpha, a human cytokine that was the first line therapy for CML[31]. A study found that 87.1% of patients had a major cytogenetic response with imatinib in comparison to 34.7% response from interferon-alpha[32]. Imatinib replaced interferon-alpha as the first line therapy for CML.

Although targeted therapies are selective for cancer cells, these agents can have poor efficacy. To improve the targeting ability as well as ability to kill cancer cells, antibody-drug conjugates, agents which consist of a monoclonal antibody attached to a cytotoxic drug, have risen in popularity, with several having been approved for clinical use in recent years [33]. For example, Mirvetuximab soravtansine and Tisotumab vedotin-tftv were developed to treat Platinum-resistant ovarian cancer and recurrent or metastatic cervical cancer, respectively[34].

Ultimately, cancer continues to dominate due to heterogeneity in the disease and acquired drug resistance[35].

### **1.1.7 Heterogeneity and Drug Resistance**

The genetic complexity of cancer cells becomes harder to treat due to heterogeneity in the disease; a tumour of the same type can have different genetic and epigenetic make-up in different patients, thereby rendering some treatment plans ineffective[36]. When Herceptin, a monoclonal antibody, which attaches itself to the human epidermal growth factor receptor 2 (HER-2), was

given as treatment to patients overexpressing the HER-2 receptor in breast cancer, only 26% had an objective response rate[37]. The low response rate can be further attributed to intra-tumour heterogeneity, which suggests that genetic differences can also be found in sub-populations of cancer cells within the same tumour[36]. Heterogeneity also plays a role in acquired drug resistance.

Heterogeneity results in inherent drug resistance, where cancer cells are predisposed to resistance due to their genetic makeup. Furthermore, acquired drug resistance can occur, where the cancer initially responds to treatment but then develops resistance[38]. Although caused by multiple factors, common reasons include changes in the drug target, DNA damage repair, an inability of the cells to uptake drugs with the same vigour and the utilisation of alternative pathways that reduce cell death and promote survival[39], [40].

This was seen with non-small-cell lung cancer, where mutations were found in the epidermal growth factor receptor (EGFR) in approximately 40% of the cancer cases; the mutation resulted in activation of EGFR signalling and therefore had oncogenic properties. These cancers were successfully treated with EGFR tyrosine kinase inhibitors, drugs targeting the tyrosine kinase domain of the EGFR[41], but approximately 60% of patients developed acquired drug resistance due to a further mutation in the EGFR. This mutation resulted in an increase of ATP which then competed with the treatment drugs in binding to the mutated EGFR[42].

Although there is a large variety of treatment options available, the clinical success rate for pharmaceutical agents reaching registration is extremely poor.

From approximately 10,000 compounds sourced at the drug discovery phase, only 1 – 5 will be submitted for approval to be tested in clinical trials and the process can take up to 10 years[43]. Cheaper, more efficient and accurate preclinical screening methods in the drug development process could improve the output of drugs reaching clinical testing in humans.

## 1.2 Preclinical Screening

Therapeutics can be sourced through screening natural products, creating derivatives of compounds already established and through purposeful design to target a specific biological target. The therapeutics are subjected to a screening cascade, including *in silico*, *in vitro* and *in vivo* stages, designed to test if the drug is active in producing an effect.

Increasing knowledge of bioactivity and drug interactions has facilitated extensive data collection. In addition to biological targets, targets with known clinical side-effects, termed 'anti-targets', have also been found[44]. This information has been collated in multiple online databases, such as ChEMBL, which contains the data of over 2 million compounds and over 14 million bioactivity data points[45]. These databases facilitate *in silico*, computer based, testing of compounds against protein targets and anti-targets to find the best combination to produce a potential therapeutic effect with minimal side effects[46]. Mestres et al collated interactions between 802 drugs and 480 biological targets from 7 different databases finding 4767 unique interactions thus concluding an average of 6 targets for every drug[47]. However, these are only predictions. For a drug to be

approved for testing in clinical trials, more robust testing is required, moving away from simulations and predictions to physical evidence.

When analysing drug-target interactions, testing is done *in vitro*, where assays and analytical procedures, are performed in cells or tissue samples maintained outside a living organism[48]. Initial screens in cell-free biological assays and cell-based assays are used to determine target specificity, activity and mechanism of action[49]. Furthermore, information can be collected on acute and long term cytotoxicity, mutagenicity, adhesion, angiogenesis, migration and invasion[49]. *In vitro* testing is advantageous as assays are simple, cheap and allow for high throughput testing of potential therapeutics against multiple biomolecules or cell lines[50]. Unfortunately, *in vitro* conditions are not reflective of those in human patients and so have limited predictive value. Increased use of more sophisticated 3D and mixed cell models have resulted in improved analysis but this cannot account for systemic complexities found in whole living organisms.

A therapeutic agent's pharmacokinetic profile must be analysed, looking at how the body interacts with the agent once it is inside the body. This includes the agent's absorption into systemic circulation, how it is distributed from the bloodstream to tissues, the metabolism of the agent and its elimination from the body. Therefore, once lead molecules have been identified, they must be tested *in vivo*, inside whole living organisms, animals[50].

### 1.2.1 Ethics of Animal Use

If testing is to be carried out on animals, it is important that an ethical framework is in place which can ensure that there is no mistreatment or unjustified use of animals. One of the cornerstones of such framework is the 3Rs. These come from a book written by Russell and Burch 'The principles of humane experimental technique', published in 1959, which advocated safe and ethical experiments on animals[51]. They proposed the 3Rs, replacement, reduction and refinement, which endeavour to eradicate animal use by ensuring every effort is made to find suitable alternatives. Furthermore, Russell and Burch went beyond focussing on experimental methodology, encouraging ethical thinking and empathy, by referring to the concepts of humanity and inhumanity, with inhumanity being described through distress, anxiety and frustration of a need[51]. It encouraged a holistic view regarding animals as opposed to viewing them as an empty vessel used to test drugs[52].

Since then, the 3Rs have been incorporated in legislation worldwide. In the UK, the Animals (Scientific Procedures) Act 1986 specifies that the 3Rs must be justified in every experiment using animals. The first of the 3Rs is replacement. It must be proved a suitable alternative to using protected animals, living vertebrates, is not available. Full replacement involves not using any animals at all[53]. Animals can be replaced via established animal cell lines, human volunteers or mathematical computer models[54]. Partial replacement involves using animals incapable of experiencing suffering, including immature forms of vertebrates and invertebrates such as amoebae and nematode worms[53].

Next of the 3Rs is reduction, whereby the number of animals used in a study must be reduced to as few as possible. It is imperative that good planning precedes animal experiments to ensure maximum information is obtained from a limited number of animals. Statistical design has a key role in reduction through predicting the number of animals required to form an accurate conclusion from the data[55].

Finally, the last of the 3Rs is refinement and relates to animal welfare and suffering, ensuring that living conditions are satisfactory and all attempts have been made to minimise any pain or stress the animal is subjected to, as well as incorporating humane endpoints. Ultimately, it is ensuring the animal is as happy as possible[56]. Hurst and West found that reducing stress for animals through improved husbandry techniques positively impacted experimental results as the data was less-variable[57]. For instance, mice exhibited lower anxiety levels when handled with a cupped hand where they were free to roam around the glove or were led through a clear acrylic tunnel in comparison to being lifted by their tail. Allowing them to choose to enter a tunnel or cupped hand gave them autonomy thus making human contact less intimidating and threatening[57]. *In vivo* models are an essential step in evaluating a drug before it goes into clinical testing in humans.

Whilst *in vitro* tests can give very specific information about drug-target interactions, they aren't able to cover the complexities of tumour growth *in vivo*. Tumours *in vivo* have an altered microenvironment and extracellular pH, with poor blood flow resulting in hypoxia and a change in interstitial fluid pressure. These conditions can alter properties of the agent culminating in poor conditions

for drug penetration and delivery[58]. Tumour cells tested *in vitro* are exposed to the full concentration of agent applied whereas *in vivo* the drug circulates around the body and is subject to metabolising enzymes before reaching the tumour. *In vivo* models provide important safety information regarding dosage and toxicity ahead of clinical trials with humans[58]. Ultimately, *in vivo* animal models have higher predictive values than *in vitro* models and therefore, guidelines worldwide stipulate that all new medicines must be tested on animals to ensure patient safety[59].

Furthermore, not testing drugs adequately *in vivo* can lead to catastrophic events, as was seen with the drug thalidomide, originally an antiepileptic drug which was then prescribed to alleviate morning sickness in pregnant women[60]. Unfortunately, due to a lack of thorough, appropriate preclinical testing, it was unknown that the drug was able to travel through the placenta with damaging results; over 5000 infants were born with malformations[61]. Whilst it is recognised that animal research is crucial, it is closely governed and the 3Rs, play an important role in achieving this.

In terms of preclinical cancer screening, mice are the most commonly used species. With similar nervous and reproductive systems, the mouse genome has many comparable genes with humans. Consequently, mice suffer from diseases similar to humans, including but not limited to cancer, diabetes and anxiety disorders[62]. The anatomy and physiology of mice is well understood and thus any changes are easily observed. Their rapid metabolism, short lifespan and high reproductive rates allow analysis through different stages of the life cycle in multiple generations[63].

## 1.2.2 Preclinical Animal Models

*In vivo* preclinical screening cancer models play an essential role in the drug discovery process and are involved in target validation, lead compound identification and lead compound optimisation[58]. The main screening models are: subcutaneous transplantation models, where cells are injected under the skin; orthotopic models, where cells are transplanted at the tumour site of origin, and genetically engineered models (GEM), in which genes are manipulated resulting in spontaneous tumours[58], [63]. Subcutaneous models are the most commonly used.

### 1.2.2.1 Subcutaneous Models

Subcutaneous (s.c.) models involve implanting established tumour cell lines or tissue underneath the mouse skin[64]. The three main types of subcutaneous models are allografts, xenografts and patient derived xenografts (PDX). Allograft mice models have a full immune system. The implanted tumour is of the same genetic background as the host mouse, therefore, tumour rejection is less likely and the tumour micro-environment closely resembles the original tumour microenvironment[65]. However, a study by Voskoglou found murine allografts had very poor translation from preclinical to clinical activity in humans in comparison to xenograft models[48].

Xenograft models consist of established human, or other species, tumour cell lines implanted subcutaneously in immunocompromised mice. This model enables testing of a variety of tumour cell lines and provides fast and reproducible



results. Furthermore, the mice are nude, hairless, thus the tumour sits under the skin and is visible and easily accessible. Issues with these models are that the majority of cell lines are not implanted at their site of origin, xenografts suffer from a loss of heterogeneity, due to continual passage of the cell lines, and low metastatic rates[66].

PDX models improve on xenografts as human tumour tissues are excised and implanted in the mouse; the genetic, histological and phenotypic features of the tumour are preserved and stabilised[67]. Although this model has high physiological relevance, it is difficult to access freshly excised human tumours. The limited engraftment rates and long latency time for tumour development is unfortunate considering the process carries high costs and is labour intensive[58].

Human tumour grafting highlights the potential for personalised medicine as the patient's own tumour is propagated in the mice and used for testing. In one study, a patient with advanced pancreatic cancer was initially treated with conventional gemcitabine treatment but was found to be resistant and the disease progressed significantly[68]. Human tumour graft analysis of the patient's tumour found a biallelic (mutation in both alleles of the gene) inactivation of the PALB2 gene, a gene which functions in tumour suppression. The tumour responded to treatment with mitomycin C, even though this is not a common second line therapy for patients resistant to gemcitabine[68]. When the patient was treated mitomycin C, a DNA damaging agent, the patient remained symptom free for 3 years which is a big improvements from the median 3 month survival expected for patients with gemcitabine resistance[68]. In another study, PDX were developed for

gynecologic cancer and provided information on the patient's prognosis. The more successful the PDX engraftment rate, the poorer the prognosis of the ovarian cancer[69]. Interestingly, genes that were up-regulated in the PDX models correlated with the patient outcome, highlighting the usefulness of this model in personalised medicine[69].

### **1.2.2.2 Orthotopic Models**

Where cell lines in s.c. models are implanted under the skin, cell lines in orthotopic tumour models are implanted at the site of the tumour origin thus reconstituting the original tumour microenvironment. Orthotopic tumour models also use established cell lines in immune-competent or immune-deficient rodents. Models are site-specific where cells are injected so that they form tumours in specific locations and so can also be used to mimic metastatic tumours. As the model facilitates local and metastatic spread, the interactions between the stroma and tumour can be observed [70].

Testing Bevacizumab (Avastin) in an orthotopic model gave positive results. Avastin is a very popular drug used to treat breast carcinoma, colorectal cancer, renal cell carcinoma, and non-small cell lung cancer[71]. The drug is a humanized monoclonal antibody targeting the VEGF protein, therefore inhibiting the growth of new blood vessels. When tested in an orthotopic model the drug significantly reduced the disease progression. Most patients were seen to have a prolonged survival rate of up to 3 years[71]. Using orthotopic models to predict disease progression was highlighted in another study looking at metastatic prostate

cancer. Resecting the primary tumour resulted in significantly longer survival rates due to slower disease progression and reduced metastasis[72].

### 1.2.2.3 Genetically Engineered Mouse Models

Genetically engineered mouse (GEM) models involve inducing specific genetic changes in a mouse to increase the chances of developing cancer or resistance. This can include oncogene expression, gene knockout or knock in, promoter genes or mutations in tumour suppressor genes. The mice are usually immune-competent and the tumour growth is spontaneous and autochthonous[73]. Aspirin and Enalapril work by inhibiting the angiotensin-converting-enzyme inhibitor and are commonly used to treat high blood pressure and kidney disease[74]. Aspirin is used to treat pain and inflammation mainly used for cardiovascular patients. These drugs are shown to have chemo-preventive properties in genetically engineered mouse model that had pancreatic cancer. Treatment after 3-5 months showed a delay in progression of pancreatic cancer[74].

In another study, GEM models were used to identify proteins involved in metabolic and inflammation processes as potential biomarkers in saliva in relation to tumour development in breast cancer with the aim to detect the cancer early thus leading to an improved diagnosis as well as prognosis. [75]. A major advantage of using GEM models is the tumours develop spontaneously and are not implanted, and so are histologically and genetically accurate[76]. Therefore, the tumours are more readily translated to human tumours in comparison to subcutaneous xenograft models. Conversely, GEM models are subject to

heterogeneity and variability and so can be difficult to co-ordinate for longitudinal studies[76].

#### 1.2.2.4 Summary

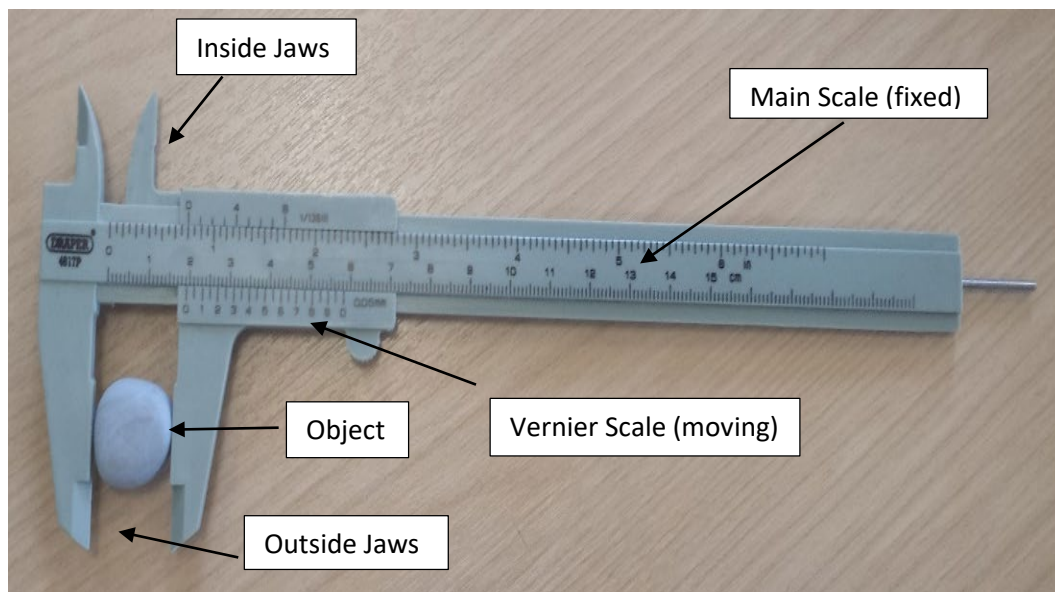
The s.c. tumour transplantation model is the gold standard for *in vivo* screening. It is the most commonly used to demonstrate drug anti-tumour efficacy, dosing regimens and toxicology as the model is simple and provides fast, cheap, reproducible results with easy access to the tumour[66]. S.c. models affirm that sufficient amounts of a therapy can get through the body and to a solid tumour without getting metabolised or causing off-target effects in order to shrink the tumour or delay its growth compared to untreated controls. Unlike orthotopic and GEM models, the tumour is not at its site of origin and therefore the native tumour microenvironment is not present and can result in s.c. models having poor clinical predictive power.

Although orthotopic and GEM models boast higher metastasis rates, they are hindered by long tumour development times and high expenses. Access to tumours is difficult as the primary and secondary tumours can be located anywhere in the mouse body whereas the tumour sits on the surface of the mouse in the s.c. model. Ultimately, s.c tumour transplantation models are still most commonly used due to their numerous practical advantages over orthotopic and GEM models. The next section discusses in detail the methodology for s.c tumour transplantation models [58], [63].

## **1.3 Evaluation of Efficacy in Subcutaneous Tumour Transplantation Models**

### **1.3.1 Calliper Measurements**

For an agent to be considered successful in these models, it must be able to significantly delay tumour growth in treated tumours in comparison to control tumours. Tumour growth is monitored through measuring tumour volume. This is normally done by taking 2D measurements of the tumour and translating them into a volume using a geometrical formula. The measurements are usually taken with Vernier callipers. 2D measurements are taken with a calliper in the X and Y axis to find the length and width. It is difficult to measure the Z axis, height, as tumours are rounded or abstract in shape[77]. Vernier callipers, Figure 2, have a main scale and a secondary moving vernier scale offering increased precision benefits[78]. The longest length of the tumour and its perpendicular width are measured; both measurements are decided upon by the investigator[78]. When taking measurements, the tumour is often squeezed until resistance is felt. Finally, the measurements are inputted into a formula to calculate volume.



**Figure 1.1:** Vernier calliper

Image highlighting the inside and outside jaws which hold the object being measured and the main and Vernier scale from which the measurements are read.

Before the 1980's, multiple formulas were used to calculate volume; Tomayko and Reynolds grouped and tested the various formulas and these are summarised in Table 1.1[78]. The absolute volume, used for comparison, was found through excising the tumours and weighing them on a digital balance to find their mass. Formulas based on the tumour area and diameter performed poorly whereas spheroid, rectangular and ellipsoid formulas had improved predictive potential[78]. From these, Tomayko and Reynolds found the ellipsoid formulas had the lowest mean square errors and boasted the best correlation with tumour size. Furthermore, every formula in Table 1.1, including the ellipsoid formula, was poor in predicting the volume for larger tumours and there was no significant difference between single and multi-lobed tumours[78].

**Table 1.1:** Formulas used to calculate tumour volume.

Formulas used to calculate tumour volume are shown in the table below with their correlation to true volume and mean square error[78]. L= length, W = width and H = height. True volume was calculated by excising the tumours and weighing them on a digital balance to find mass.

Category	Formula	Correlation with tumour size	Mean square error
<b>Area</b>	L x W	0.88	6.7
	$\pi/4 \times L \times W$	0.88	11.4
<b>Diameter</b>	L	0.63	53.0
	$(L + W)/2$	0.65	59.4
	$(L \times W \times H)^{1/3}$	0.66	68.0
<b>Rectangular solid volume</b>	L x W x H	0.93	121.9
	L x W <sup>2</sup>	0.82	484.2
<b>Spheroid volume</b>	$4/3 \pi \times r^3$		
	r = L/2	0.77	291.7
	r = (L + W)/2	0.85	79.8
<b>Ellipsoid volume</b>	$\pi/6 \times L \times W \times H$	0.93	5.2
	$\pi/6 \times [(L + W)/2]^3$	0.85	79.8
	$\pi/6 \times L \times W^2$	0.82	52.0
	$\pi/6 \times (L \times W)^{3/2}$	0.85	71.8
	$0.4 \times L \times W^2$	0.82	18.4
<b>Ellipsoid volume assuming <math>\pi=3</math></b>	$1/2 \times L \times W^2$	0.82	43.0
	$1/2 \times L \times W \times H$	0.93	5.1

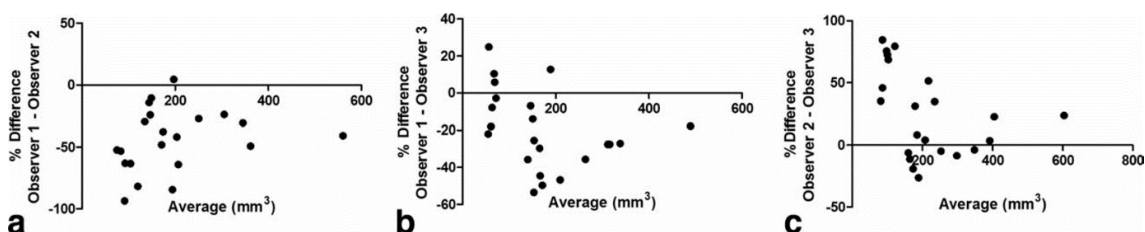
From the ellipsoid formulas that required only two measurements of length and width, omitting the height measurement, the formula with the highest correlation with tumour size and lowest mean square error became the industry standard:

$$\text{Modified ellipsoid formula: } \frac{1}{2} (L \times W^2)$$

Calliper measurements are highly subjective as the location of the longest length and perpendicular width are decided upon by the investigator. Additionally, the tumour is often squeezed until resistance is felt before the measurement is taken.

In one study, inter and intra-observer variation in calliper measurements was tested using experienced investigators[79]. The average difference between any two observers was found to be 15%. The intra-observer variation was tested through repeated measurements of the same tumour, whilst investigators believed the tumours were different to each other. The difference was 27% and 7% for small and large volumes, respectively[79].

Evidently, callipers carry poor reproducibility and accuracy as a result of human error and observer bias. Kersemans et al constructed a Bland-Altman analysis, also known as a difference plot showing the agreement between two quantitative measurements, which in this case highlights the differences in calliper measurements between three observers shown in Figure 1.2[80]. The findings illustrate a high and inconsistent percentage difference between observers in taking calliper measurements[80].



**Figure 1.2:** Interobserver Bland–Altman analysis for calliper measurements.

Graphs (a), (b) and (c) show a large percentage difference between the three different observers highlighting the interobserver variation[80].

In addition to this, using callipers can cause refinement issues as the animal is subjected to handling and its tumours are squeezed; the process of squeezing could result in damage to the tumour and its surroundings thus defeating the



purpose of the experiment[79], [81]. Although the calliper is the gold standard instrument, it is important to explore alternative methodologies to find accurate tumour volume whilst ensuring the mice gain maximum refinement benefits.

### 1.3.2 Alternatives to Calliper Measurements

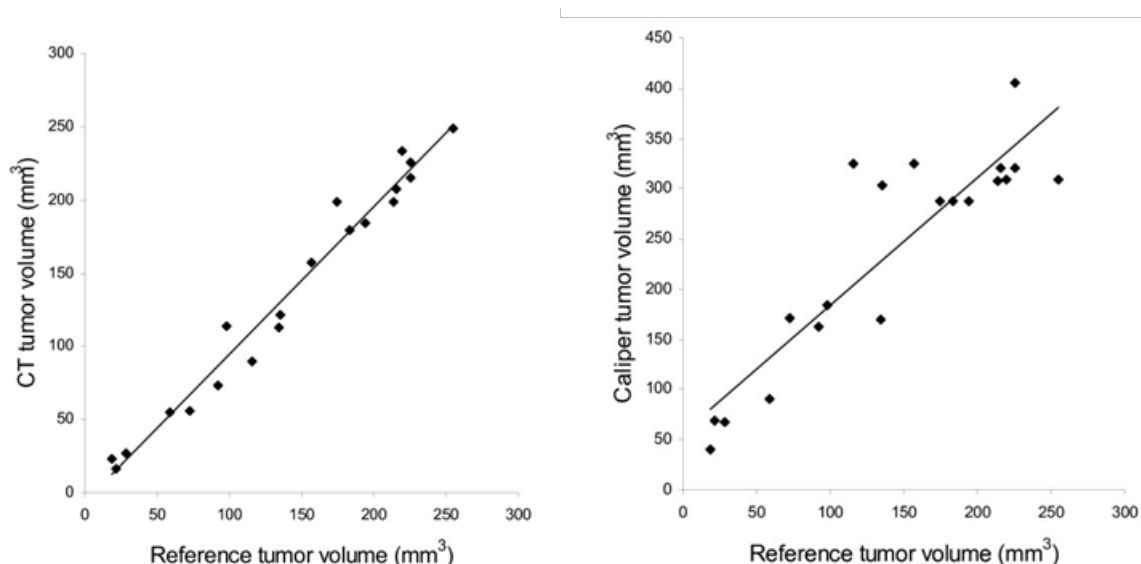
The ideal replacement for callipers would be a non-invasive system where minimal to no contact with the animal is required. From electromagnetic fields to sounds waves, tumour information can be collected without compromising tumour tissue via handling with callipers. Additionally, these methods are reverse translated from clinical use. Techniques used to find s.c. tumour volume are reviewed below and in Table 1.2 and compared with calliper measurements.

#### 1.3.2.1 MicroCT

Micro computed tomography (microCT) is an imaging system using powerful electromagnetic X-rays to provide cross sectional slices in both 2D and 3D. The different tissues and densities in the body absorb or block the X-rays at different levels and therefore the tissues can be differentiated from each other, facilitating the visualisation process[82]. Proven to be successful on bone studies, mapping the bone architecture and its different densities, the imaging modality is fast establishing itself in preclinical imaging in small animals, being used in both *in vitro* and *in vivo* studies. *In vitro* uses involve providing microscopic histological information and *in vivo* studies range from whole body to vascular imaging[83]. Furthermore, microCT is able to detect metastasis as well as measuring response

to treatment[84]. Whilst this technique has the penetrative power and resolution to image deep-lying tumours, it has also been used to evaluate surface tumours. Jensen conducted a comparative study to ascertain if the microCT imaging system could improve upon volume measurements of s.c. xenograft tumours found using callipers and the modified ellipsoid formula[81]. Xenografts were implanted subcutaneously in mice with the human breast adenocarcinoma cell line. Calliper measurements were taken and the mice were then anaesthetised and fixed onto a bed for a 7 minute CT scan. For the reference tumour volume, tumours were excised and the volume was calculated from weight and density[81].

When compared with reference tumour volume, volumes obtained from both microCT and callipers gave  $R^2$  values of 0.97 and 0.8, respectively; results can be seen in Figure 1.3. Interestingly, intra-observer variance was twice as high in calliper measurements. Evidently, microCT surpassed callipers in accuracy, reliability and reproducibility[81].



**Figure 1.3:** Linear regression with CT and calliper tumour volume.

Showing calliper tumour volume ( $y = 1.27 \pm 0.15x + 56.9 \pm 24.2$ ,  $R^2 = 0.80$ ) and microCT tumour volume ( $y = 1.01x - 6.1$ ,  $R^2 = 0.97$ ) against the reference tumour volume[81]. The microCT volume measurements have higher accuracy in comparison to the calliper volume measurements.

Although microCT surpassed callipers in volume measurements, the mouse was anaesthetised, fixed onto a bed and subjected to long imaging times[83]. This was improved upon by Abou-Alkacem who used both microCT and calliper measurements to measure the tumour growth for human colon cancer cell line implanted subcutaneously in 9 mice[85]. Although the study still used anaesthesia, the imaging time was reduced to 90 seconds per mouse[85]. When compared to the reference volume, found after excising the tumour, volumes obtained from both microCT and callipers gave  $R^2$  values of 0.93 and 0.8, respectively. Therefore, MicroCT exceeded callipers in accuracy in predicting tumour volume.

### 1.3.2.2 PET Scanning

Positron emission tomography (PET) is a type of non-invasive imaging utilising a tracer[86] to monitor physiological functions. A tracer is a substance that can be tracked via distinctive properties such as colour or radioactivity. A commonly used tracer is [<sup>18</sup>F]-fluorodeoxyglucose (FDG) which is an analogue of glucose. Glucose absorption is increased in the majority of cancers because cancer cells consume more glucose in comparison to normal cells. Therefore, there is increased expression of GLUT-1 glucose transporters[87]. PET scanning is useful in providing 3D images showing functional processes in the body and is able to identify benign tumours from malignant and the staging of the cancer[88].

Jensen used FDG-PET scanning and callipers to measure and compare s.c. tumour volume measurements. When the FDG-PET scanning and calliper measurements were compared to the reference volume, found after excising the tumour,  $R^2$  values were 0.75 and 0.8, respectively[81]. The FDG-PET was unable to reproduce the accuracy of the callipers. Moreover, the FDG-PET scanning process required the mice to be anaesthetised, fixed onto a bed and scanned for 20 minutes offering no refinement benefits[81]. Furthermore, there is increased uptake of FDG in response to an infection thus activity is recorded at the site of the infection as well as in healing bone, joints and lymph nodes[89]. Ultimately, the FDG-PET scanner is not a viable replacement for callipers.

### 1.3.2.3 Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is commonly used to measure tumour volume and track metastasis, finding the size and position of the tumour and tracking longitudinal treatment effects[84]. Tumour characteristics such as vascularisation, apoptosis and diffusion can also be determined[90]. MRI does not use any harmful ionising radiation. Instead, it exploits hydrogen atoms from fat and water molecules found in most tissues by subjecting them to a strong magnetic field and radio frequency[91]. Altering radio frequencies and magnetic field strengths enable data collection from different parts of the body allowing visualisation of different layers of the tissue[92].

MRI was used in a study measuring s.c. tumour xenograft volume, directly comparing calliper measurements with high resolution MRI images[80]. This high resolution scan took approximately 9 minutes and the mice were anaesthetised. The volume found from the high resolution images served as the reference volume; the tumours were not excised as the volumes were measured over time. The study found callipers to have a  $r < 0.92$  correlation with the reference volume[80].

Furthermore, the study went on to improve refinement benefits to mice by also finding volume from images obtained from awake, manually constrained mice where no anaesthesia was administered. A correlation of  $r > 0.96$  was found. Not only did this give improved volume measurement compared to callipers, the MRI imaging time was reduced from 9 minutes from the high resolution images to 1 minute[80].

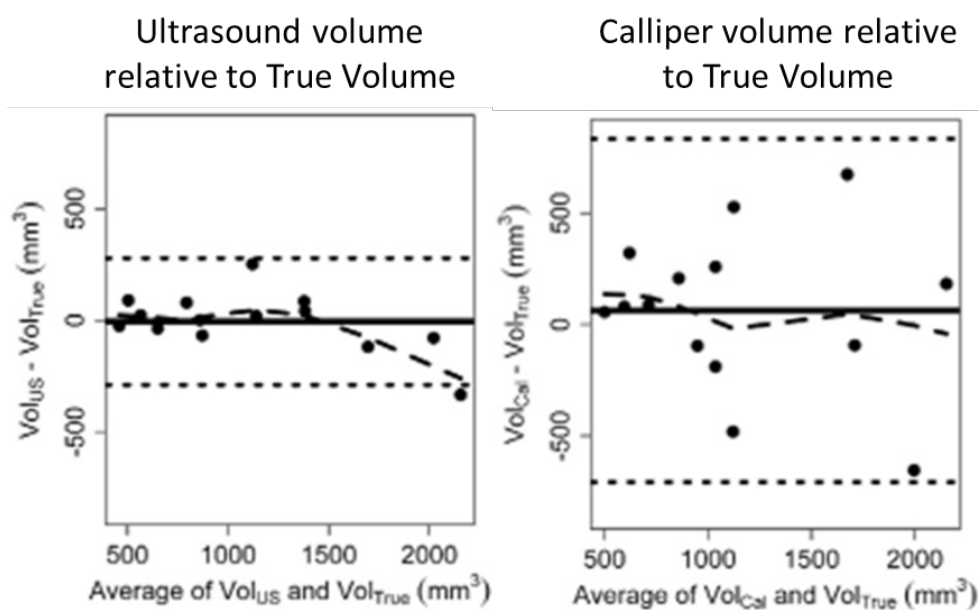
A study by Abou-Elkacem also used both MRI and calliper measurements to measure volume[85]. The reference volume used to compare the volumes found was found by excising the tumours and finding their mass. The volume found from MRI had a R2 of 0.96, whereas the calliper measurements, when compared to the MRI volume, had a R2 volume of 0.80[85]. Although MRI imaging equipment is expensive and requires a high level of training, they provide improved volume measurements compared to callipers[93].

#### **1.3.2.4 Ultrasound**

Like MRI's, ultrasound does not use harmful ionising radiation[94]. Instead, it uses high frequency sound waves. Ultrasound machines are known for their role in prenatal pregnancy but are becoming increasingly popular in preclinical imaging, providing information on anatomical imaging; more sophisticated ultrasound machines can quantify tumour blood flow and perfusion. The real-time imaging, which carries fewer costs compared to the MRI equipment, is able to provide fast information and is easier to use[84].

Ayers conducted a study comparing calliper measurements and 3-D ultrasound imaging for finding s.c. xenograft volume[94]. Once calliper measurements were taken, the mice were anaesthetised and placed on a heated stage where the xenograft was coated with a warm ultrasound gel and the scan took place producing 3-D images. Four trained observers manually identified and segmented tumour regions in the images to determine the volume. True volume was calculated once the tumour was excised and its mass determined through

direct correlation between tumour volumes found by water displacement, and the mass itself. Ultrasound deviation from the true volume was  $P=0.0005$ , significantly lower than calliper deviation at  $P=0.001$ , (Figure 1.4)[94]. Furthermore, although volume determination for ultrasound involved four observers manually determining the tumour region, the calliper volumes had a volume variance 1.3 fold higher than ultrasound measurements.



**Figure 1.4:** Findings from Ayers et al showing deviation in subcutaneous xenograft volume determined via ultrasound and callipers from the true volume.

Ultrasound deviation from the true volume was seen to be  $P=0.0005$ , significantly lower than calliper deviation at  $P=0.001$ [94].

Pfanzler et al built upon the previous study by combining 3D ultrasound imaging with speckle and 3D wall motion tracking algorithms to determine s.c. xenograft volume[95]. The study determined ultrasound measurements both *in vivo* and *ex vivo* but mice were sacrificed for both measurements, with *in vivo* measurements taking place immediately after the sacrifice. True volume was calculated *ex vivo* through determining tumour mass density, following the same method used by

Ayers et al[94]. Volume determined by ultrasound was closer to true volume in comparison to calliper measurements for 5 out of 7 tumours[95]. In summation, although ultrasound requires anaesthesia and increased animal handling when compared to callipers, it has better reproducibility and accuracy for volume prediction.

#### **1.3.2.5 Thermal Imaging**

Thermal imaging was first used clinically in the 1960's where an elevation of 1-2°C in human breast cancer tissue was seen[96]. This was believed to be due to hyper-vascularity occurring at tumour sites to accommodate for growing cell masses. In the 1970's, the Breast Cancer Detection Demonstration Projects tested thermography, as well as mammography, for diagnostic purposes but thermography was eventually abandoned in 1977[97] as thermal images were difficult to interpret due to lack of experience and technical training of staff[98], [99]. Digital infrared imaging devices are now more readily available and offer a cheaper alternative to other non-invasive imaging devices such as CT, MRI and PET scans.

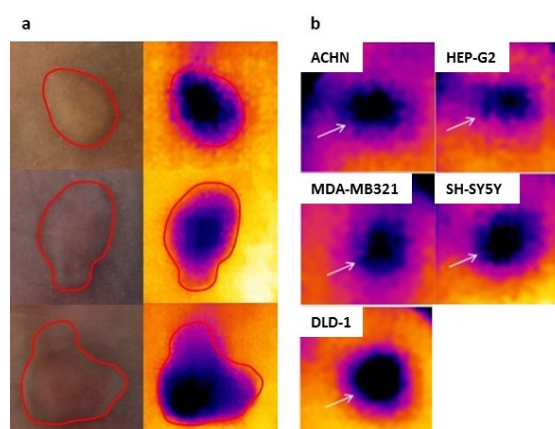
Where previous studies demonstrated a temperature elevation clinically at a human breast tumour site, studies on subcutaneously implanted experimental human breast (MDA-MB-231, MCF7)[100] and rat mammary adenocarcinoma (13762 MAT)[101] xenograft models, showed a significant decrease in temperature. Furthermore, Song et al found that tumours were detected using thermal cameras up to 3 days before they were visible to the human eye, highlighting the sensitivity of the thermal camera to pick up vascular changes very



early in tumour development[100]. They found that the temperature decreased further as the tumour grew indicating that thermal imaging could monitor tumour volume[100]. A study by Xie et al then went on to link the temperature decrease to poor vasculature in the xenograft tumours caused by rapid growth, the effect was intensified through using anti-angiogenic drugs which damaged the vasculature further and caused a further decrease in the temperature[101].

A decrease in temperature at the site of subcutaneous xenograft tumours, in multiple cell lines, was confirmed in laboratories here at Bradford by Hussain et al, as can be seen in Figure 1.5[102]. In the same study it was demonstrated that thermographic imaging can be used to not only monitor tumour volume, but also differentiate between treatment types and measure drug toxicity[102].

It took only 3 seconds to image each tumour and there is no requirement for anaesthesia or any contact with the tumour itself[102]. Although simpler than CT, MRI and PET scans, thermographic imaging equipment and relevant training would still be required. Moreover, it can be difficult to isolate the tumour outline in thermographic images (Figure 1.5b) as there is also a temperature change in the tumour microenvironment.



**Figure 1.5:** Thermographic imaging of tumours

Figure (a) shows thermal images alongside corresponding digital images, clearly showing a reduction of temperature at the tumour site. (b) Shows the effect is consistent with different tumour types[102].

### 1.3.2.6 Optical Imaging

Optical imaging is a form of non-invasive imaging, using luminescence for preclinical testing. Fluorescence imaging (FLI) is a form of these. The cancer cell lines used to create the tumour in the mice are infected with fluorescence proteins (FP). The tumours are then inflicted with light, the FP absorbs light and reaches an excited state. Next, the FP decay from the excited state and whilst doing so, emit a fluorescent light which can be measured with specialised equipment. FLI is commonly used in preclinical testing, from drug efficacy to monitoring the growth of tumours[84].

A study compared multiple modalities and found that ex-vivo tumour volume correlated highly with volume found using MRI ( $r^2 = 0.96$ ). In-vivo volume determined using callipers correlated with the MRI at ( $r^2 = 0.80$ ). The study went on to then compare MRI volume with volumes found from optical imaging, using both green and red fluorescent probing for testing. Where MRI also finds depth in an image allowing it to find a 3D volume, the FLI, similar to in vivo calliper measurements, can only find two dimensions of length and width and therefore, the modified ellipsoid formula was used to determine volume. The study found that FLI had a high correlation with MRI volume ( $r^2=0.73$ ), although not as high as the in vivo volumes ( $r^2 = 0.80$ ). Furthermore, it was found that both FLI and in vivo calliper measurements overestimated the size of tumours, indicating the

flaws in basing volume on an ellipsoidal measurement. Bioluminescence imaging (BLI) is an alternative optical imaging method using the enzyme luciferase to emit light, but no comparative studies have been done in using BLI to determine s.c. xenograft volume and compare this with calliper measurements.

### **1.3.2.7 Summary of Alternative Methods**

Whilst the non-invasive methods mentioned above all offer some advantage over the use of callipers to measure s.c. tumours, these are offset by the issues with most methodologies as discussed and summarised in, Table 1.2. The high level of training required as well as the expensive equipment has resulted in the callipers maintaining the gold standard position for measuring s.c. tumour volume, although the methodology is clearly flawed. Furthermore, some non-invasive imaging modalities require anaesthesia and some require animals to be restrained and subjected to long imaging times. This project proposes using basic digital imaging involving taking a 2D digital image on a basic personal-use digital camera, which is both cheap, accessible and can be purchased from a non-specialist outlet, to replace all previous methods. As well as being inexpensive and easy to capture images, digital imaging is easily accessible for labs throughout the world.

**Table 1.2:** Non-invasive imaging methods

Advantages and disadvantages of various non-invasive imaging modalities.

<b>Non-invasive Imaging Modality</b>	<b>Advantages</b>	<b>Disadvantages</b>
<b>CT</b>	High sensitivity anatomical images, provides 3D image.	Lower resolution, limited functional information, poor soft tissue contrast, requires expensive equipment.
<b>PET</b>	High sensitivity, 3D imaging, biochemical information.	Limited anatomical information, specialised equipment, need radio-nucleotide facilities, equipment is expensive.
<b>MRI</b>	High spatial resolution, good soft tissue contrast, provides anatomical and functional information.	Low sensitivity, relatively long acquisition time, requires expensive equipment.
<b>Ultrasound</b>	Good resolution, provides anatomical and functional information, fast and portable technique, relatively inexpensive, amenable to smaller research laboratories.	Inability to image through bone.
<b>Thermography</b>	Fast and portable technique, relatively inexpensive, amenable to smaller research laboratories.	Preclinical results differ from results found with humans.
<b>Optical (BLI and FLI)</b>	Wide applicability, relatively inexpensive, amenable to smaller research laboratories.	Requires genetic manipulation of investigated cells, limited anatomical information, reduced sensitivity with increased imaging depths.

## 1.4 Digital Imaging

Digital image capture devices ranging from non-specialist shop-bought digital cameras to smartphones, are globally accessible to all. Advantages include portability, low costs and ease of use. A digital image can be considered as a 2D function of the spatial plane co-ordinates of X and Y. Therefore, a digital image consists of a number of X and Y co-ordinates, where each co-ordinate represents an element of the image with an intensity level (or grey level).

When taking a digital image, clicking the button to take the image results in the opening of a lens which allows light through to a sensor, fragmenting the incoming light source into pixels before measuring and storing the pixel colour and brightness as a number. Thus, a digital image is collection of numbers describing pixels which form together to give an image. Images can be examined in 2 ways: image processing and machine learning. Image processing involves accessing the stored information and extracting useful information including, but not limited to, the intensity of colours or any stark differences in colours which could be indicative of boundaries. Conversely, machine learning doesn't look at singular images, but finds relationships and common patterns in multiple images. This section discusses image processing and machine learning techniques that can be utilised to enable the detection of the tumour and deciphering of its tumour volume.

## **1.4.1 Image Processing Techniques that can be applied to find 3D measurement from 2D image**

Image processing involves analysing and manipulating image features to extract information. Various image processing techniques such as colour segmentation, edge detection and histograms are explored for their usefulness in fulfilling the aim of this thesis which is to find a 3D s.c. tumour volume measurement from a 2D digital image.

### **1.4.1.1 Colour Segmentation**

Colour segmentation is an image processing technique that uses colour spaces, or defined sets of rules to define colours, to find colour differentiation within an image. This could be useful to detect s.c. tumours as there may be differences in colour between the mouse flank and the tumour. Colour spaces can be categorised into RGB, YUV, CIE and HSV groups[103]. The RGB model is popular and well known because it is commonly used in displays technology, including mobile screens, computer displays and television screens[104]. Table 1.3 shows the commonly used colour spaces, with the parameters used to determine colour and its intensity.

**Table 1.3:** Showing the commonly used colour spaces RGB, YUV, CIE and HSV with their colour definitions describing how they define and segment colour

Colour Space	Parameters
RGB	Red, green, blue
YUV	Luminance, chrominance
CIE	X, Y, Z
HSV	Hue, saturation, value

As the colour of nude mice closely resembles human skin colour, skin detection models can potentially be used to identify the mouse and possibly the tumour from a digital image. A study measured the success of the different colour spaces in skin detection[103].

Skin detection involves finding pixels corresponding to skin colour and the ability to detect these in an image or video thus forming the basis of facial recognition. The CIE and RGB models performed poorly, CIE was successful in 86.2% of cases, whilst RGB and CMY models detected 86.55% and 86.15%, respectively. The YUV colour space models ranged in their success, with YIQ having the lowest detection at 86.93% and YCgCr was best at 92.63%. The best skin detection was delivered by the HSV colour space models at 95.06%[103].

Additionally, colour segmentation allows the reduction of information in an image by separating it out into separate colour channels. The colour channels can be used to threshold an image, ensuring that any pixels outside the colour range are removed. This was seen in the work of Arslan on segmenting white blood cells in bone marrow and peripheral blood images[105]. The work found that by

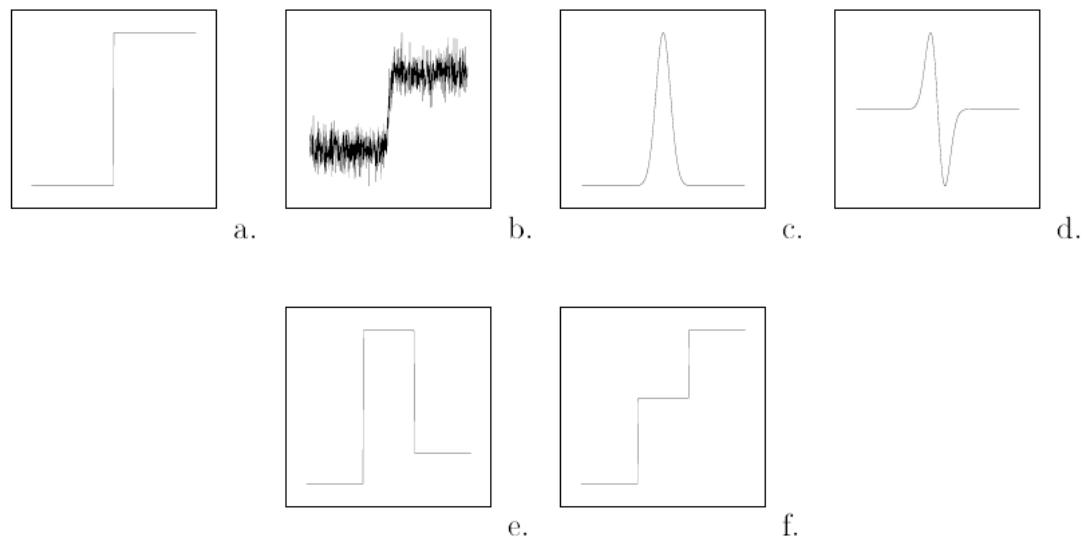
converting the RGB image to greyscale for processing, multitudes of information were lost. Instead, they applied a threshold to the images to only focus on the green band, in which the white blood cells had a greater colour intensity[105].

Focusing on different colours in an image is useful for extracting features from digital images. As well as looking at different colours, the intensity of boundaries and image regions in grayscale images can be used to outline and isolate image features using edge detection.

#### **1.4.1.2 Edges and Gradients**

Edge detection and finding image gradients are useful tools in image segmentation and can potentially be useful in differentiating the tumour from the mouse body. The technique utilises the discontinuities in image intensities to decipher boundaries once it has been converted from colour to grayscale[106]. Detectors also incorporate gradients, which look for the magnitude and directional change of the pixel intensity. Edges take the forms of steps, lines and junctions; steps are the most common form. Steps represent the difference in grey levels within a segment of the image and therefore, the grey level will reach both a maximum and minimum point, hence showing the edge[107].





**Figure 1.6:** Types of step edges

(a) Ideal step showing difference in two grey levels in an image, (b) step is corrupted with noise resulting in a smoothed step edge, (c) first order derivative of smoothed edge in (b), (d) second order derivative of (c), (e) pulse double step edge and (f) staircase double step edge[107].

When images are captured, the response of the image system to the object being captured is called a point spread function (PSF). The PSF is convoluted with the light source and the object to give a final image. This results in blurring of the object. Decreased blurring is indicative of an imaging system with better quality. Due to this, the step edges found in images are blurred and so produce smooth edges[107]. This is known as noise. Figure 1.6a. and b. represent this[107]. An image is a continuous function, and gradients are a measure of any changes in this function thus finding the gradients leads to edge detection[108].

Other ways of representing step edges is using first and second order derivatives, but these fail to show the distribution of space around the edges, shown in Figure 1.6c. and d. Finally, steps can be represented as inflection points. This means double edges are shown, showing both the high and lows in the differences

between the grey levels. The two types are pulse and staircase, shown in Figure 1.6e. and f, respectively[107]. The double step edge model is most commonly used.

Edge detectors use convolutional masks to find edges in an image, convolution masks are created by a kernel, a specified matrix, or sequence of numbers in a square format, used by the detector. Various edge detectors along with their method and specified convolution masks are described in Table 1.4.

**Table 1.4:** Edge detectors

Commonly used edge detectors including those in this study are described. Although they use different techniques, there is clear overlap in the overall method of extraction.

Edge Detector	Method (Defines Edge)	Convolutional Mask and its Use																		
<b>Sobel</b> [109]	Where image gradient is at maximum intensity	<table style="display: inline-table; margin-right: 20px;"> <tr><td>-1</td><td>0</td><td>+1</td></tr> <tr><td>-2</td><td>0</td><td>+2</td></tr> <tr><td>-1</td><td>0</td><td>+1</td></tr> </table> $X$ <table style="display: inline-table;"> <tr><td>+1</td><td>+2</td><td>+1</td></tr> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>-2</td><td>-1</td></tr> </table> $Y$ <p>Pixels in the centre. Uses mask to find edges vertical and horizontal to the pixel grid in the image.</p>	-1	0	+1	-2	0	+2	-1	0	+1	+1	+2	+1	0	0	0	-1	-2	-1
-1	0	+1																		
-2	0	+2																		
-1	0	+1																		
+1	+2	+1																		
0	0	0																		
-1	-2	-1																		
<b>Prewitt</b> [110]	Where image gradient is at maximum intensity	<table style="display: inline-table; margin-right: 20px;"> <tr><td>-1</td><td>0</td><td>+1</td></tr> <tr><td>-1</td><td>0</td><td>+1</td></tr> <tr><td>-1</td><td>0</td><td>+1</td></tr> </table> $X$ <table style="display: inline-table;"> <tr><td>+1</td><td>+1</td><td>+1</td></tr> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>-1</td><td>-1</td></tr> </table> $Y$ <p>No focus on central pixels in the mask. Finds vertical and horizontal edges to the pixel grid in the image.</p>	-1	0	+1	-1	0	+1	-1	0	+1	+1	+1	+1	0	0	0	-1	-1	-1
-1	0	+1																		
-1	0	+1																		
-1	0	+1																		
+1	+1	+1																		
0	0	0																		
-1	-1	-1																		
<b>Roberts</b> [111]	Where image gradient is at maximum intensity	<table style="display: inline-table; margin-right: 20px;"> <tr><td>+1</td><td>0</td></tr> <tr><td>0</td><td>-1</td></tr> </table> $X$ <table style="display: inline-table;"> <tr><td>0</td><td>+1</td></tr> <tr><td>-1</td><td>0</td></tr> </table> $Y$ <p>Finds edges at 45 degree angle to the pixel grid in the image.</p>	+1	0	0	-1	0	+1	-1	0										
+1	0																			
0	-1																			
0	+1																			
-1	0																			
<b>Laplacian</b> [112]	Approximates the second derivative of the gradient and finds the zero crossing; where the first derivative is at maximum. Uses Gaussian smoothing.	<table style="display: inline-table; margin-right: 20px;"> <tr><td>1</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>8</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> </table> $X$ <table style="display: inline-table;"> <tr><td>-1</td><td>2</td><td>-1</td></tr> <tr><td>2</td><td>-4</td><td>2</td></tr> <tr><td>-1</td><td>2</td><td>-1</td></tr> </table> $Y$	1	1	1	1	8	1	1	1	1	-1	2	-1	2	-4	2	-1	2	-1
1	1	1																		
1	8	1																		
1	1	1																		
-1	2	-1																		
2	-4	2																		
-1	2	-1																		
<b>Canny</b> [113],[114]	Uses Gaussian smoothing[115]. Uses Sobel detector to find gradient magnitude, which is used to find the gradient direction.	<table style="display: inline-table; margin-right: 20px;"> <tr><td>-1</td><td>0</td><td>+1</td></tr> <tr><td>-2</td><td>0</td><td>+2</td></tr> <tr><td>-1</td><td>0</td><td>+1</td></tr> </table> $X$ <table style="display: inline-table;"> <tr><td>+1</td><td>+2</td><td>+1</td></tr> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>-2</td><td>-1</td></tr> </table> $Y$ <p>Pixels in the centre. Uses mask to find edges vertical and horizontal to the pixel grid in the image.</p>	-1	0	+1	-2	0	+2	-1	0	+1	+1	+2	+1	0	0	0	-1	-2	-1
-1	0	+1																		
-2	0	+2																		
-1	0	+1																		
+1	+2	+1																		
0	0	0																		
-1	-2	-1																		

Although edge detectors are sophisticated tools in finding boundaries various other image processing techniques, such as morphology operations and histogram of gradients, can be used to contribute towards isolating an object from an image. In this case, this would be the tumour from the mouse and surrounding background.

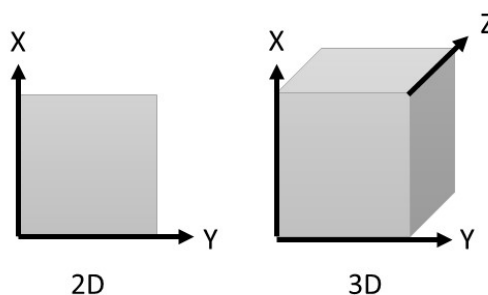
#### **1.4.1.3 Histograms and Morphology Operations**

Where edge detection uses gradients in conjunction with convolutional masks, the technique known as histogram of gradients focusses on regions of interest within an image. Within a region of interest, the area is divided into cells. Histograms are made of each cell, these are then combined and normalised. This technique shows how the gradient intensity and direction is distributed in an image.

Morphology operations process images based on their shapes. This operation consists of implementing structuring elements to an input image which results in an output image which is also the same size. Dilation and erosion are the two basic types of morphological operations. Dilation differs to erosion as it adds pixels to boundaries of an object whereas erosion removes them. The size and shape of the structuring element determines the amount of pixels that are added in dilation and removed in erosion. In dilation and erosion, a rule is used to establish the state of a pixel in the output image including the neighbouring input image. This rule is used to classify the operation as a dilation and erosion.

### 1.4.2 3D Measurement from 2D image

This project proposes to obtain volume measurements with digital imaging. However, standard digital images are 2D and volume is a 3D measurement. 2D images have two axes, width and height, represented by X and Y, respectively. 3D has a third axis, depth, represented by Z; this is visually represented in Figure 1.7. Thus, applying depth to 2D digital images of a mouse tumour would enable 3D measurement of the tumour. Multiple factors enable the extraction of depth information from a 2D image, including height, texture, light intensity, size of an object and perspectives. These factors are known as depth cues; these are key to processing the digital images and providing depth information, facilitating the 2D to 3D transformation[116]. The number of image inputs is crucial in deciding which depth cues to be used, this is summarised in Table 1.5[116].



**Figure 1.7:** 2D and 3D axes

2D showing X and Y axis and 3D showing X, Y and Z axis.

**Table 1.5:** Depth cues in single and multiple images

Number of Image Inputs	Depth Cue	Description
<b>One image</b>	Defocus	Closer objects are clearer and distant objects are blurred
	Linear perspective	Parallel lines appear to converge to distance
	Atmospheric scattering	Distant objects are less distinct and more bluish in colour
	Shading	Pixels with a brighter light intensity are closer
	Symmetric patterns	Bilateral symmetric objects from different views can be seen as different images
	Patterned textures	Distortion in image and rate of change of distortions, with distortions further away from camera appearing smaller
	Occlusions	Objects which overlap other objects are closer
<b>Two images or more</b>	Binocular disparity	Take two images from different angles and find corresponding points
	Motion	Near objects move faster than far objects do
	Focus	Multiple images with different focus levels and different distances are registered
	Defocus	Closer objects are clearer and distant objects are blurred
	Silhouette	Contour of objects which separate them from the background

The focus of this project is on obtaining depth information from a single 2D image, for maximum refinement benefits to the mouse, and therefore, depth cues from single images must be explored further.

#### 1.4.2.1 Depth Cues from Single Images

One way of finding depth is through the identification of parallel lines. Parallel lines can be analysed through linear perspective; parallel lines will appear to converge with distance, thus increased convergence represents increased distance and depth[117]. Patterned texture is another depth cue. This approach builds a 3D impression from the image, by analysing the texture of the pixels,

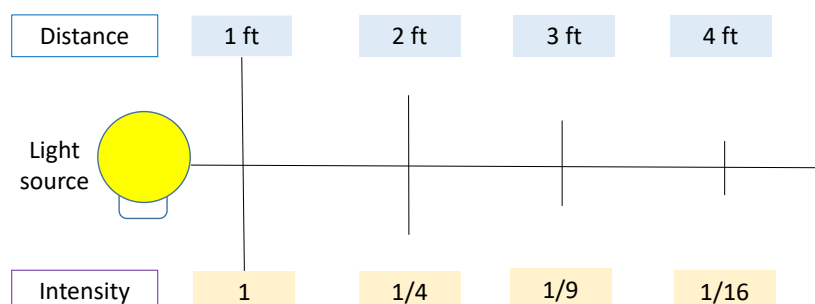
'texels'. The distortion intensity of the texels and the rate of change of distortion are used to build the 3D impression; texels at increased distance seem smaller and texels can also appear to be shortening if the texel is not parallel to the image plane[118].

Occlusions are depth cues which look at the overlapping of objects; an object is overlapped if its view is disrupted by another object. The object with its view disrupted is seen as being further away from the camera. The two depth cues in the category of occlusions are curvature and simple transform. Curvature involves segmenting the image, separating objects and shapes, and then highlighting curved objects, whereas simple transform involves mapping the path of pixel intensities[119].

#### **1.4.2.2 Shape from Shading**

Shape from shading is another depth cue that can be found in single images and instead of focussing on the shapes found in the image, this depth cue involves creating and finding depth using light intensity. It is a concept that involves using the difference in shading and light intensity on the 2D image to generate a 3D image[120]. The process uses Inverse Square Law and Lamberts Cosine Law. The inverse square law states that an increase in distance from a light source results in decreased intensity; the light intensity decreases by the square of the distance, this can be seen in Figure 1.8[121]. Light is reflected from a surface using either spatial or diffuse reflection. Spatial reflection involves a single ray being projected onto a surface and then being reflected away in a single ray with an identical angle. In diffuse reflection, the incoming ray onto the surface is

reflected back in multiple rays at varying angles and it is assumed that the ray intensity from the reflected rays is equal in all directions[120].

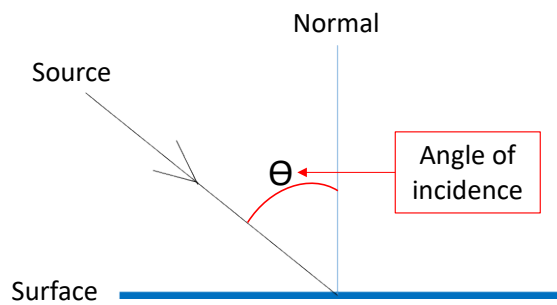


**Figure 1.8:** Diagram illustrating the inverse square law regarding light intensity.

The inverse square law states that as the distance away from the light source is increased, the light intensity decreases by the inverse square of the distance. For example, when the distance from the light source is 2ft, the light intensity is 1/4[122].

Lamberts cosine law involves the use of a surface which has diffuse reflection, and it states that the cosine of the angle of incidence between the light source, and the surface normal, which is the line perpendicular to the surface, is always directionally proportional to the luminance intensity being reflected of the surface. Figure 1.9 shows the angle of incidence ( $\Theta$ ). Shape from shading uses Lamberts law to formulate algorithms which generate a 3D shape using the lighting intensities[120]. The algorithms vary depending on the task and can be adapted according to the needs of the image.





**Figure 1.9:** Diagram to show the angle of incidence

The angle of incidence is the angle between the ray from the light source and the surface normal[122].

An example where shape from shading has been applied to a biological sample was to generate depth in images captured of small intestine villi morphology via video capsule endoscopy[123]. Coeliac disease is an autoimmune condition where the immune system attacks the component gliadin in the protein gluten, commonly found in wheat, barley and rye[124]. A characteristic feature of patients with coeliac disease is villous atrophy and instead of having uniform folds in the villi, coeliac patients have a mosaic appearance[124]. A sequence of images were captured using video capsule endoscopy in which a camera and light are placed adjacent to each other. There is a direct relationship between the reflectance map and image brightness. Brighter pixels are closer to the camera in comparison to darker pixels which are further away[122].

In order to apply shape from shading an algorithm was developed to detect the protruding areas in the intestinal lumen, which are the villi. A 9x9 pixel square was moved around the image and when the central pixel was brighter than the surrounding pixels, still contained in the 9x9 square, it was identified as a

protrusion[123]. The mean pixel brightness within the 9x9 square was measured. Successively larger pixel squares (10x10, 11x11 etc.) were made and they also had their mean pixel brightness measured. The size of the square was increased until a minimum brightest was reached. The dimension of the pixel square with minimum brightness was recorded as the width. The difference in average brightness levels between the maximum and minimum brightness squares was recorded as the height. The spatial relationship between protrusions was also quantified by finding the mean absolute difference between a protrusion and its four nearest neighbours[123].

The rendered image was successful in creating depth from the corresponding grayscale 2D image. There was a clear difference in the villi morphology between coeliac patients and control patients with coeliac patients having larger and more irregular sized protrusions[123]. As previously mentioned, the principles of shape from shading can be applied to different images and output different information. As an s.c. tumour has increased height in comparison to its surrounding, it is expected that the introduction of light would cause changes in the pixel intensity that could differentiate the tumour from the mouse flank.

Where the techniques mentioned above are based on image processing where the focus is on a single image feature at a time, the next section discusses machine learning where multiple image features are utilised to achieve a task.

### 1.4.3 Machine Learning

Machine learning involves teaching a computer to identify patterns in data in comparison to hard-coding instructions to find specific features in data, as is the case with image processing. It is a sophisticated classification process where objects can be categorised. The two types of machine learning algorithms used in this project are cascade classification and convolutional neural networks, furthermore an extra layer of intelligence is added to the two via the use of support vector machine (SVM). SVM classifiers separate data using an optimum hyperplane usually computed by solving a quadratic optimisation problem. For instance, two input variables create a two dimensional space, and in such a situation, the hyperplane is a line that separates the two input variables into two different categories. For the purpose of this project, the categories would be tumour and non-tumour.

#### 1.4.3.1 Cascade Classification

Cascade classification is a method of object detection that uses multiple cascades, or stages, to train classifiers, known as weak learners, with positive and negative images. The classifiers are boosted through learning with every cascade thus transforming the classifiers into strong learners with increased predictive power. The first cascade classification algorithm used in object detection on grey images was developed by Paul Viola and Michael Jones in 2001 for face detection[125]. The novelty was in the classifiers searching for features in the image rather than looking at the pixels themselves. The features were rectangle based shaded in either grey or white, where the sum of white

pixels are subtracted from the sum of grey pixels. The picture itself from which the object detection occurs, is transformed into an integral image, where the image is separated into quadrants and the lower right hand corner of each quadrant is a sum of all the pixels in the quadrants above and to the left of it. The integral image allows for fast and efficient computation[125].

A final step is cascading the classifiers using AdaBoost, a technique developed by Freund and Schapire. It is known as a boosting algorithm because multiple weak classifiers are trained with random sub-sets of the training data set. The classifiers with higher accuracy are given higher weights and so will have bigger impact on the next stage of the classification whereas the classifiers with low weights will have low impact on the next stage of the classification. Thus, regions which do not contain the region of interest are quickly discarded as the classifiers working on such areas will have low weights assigned to them. Viola and Jones used the AdaBoost algorithm to not only train the classifiers but also select the best features in the image, as with every classification, the rectangle that performed best was selected. Thus a combination of the best features could be used with each iteration of the classifications. The study achieved object detection 15 times faster than any previous method described[125].

#### **1.4.3.2 Convolutional Neural Networks**

A neural network transmits information from an input to an output through many different directions. This function is performed through a series of hidden layers which consist of a group of neurons. Despite each neuron being linked to all the other neurons in the preceding layer, it still functions independently. The output

layer is located in the final layer of the neural network. This is where the information is classified and a raw score is obtained. Although the classical neural networks (multilayer perceptron) can classify information to detect objects, they are not advanced enough to scale to full size image. Convolutional neural networks (CNN) have addressed this issue by arranging the groups of neurons in 3 dimensions to represent the width, height and depth of an image. Furthermore, CNNs have achieved success due to weight sharing, and feature space reduction (pooling).

A CNN architecture is derived from 3 main layers: Convolutional (CONV) Layer, Pooling Layer, and Fully-Connected Layer. A possible CNN architecture can consist of 5 layers; INPUT, CONV, RELU, POOL and FC. The neurons in the INPUT layer will store the raw pixel value of an image. The, CONV layer, produces a dot product by calculating the output of the neurons from the weights and the location of the neuron connection during input. In the RELU (rectified linear unit) layer an activation function is applied where the input signal is converted to a nonlinear output signal which is then used in the next layer. Each layer transforms information to the next and by doing so it increases the number of dimensions which can result in the distortion of an image. Therefore, before the final step, a down sampling operation is performed in the POOL layer to reduce these number of dimensions, this technique has a positive effect of making the neural networks rotation invariant. Finally the FC (fully-connected) layer will calculate the raw score. This output is used to predict detection.

The Oxford Visual Geometry Group (VGG) created the popular and commonly used CNNs VGG-16, VGG-19 and VGG-face[126]. VGG-16 has a total of 37

layers out of which there are 13 convolution layers and 3 FC layers and all the convolution layers are equipped with 3 x 3 size filters as depicted in figure 1. The remaining layers are activation and pooling layers and the last layer is the decision layer. This deep neural network was trained on ImageNet database containing over 1.2million images of 1000 categories of objects. In 2015, the VGG-16 model was retrained on a dataset of 2.6 million human faces, this they named the VGG-Face; it has been reported to have achieved state of the art face recognition accuracy. VGG-19 is a deeper version of the VGG-16, the depth of the network was increased reaching up to 43 layers comprising of 19 convolution layers, 3 FC layers, the rest are interweaved activation and pooling layers.

### **1.5 Aim and Objectives**

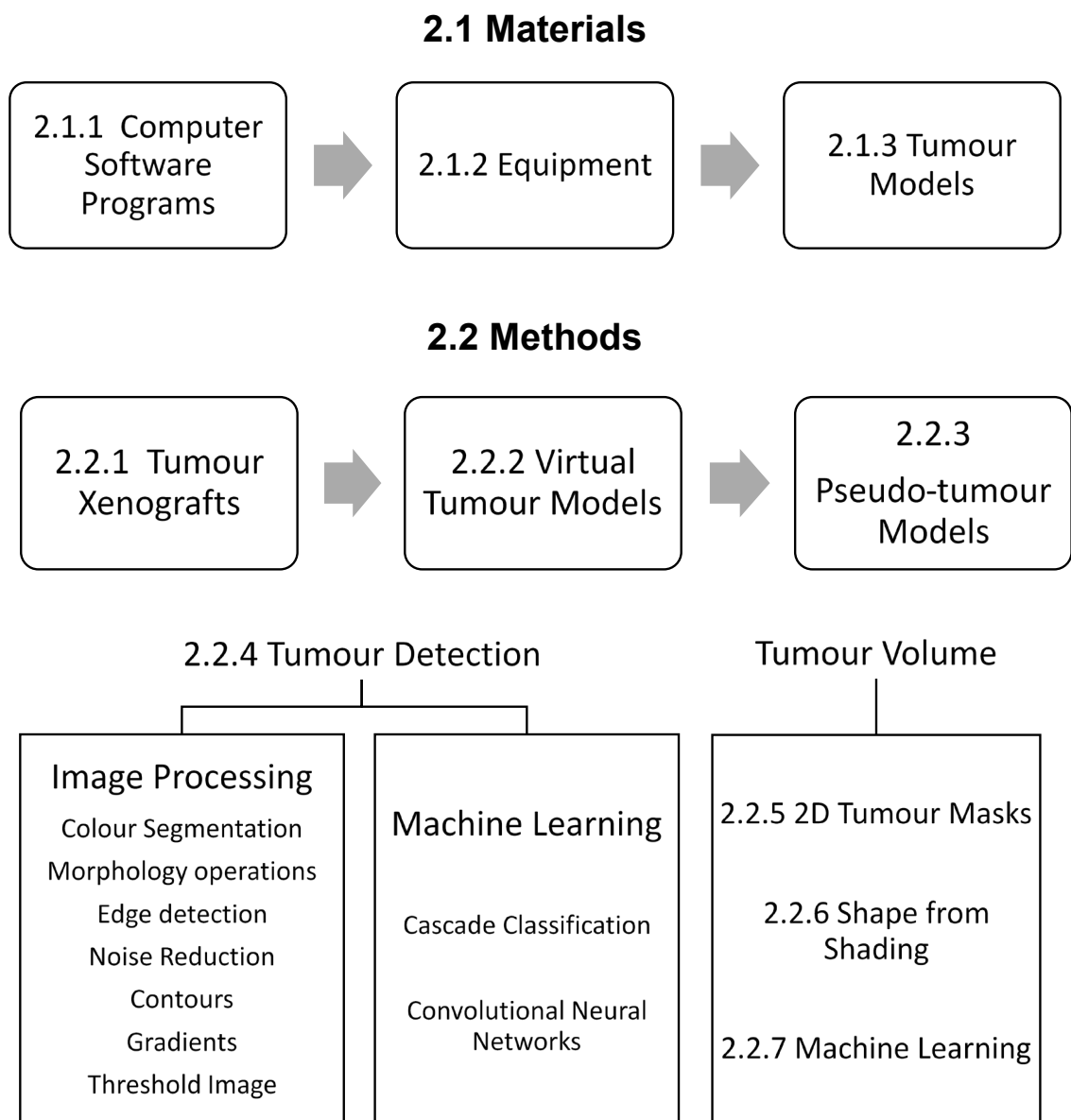
As has been reviewed there are several problems using callipers to monitor tumour growth in subcutaneous xenograft models. Therefore the aims of this project are to develop image analysis software which can calculate the xenograft volume in 3D from 2D images thus refining the method for tumour measurement and removing physical contact with the tumour. Nude mice were used in this thesis as the challenges in differentiating the tumour from the surrounding skin in hairy mice would be increased due to the overlying fur causing further disruption of the contour of the tumour thus increasing the difficulty of extracting tumour measurements. The aim will be achieved through the following objectives:

- Use of image processing and machine learning to detect and isolate the tumour from the image through an automated process.

- Calculation of volume from the 2D image through using relatively simple image manipulation techniques to correlate 2D image analysis using the modified ellipsoid formula with calliper measurements.
- Validation of the process using pseudo-tumours where the true volume is obtained through 3D scanning, and this is compared with calculation of the volume obtained from the 2D image and callipers.
- Using shape from shading, to create virtual models with different lighting intensities to calculate volume through using regression analysis with volume and pixel features.
- Predict tumour volume from 2D digital images through using machine learning and partial least squares regression analysis and compare these with tumour volumes found from calliper measurements.

## 2. Materials and Methods

This Chapter presents the materials and sequential methods used to achieve the aim of detecting an s.c xenograft tumour in a 2D image and deciphering its volume. As can be seen in the flowchart below Figure 2.1, the sections start with describing techniques used for tumour detection leading onto methods to determine tumour volume.



**Figure 2.1:** Flowchart showing the techniques used in determining tumour detection and tumour volume and the materials used to achieve this.



## 2.1 Materials

### 2.1.1 Computer Software Programs

A variety of specialist computer software programs were used to ensure best practice in achieving the aims and objectives. Program specifications and contributions to the project are disclosed in Table 2.1.

**Table 2.1:** Summary of the computer software programs used in this project.

<b>Computer Software Program</b>	<b>Version Used</b>	<b>Description</b>	<b>Contributions</b>
<b>ImageJ</b>	ImageJ 1.49, open source, National Institutes of Health, USA	Image processing program designed for scientific images	Generate tumour masks from digital images of xenograft and pseudo-tumours.
<b>Matlab</b>	Matlab 2014a, developed by MathWorks, USA	Computing language used to develop algorithms, visualise and analyse data and numeric computation	Build algorithms for image processing, machine learning and statistical analysis.
<b>Maya</b>	Maya 2015, developed by Autodesk, USA	3D animation and modelling software	Find volume of 3D pseudo-tumour models; create polygons with different degrees of shading and volumes.
<b>PolyWorks</b>	PolyWorks 2014 IR8, developed by InnovMetric Software, Canada	3D scanning visual feedback and 3D modelling	Used to generate 3D polygons from scanning pseudo-tumour models.

### 2.1.2 Equipment

The specifications and contributions of equipment utilised in the project are detailed in Table 2.2.

**Table 2.2:** Description of equipment used in the project

<b>Equipment</b>	<b>Specifications</b>	<b>Contributions</b>
<b>Laptop</b>	Latitude E5540, Dell, UK	Data collection, analysis and reporting
<b>Camera</b>	Panasonic Lumix DMC-SZ1, UK	Captured all digital images of Xenograft tumours, pseudo-tumours and artificial models
<b>3D Scanner</b>	Faro Quantum Arm, UK	Scanned pseudo-tumours, with results plotted on the accompanying software, PolyWorks.

### 2.1.3 Tumour Models

The project aims are to replace calliper measurements of s.c. tumours in mice with a non-invasive digital imaging system. Nude mice were used in this thesis as the challenges in differentiating the tumour from the surrounding skin in hairy mice would be increased due to the overlying fur causing further disruption of the contour of the tumour thus increasing the difficulty of extracting tumour measurements. Access to images of s.c. tumours in mice is limited, since it is not ethically acceptable to grow tumours solely for the purpose of acquiring imaging data. Therefore, in addition to tumour xenografts, pseudo-tumours and virtual tumours were developed to allow for a larger dataset to be studied. The following table details the different tumour types used in this project, specifying the materials they were developed with and their uses in the project.

**Table 2.3:** Description of the different tumour types used in the project

<b>Tumour Type</b>	<b>Materials</b>	<b>Uses</b>
<b>Tumour Xenografts</b>	H460 cell line derived from a non-small cell lung carcinoma. Purchased from LGC Promochem (Middlesex, UK)	To compare volume derived from calliper measurements with volume found through digital images using computational techniques.
<b>Virtual</b>	Created using Autodesk Maya as polygons (shapes with vertices and edges), or non-uniform rational B-splines (Nurbs), a type of geometry used to create 3D curves. The models were textured with Lambert material. Light sources included an ambient light to represent the ceiling and a directional light to create shading.	To investigate the relationship between pixel intensity and volume. To provide images of tumours with known volumes.
<b>Pseudo-tumour</b>	Blu-Tack, Bostik, UK. PlayDoh, Hasbro, USA Plasticine, Harbutt, UK	To compare volume measurements derived from calliper measurements and 3D scanning with volume calculated directly from a digital image of the pseudo-tumour. To investigate the effect of pixel intensity using height. To provide images to increase the dataset of tumours.

## 2.2 Methods

### 2.2.1 Tumour Xenografts

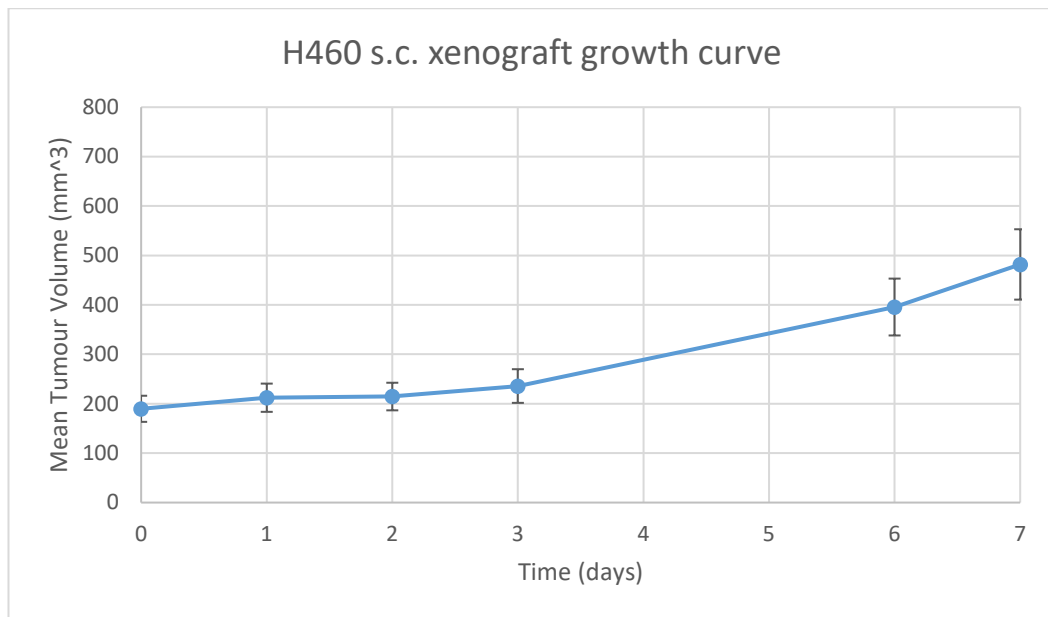
#### 2.2.1.1 Establishing Tumour Xenografts

Balb/c immunodeficient nude mice (Envigo, Loughborough, U.K.), between the ages of 6 and 8 weeks were used. Throughout the study, all mice were housed in air-conditioned rooms in facilities approved by the United Kingdom Home Office to meet all current regulations and standards, with mouse having access to food and water *ad libitum*. All procedures were carried out under a Project

Licence (PPL 40/3670) issued by the UK Home Office according to government legislation, following approval of the work by the local Animal Welfare Ethics Review Board at the University of Bradford, and in accordance with the UK National Cancer Research Institute Guidelines for the Welfare of Animals[127]. Tumour xenografts were established as follows: under brief general inhalation anaesthesia, 2 to 3 mm<sup>3</sup> fragments of donor human tumour xenograft tumours were implanted subcutaneously in the abdominal flanks. Once tumours were palpable, tumours were measured frequently by callipers as described below, and animal bodyweight also recorded to check for any deleterious effects.

As the focus of this thesis was to collect images of sample images of tumours that could be used for image processing. It was always the intention to 'piggy-back' on ongoing studies, so that no tumours were being grown up especially for this purpose. Therefore images were collected from animals being used in cancer research studies ongoing at the Institute of Cancer Therapeutics, Bradford. Images were captured from several tumour types being used in their studies, including xenografts grown using renal (ACHN), colon (DLD-1), liver (HEP-G2), breast (MDA-MB-231), head and neck squamous cell (OSC-19), neuroblastoma (SH-SY5Y) and non-small cell lung (H460) human tumour cell lines.

The H460 study was the only study where images were captured from each tumour over the duration of the study, and fig 2.2 shows the growth curve as measured by callipers.



**Figure 2.2:** H460 s.c xenograft growth curve as measured by callipers

### 2.2.1.2 Measuring Volume with Callipers

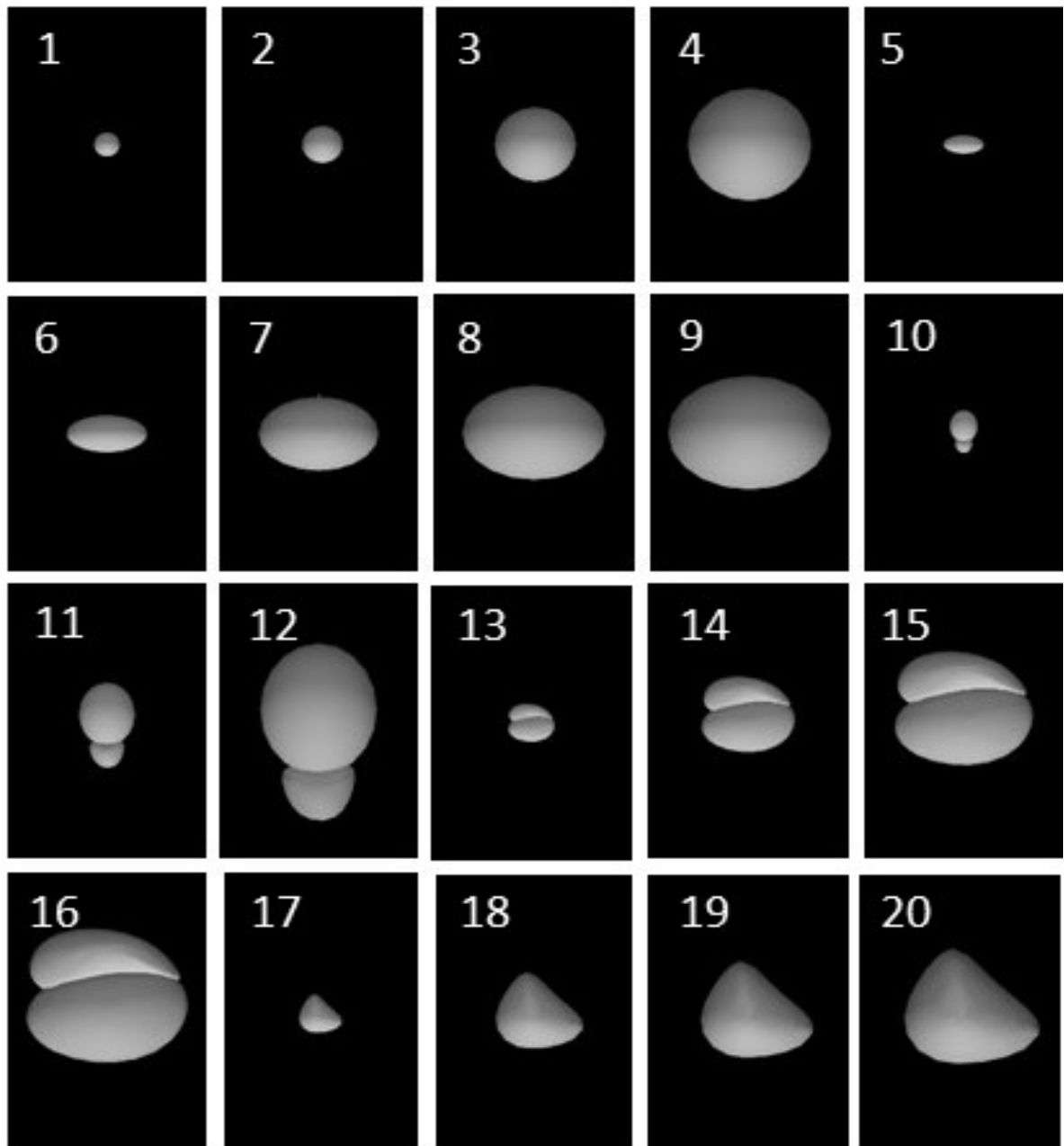
Calliper measurements were taken of the tumours using Vernier callipers. The maximum length (L) of the tumour and the perpendicular width (W) were measured. Readings were taken to the nearest 0.5mm and inputted into the modified ellipsoid formula[78] generating a volume in mm<sup>3</sup>.

$$\text{Modified ellipsoid formula: volume} = \frac{1}{2} (L \times W^2)$$

## 2.2.2 Virtual Tumour Models

### 2.2.2.1 Model creation

Virtual models were created to investigate the relationship between lighting intensity and volume measurements, and to determine if shading could be used to predict tumour volume. Virtual models (Figure 2.3) were fashioned to represent the variety of tumour sizes and shapes found *in vivo*.



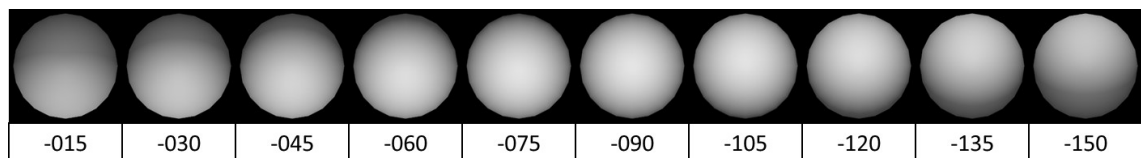
**Figure 2.3:** Virtual Tumour Models

20 models created using software Autodesk Maya to represent the variety of shapes and sizes found *in vivo*. Models have an ambient light shining on them to represent a ceiling light.

Models were constructed with polygons, which are pre-set shapes with defined vertices and edges, and Nurbs, which are shapes generated with curves, using the software Autodesk Maya. A polygon with the pre-set shape of a sphere was

used to create a set of spherical tumours through increasing and decreasing the size of the sphere. As Nurbs are easy to manipulate, these were used to create abstract tumour shapes. The models were textured with Lambert[128], a type of material displaying a matte surface, representative of the skin of nude mice.

To simulate the imaging environment of s.c. tumours, every model was exposed to ambient light, a central light source representative of the position of the ceiling light present in the theatre where the mouse tumour xenograft images were captured. An additional directional light was projected at 10 different angles on the X axis to create varying degrees of shading on the surface of the model. The angles, Figure 2.4, were -15, -30, -45, -60, -75, -90, -105, -120, -135 and -150. Rather than selecting a base model, all the angles were used in testing to maximise information obtained from the models.



**Figure 2.4:** Representative virtual model showing lighting angles along the X axis

The virtual models were subjected to a directional light applied to the surface of the model at angles ranging between -015X and -150X in increments of 10. An image was captured for each lighting angle and this enabled the measurement of the relationship between pixel intensities caused by lighting angles with the model volume.

#### 2.2.2.2 Volume Determination

Virtual model volume was calculated in Maya with the divergence theorem, through calling the function '*computePolysetVolume*'. Divergence theorem, also

known as Gauss's theorem, is a mathematical concept showing the relationship between the flux of a vector field through a surface and its actions inside a surface[129]. The divergence theorem states that change in density within a given space is dependent on flow into or out of the region through a boundary. As this algorithm only works on polygons, Nurbs were converted to polygons before the volume was calculated.

### **2.2.2.3 2D Image Capture**

Saving a 2D image from a 3D virtual model is known as rendering. To do this, the models were positioned centrally on the default grid in Maya. The viewing angle was shifted to top view, ensuring only the top of the model could be seen, making it comparable to images taken of s.c. tumours and pseudo-tumours. The 'Rendering' tab was opened from the menu and 'Render current frame' was selected; this displayed a 2D representation of the current view of the model. The image was saved by selecting 'File' from the menu tab and selecting 'Save Image'.

## **2.2.3 Pseudo-tumour Models**

### **2.2.3.1 Pseudo-tumour Model Optimisation**

As access to s.c. tumours was limited, pseudo-tumour models were also created in order to compare different methods for finding volume including calliper measurements, digital images and 3D scanning. Additionally, models were created to investigate pixel intensity through creating models of different heights, and also to increase the size of the dataset of images for further analyses.



A variety of materials, surface colours, model colours and scanning methods were tested to ensure compatibility with the 3D scanner used to measure actual model volume. The scanner was calibrated to find the best noise threshold and exposure values suited to the model feature being tested to ensure maximum performance. Additionally, a variety of model sizes were tested to ensure the models would reflect the diversity found *in vivo*.

Materials used in preliminary testing, as set out in Table 2.4 below, included Blu-tack, Plasticine and PlayDoh. All three materials were easy to manipulate and fashion into tumours. However, Plasticine did not hold its shape firmly thus making it unsuitable, as the shape could potentially distort through handling. Furthermore, although PlayDoh did hold its shape firmly, over time the material dried out resulting in a rigid inflexible model, unrepresentative of a tumour. The Blu-tack was the most suitable as it was both easy to manipulate whilst holding its shape firmly.

With regards to the colour of the pseudo-tumour, all material colours were scanned successfully. The suitability of the scanning surface was also an important factor as it had potential to interfere with the readings. Multiple surface colours were tested and all were successful except for black. The table itself was preferable as the scanning surface as the solid foundation was better for handling the delicate pseudo-tumour models.

**Table 2.4:** Preliminary testing for artificial model scanning

Finding optimum conditions for creating and scanning the pseudo-tumours to ensure accurate and reliable scanning results

<b>Testing Factor</b>	<b>Details</b>	<b>Result</b>
<b>Material</b>	Blu-Tack	Easy to manipulate and held shape firmly.
	Plasticine	Easy to manipulate but did not hold shape firmly.
	PlayDoh	Easy to manipulate but dried out, not useful for prolonged use.
<b>Material colour</b>	Blue, pink, purple, red, green, blue, yellow, orange and white.	All colours scanned successfully.
<b>Colour of scanning surface</b>	White, black, yellow, orange, green, red, blue and wooden table.	Black surface was unsuccessful, other colours were scanned successfully but the wooden table was best to add and remove models without disturbing the model shape.
<b>Model scanning method</b>	Full - The top of the model was scanned, then turned over and the bottom was scanned separately. The two were then aligned.	This did not work with the smaller tumours as the shape slightly changed when the model was turned over to scan the second view. Some alignment carried errors resulting in rough surfaces with broken edges.
	Flat - The models edges were gently spread onto a flat layer of Blu-tack to replicate tumour and the skin.	The scanner was able to accurately represent the model as a point cloud but the model would have to be separated from the surrounding material for digital image processing thus introducing error into the volume measurements.
	Raised - The model was raised on a pinhead and a single scan was done.	As the model was raised the bottom curvature of the model was also scanned and the bottom of the model was purposefully made flat so it can be hole filled later. This accurately represents the true size of the model without needing alignment.
<b>Model sizes (mm)</b>	25x40, 20x30, 12x20, 10x15, 7x10, 3x4	All sizes were scanned successfully, the dimensions were reproduced in the model.

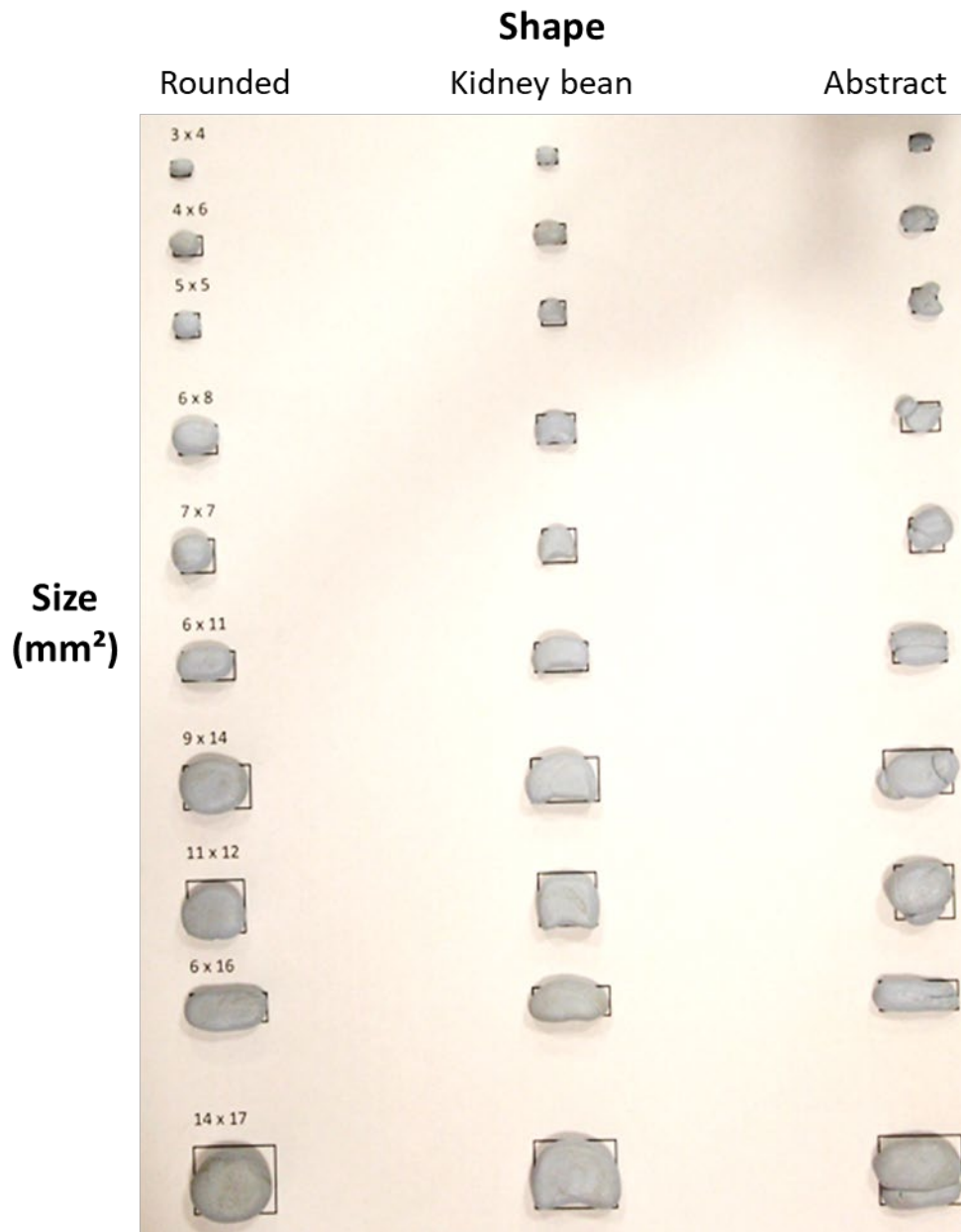
Three different methods of scanning the pseudo-tumours were tested. Firstly, the 'Full' method involved scanning the entire pseudo-tumour in two sessions, first scanning the top before flipping it over and scanning the bottom. The two scans were aligned using the software PolyWorks but this introduced errors and jagged

edges. The next method was the 'Flat' method. The pseudo-tumour was carefully placed on a layer of Blu-tack representative of a xenograft tumour and its surrounding skin. Although successfully scanned, difficulty separating the pseudo-tumour from the surrounding material using PolyWorks resulted in the loss of data from the bottom and sides of the pseudo-tumour. Finally, the 'Raised' method involved raising the model on a pinhead allowing for a single view scan of both the top and bottom in a single scanning session. The models were purposefully flattened from the bottom ensuring increased accuracy when completing the pseudo-tumour structure using PolyWorks. Finally, multiple pseudo-tumour sizes, representative of the variety found in vivo, were tested and all were scanned successfully.

### **2.2.3.2 Pseudo-tumour Volume Measurement**

#### **Pseudo-tumour Models for Volume Measurements**

Pseudo-tumour models created to allow volume measurements were designed with 10 different sizes; each size had three shapes, resulting in 30 models, (Figure 2.5). The models, measured in mm, were sized: 3x4, 4x6, 5x5, 6x8, 7x7, 6x11, 9x14, 11x12, 6x16, and 14x17 and fashioned as a rounded shape, a fat kidney bean shape and finally an abstract shape reflective of the variety found in tumours. The model was stabilised on a pinhead and a single view scan was taken. Further details are given in section 2.2.3.2.



**Figure 2.5:** Pseudo-tumour models used for volume measurements

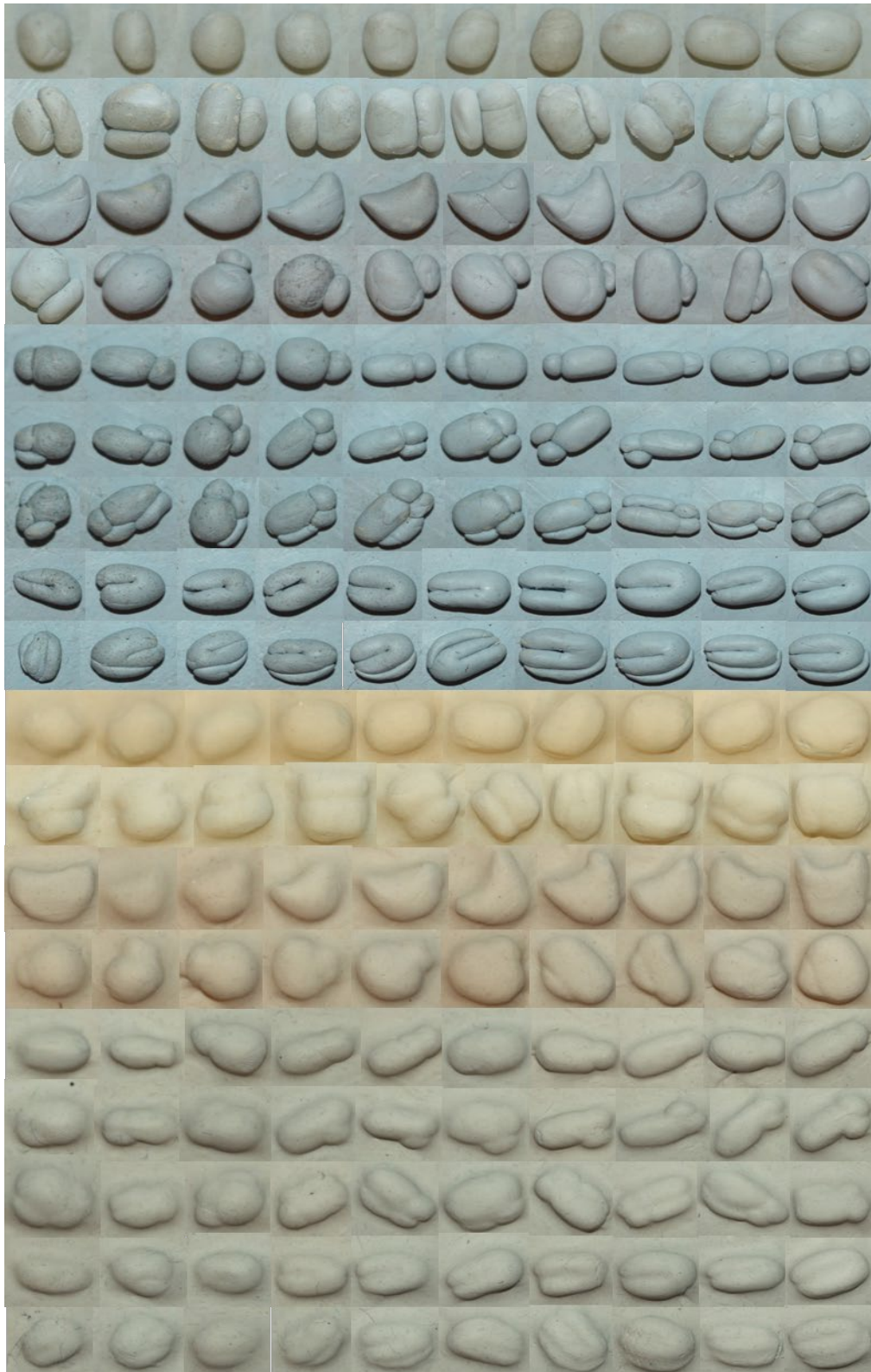
30 pseudo-tumour models were fashioned to represent the variety found *in vivo*. Each row has three pseudo-tumours of the same size, with the size defined on the first column. The first column shows a rounded shape, the second column shows a kidney bean shape and the final column shows an abstract shape.

### Models for Deciphering Height

To investigate the relationship between height and pixel intensity, 8 uniform models were created at different heights (mm): 7, 10, 13, 15, 16, 23 and 25. These were then imaged, from the same height, both with and without flash. The image without flash had no additional light source whereas the image with the flash had an additional light source emitted from the camera at the time of taking the image.

### Pseudo-tumour Models for Imaging

To increase the number of images available for processing, 180 models, Figure 2.6, were fashioned to represent the variety of tumours found *in vivo*. Both Blu-Tack and white Play-Doh were used to create the pseudo-tumours. As the purpose of the pseudo-tumours was to increase the number of images, the Play-Doh drying out had no negative impact and the white colour provided a closer match to s.c. tumours. Images of the models were captured through stabilising the camera using a tripod. A Vernier calliper was used to take maximum length and perpendicular width measurements; these were inputted into the modified ellipsoid formula (see section 2.2.1.2), to give a volume comparable to the s.c. tumour volumes found using a calliper.



**Figure 2.6:** Pseudo-Tumour models used in the image database.

180 pseudo-tumour models created using Original Blu-Tack and white Play-Doh to build up the image database of images that have both digital and calliper volume measurements.

### 2.2.3.3 3D Scanning of pseudo-tumour models

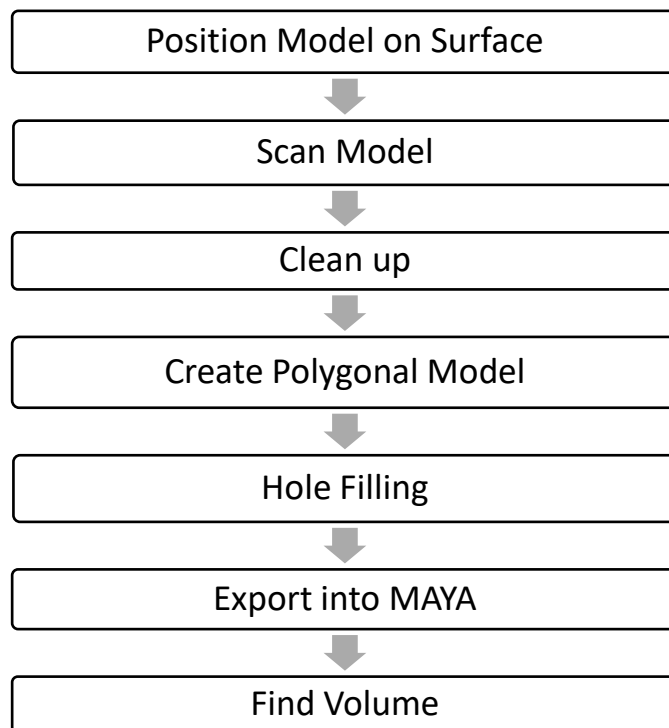
The FARO QUANTUM ARM, illustrated in Figure 2.7, is a handheld 3D scanning device that uses triangulation to scan an object. The handheld device has a laser light accompanied by a camera which can detect the position of the laser light. The distance between the laser light touching the object and the camera is known and using this information, a point cloud is plotted on the accompanying software (PolyWorks, Table 5). Calibration before the scanning session begins ensured that the points were plotted in the correct space. Figure 2.8 shows the steps for the scanning process.



**Figure 2.7:** Image of 3D FARO QUANTUM ARM scanner

Once the scanning device was calibrated the artificial model was positioned on the surface and scanned. A clean-up process in PolyWorks ensured all extra points which did not belong to the artificial model were removed. The polygon model was then created, where the plotted points were connected in the form of



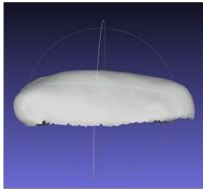
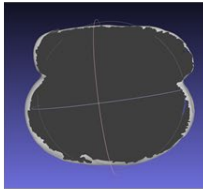
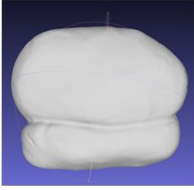

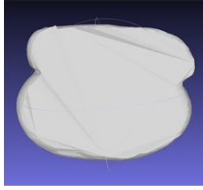
triangles. Any holes in the reconstruction were filled within PolyWorks. The process was automated and ensured any gaps or holes in the model would be covered, consistent with the model shape, ensuring accurate analysis for volume. For complex models the process was optimised by increasing the number of triangles in the polygon as well as dividing the triangles by their edge length thus making it easier to identify holes and fill them without deviating from the original structure. Figure 2.7 shows the digital image of a pseudo-tumour, with its corresponding 3D scan model both before and after the hole filling process took place.



**Figure 2.8:** Scanning Process

Once the FARO Quantum arm scanner was calibrated, the pseudo-tumour models were positioned and scanned. The scans were digitally cleaned and formed into polygonal models. Any holes found were filled before exporting the models into the software Autodesk Maya to find its volume, which was then compared with volume found through using calliper measurements and digital imaging methods.



View:	Above	Side	Below
Model: Size 14x17mm <sup>3</sup> Shape: Abstract Digital image:			
Scanned Model:			
After Hole filling:			

**Figure 2.9:** An illustration of a pseudo-tumour model and its corresponding 3D scan model, before and after the hole-filling process.

The Pseudo-tumour model was scanned using the FARO Quantum arm scanner and the figure shows the scanned model from above, the side and below. The model was then subjected to the hole filling process within PolyWorks and the resultant model is shown, again from above, the side, and bottom view.

Once the polygon model was complete, it was exported into Autodesk Maya to calculate the volume. The divergence theorem was used to calculate the volume using the command, '*computePolysetVolume.*' See section 2.2.2.2 *Volume Determination* for further information.

#### 2.2.4 Tumour Detection

This section describes the image processing and machine learning techniques used when attempting to detect and isolate the tumour from the rest of the image. All algorithms and techniques were implemented using MATLAB unless stated

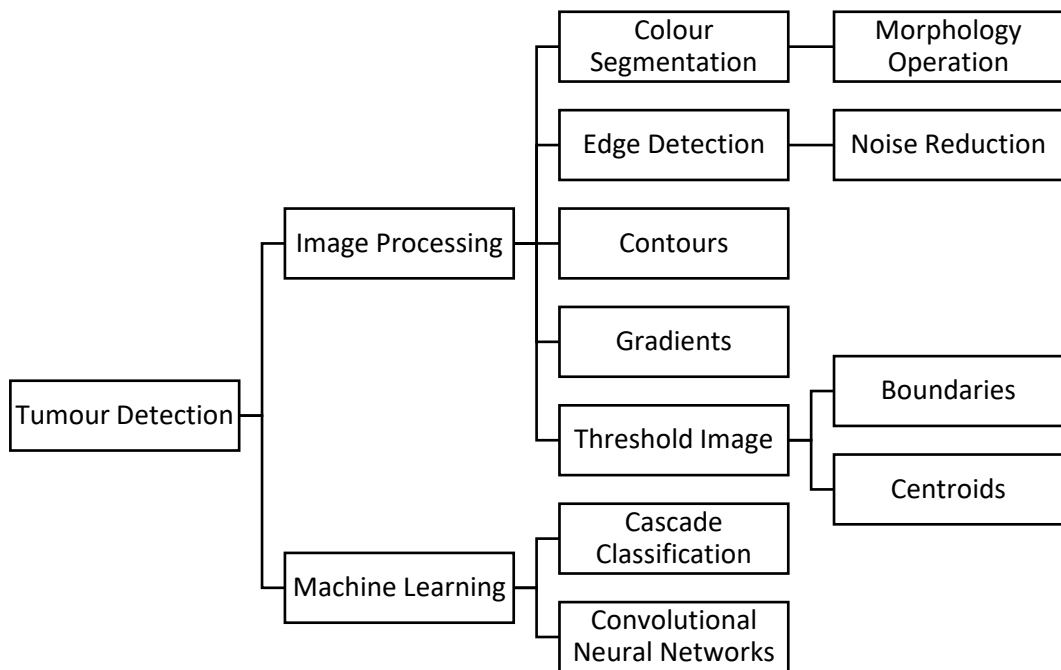
otherwise. From the s.c. xenograft images, the following captured image was used for testing (representative image in Figure 2.10). If the test was successful, the techniques were then implemented on all images collected from the study.



**Figure 2.10:** Representative primary mouse image used for image processing and detection

The following flowchart (Figure 2.11) illustrates the sequence of methodologies applied to the original image to isolate the mouse body from the rest of the image and outline the tumour boundary. Image processing techniques begin with colour segmentation which was used to isolate the mouse from the rest of image through colour differentiation. Morphological operations were applied to trim pixels at object boundaries and thus contribute to trimming the mouse body in order to centre onto the tumour. Edge detection was applied directly to the original image and used to determine the tumour boundary. Noise reduction was applied to improve the edge detection process. Contours group image regions with similar pixel intensities, and so it was used to isolate the tumour from the mouse body. Gradients were used to find the tumour boundary as they identify pixel intensity and direction. Finally, a threshold was applied to the original image to limit the

pixels on display. Further techniques for finding the image boundary and centroid, central pixel region, were used to identify and isolate the tumour. The next stage in tumour detection was to use machine learning techniques including cascade classification and convolutional neural networks (CNNs).



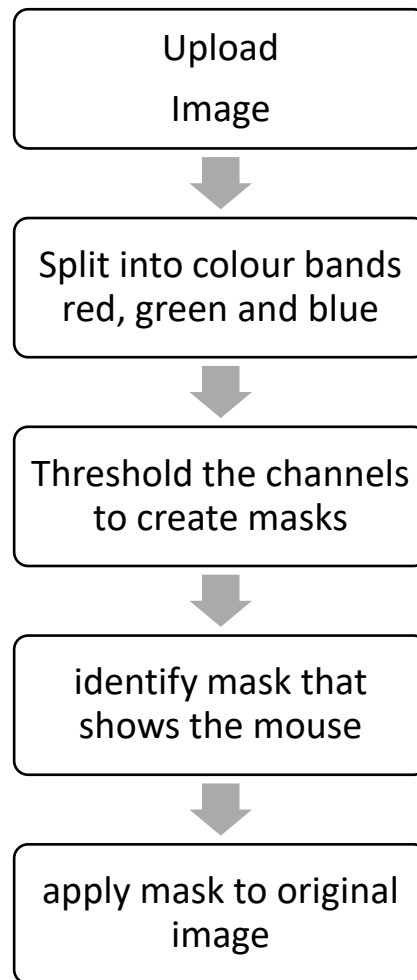
**Figure 2.11:** Flowchart demonstrating the sequence of image processing methodologies applied to the original image to isolate the mouse from the rest of the image and identify the tumour boundary.

#### 2.2.4.1 Colour Segmentation to Isolate the Mouse from the Image

In order to achieve efficient segmentation, two colour spaces were investigated, RGB and HSV. Where RGB represent colours in three layers (red, green and blue), HSV represent the colours in a conical form (hue, saturation and value). The first attempt at isolating the body of the mouse was using the RBG colour space. After the image was uploaded, it was separated into red, green and blue

channels. Each channel expresses colour with intensity between 0 and 255 where 0 is black and 255 is high intensity colour.

A threshold (a value between 0 and 255) was applied to each channel to restrict the pixels on show, in order to identify which pixel range was best for isolating the body of the mouse from the remainder of the image. The thresholds applied were 0, 50, 100, 150, 200 and 255. If the threshold on one channel was successful at displaying the mouse, the channel was combined with the remaining two channels that still portrayed their full pixel range. Together the three channels formed a binary image. A binary image is two-valued (0 and 1) black and white image, where the mouse is white due to the threshold applied and the background of the image is black, as the background pixels did not meet the threshold. The binary image was then applied to the original RGB image where only the white pixels, corresponding to the mouse, were visible.



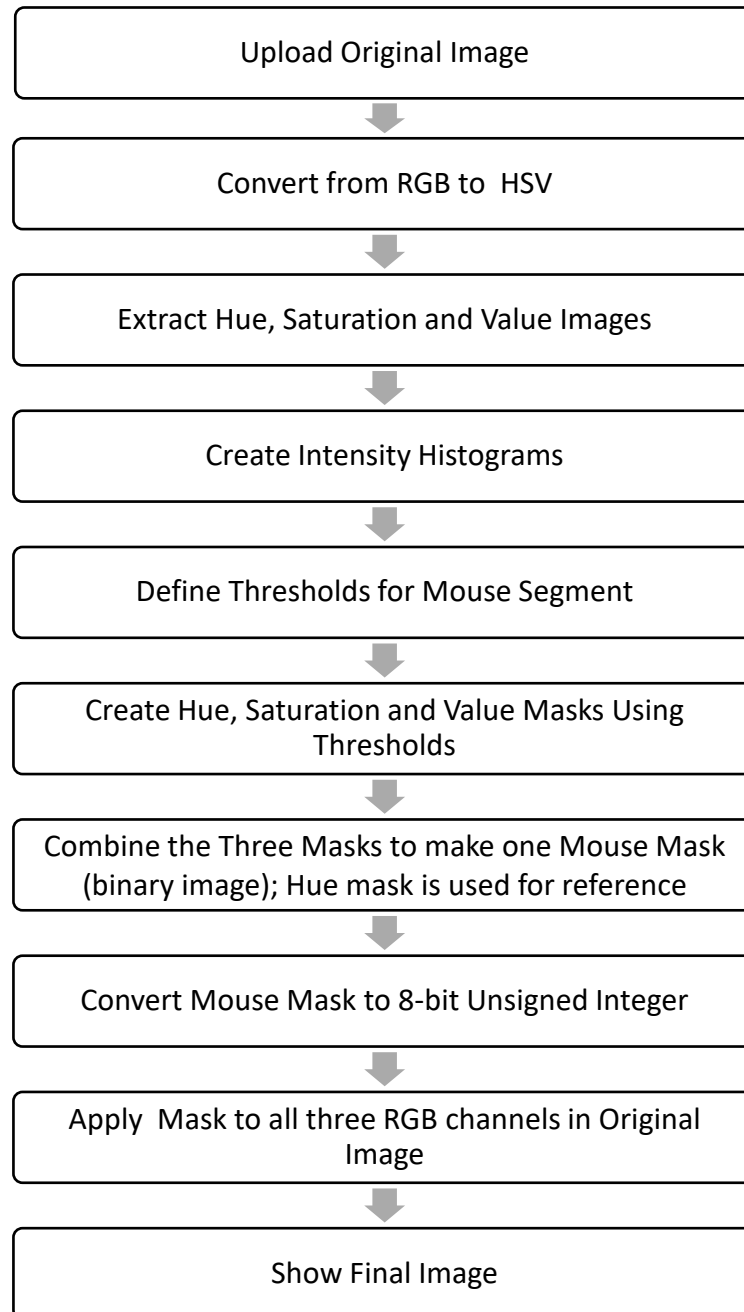
**Figure 2.12:** Algorithm for RGB colour segmentation

The steps taken to split a standard digital image in RGB format into the three colour bands to threshold them, restrict pixel range, in order to identify the pixels containing the mouse from the rest of the image.

The second colour segmentation technique applied was using the HSV colour space, Figure 2.13. The original image was uploaded into Matlab and converted from its original RGB form to HSV. The image was then split into hue, saturation and value channels and representative histograms were created, illustrating the pixel intensity in each channel. The original image has three distinct parts, the background, the hand and the mouse. As the mouse has the smallest surface

area in the image, it is expected that the mouse would have the smallest intensity peak in the histogram. The channel with the clearest visual representation of the mouse was then given a threshold dependant on the range of pixels of the smallest peak on the corresponding histogram, in order to isolate the body of the mouse from the rest of the image.

If the threshold on one channel was successful at displaying the mouse, the channel was combined with the remaining two channels that still portrayed their full pixel range. Together the three channels formed a binary image. The binary image was then applied to the original RGB image, hence only pixels that were white on the binary image were on display thus displaying the mouse. If successful in isolating the mouse body from the remainder of the image, the technique was applied to a batch of xenograft images. This was done by using the algorithm in a for-loop. A for-loop iterates through the batch, applying the algorithm to each image.



**Figure 2.13:** Algorithm for HSV colour segmentation

Splitting an image by hue saturation and value to identify if the mouse pixels can be differentiated from the parts of the image that do not contain any mouse.

#### 2.2.4.2 Morphology Operations to Trim the Mouse Body and Centre of the Tumour

Morphological operations were used to centre in on the tumour by removing outer sections of the mouse body. Erosion, a morphological operation, trims boundaries on greyscale and binary images using the function '*imerode*'. This technique was utilised to erode outlying features of the mouse body including paws and the head thus reducing information in the resultant image to aid in the isolation of the tumour. A structuring element was defined, a square with a width of 200 pixels, and used to remove pixels around the image boundaries. The image input must be in binary form thus the image should only display black and white pixels. To increase efficiency, the binary mouse mask created in the previous section when segmenting the image in the HSV colour space was eroded before it was applied to the original RGB image.

#### 2.2.4.3 Edge Detection

To isolate the tumour from the body of the mouse, edge detection methods were evaluated. These detect differences in regions of pixel intensity, and could potentially determine the outline of the tumour. Images were uploaded into Matlab and converted from RGB into greyscale as edge detection only works on black and white or grey images. Next, the edge detector function was selected, and the method for edge detection specified. The default detector in Matlab is Sobel. Other edge detectors tested were Canny, Canny approximation, Laplacian of Gaussian, Prewitt, Roberts and Zero-crossings to ensure the optimum detection



method was used to find the tumour boundary. These are covered in detail in section 1.4.1.2.

To refine the process of edge detection, a threshold was defined for the image, thus ensuring the detector then stays within the limits set by the threshold. To do this, before running the final stage of edge detection, the threshold function was executed; MATLAB automatically calculated the threshold of the image according to the method of edge detection specified. The threshold value was then set before the edge detector function was called ensuring the edge detector stayed within the limit of the thresholds. The process of edge detection was further refined through noise reduction.

#### **2.2.4.4 Noise Reduction**

Often, the quality of images is reduced due to the internal workings of a camera. This can cause blurring of regions thereby corrupting the image and making it difficult to detect an edge. Noise can be reduced in the image before the edge detector function is called using an adaptation of the Wiener method, through analysing the neighbouring pixels around each pixel. Once the image was loaded and converted to greyscale, the '*wiener2*' function was called from MATLAB. This returned an image with noise reduction and edge detection was applied to this.

#### **2.2.4.5 Contours**

Another image processing technique that could potentially aid the isolation of the tumour from the body is through creating a contour plot of the original image.

Contours create paths in an image, grouping regions with consistent image intensity values. To find the contours, the image was uploaded into MATLAB and converted into greyscale. Finally, the '*imcontour*' function was executed and the image was then visualised.

#### **2.2.4.6 Gradients**

Gradients are useful for visualising changes in pixel intensities in an image and therefore could detect the tumour from the mouse body as the tumour has increased depth in comparison to the mouse body where depth refers to the height of the tumour. This was achieved through finding the gradient magnitude and direction, the change in pixel intensity, and the direction of pixel intensity, respectively. The image was uploaded and converted to greyscale. The gradient operator was defined and the gradient function '*imgradient*' was executed. The resulting images were visualised using '*imshow*'. This gave an output of two images, one for gradient magnitude and the second for gradient direction.

#### **2.2.4.7 Threshold Image**

During colour segmentation thresholds were applied to image channels to focus on pixels corresponding to the mouse body. In this section, thresholds are found from and applied directly to the original image with the aim of finding a change in pixel intensity around the tumour boundary which would help isolate it from the body of the mouse. To threshold the image, the *imtool* function, a function in Matlab which allows the user to manually scan the image and obtain its pixel value, was used to locate the pixel value around the boundary of the tumour. The

threshold was then applied to the image by ensuring only pixels within 10 pixel values of the tumour boundary were visible. This generated a binary black and white image.

Image processing methods described previously, including edge detection and gradients, were applied to the resultant binary image. Furthermore, an attempt was made to locate the boundaries, a technique tracing exterior boundaries of regions in an image, and centroids, the central point of the region, to ultimately locate the tumour on the mouse body. To calculate the image boundaries, the '*bwboundaires*' function in MATLAB was executed. To calculate centroids in the image, the function '*regionprops*' was executed and centroids were specified. The centroids were then plotted onto the image itself and visualised.

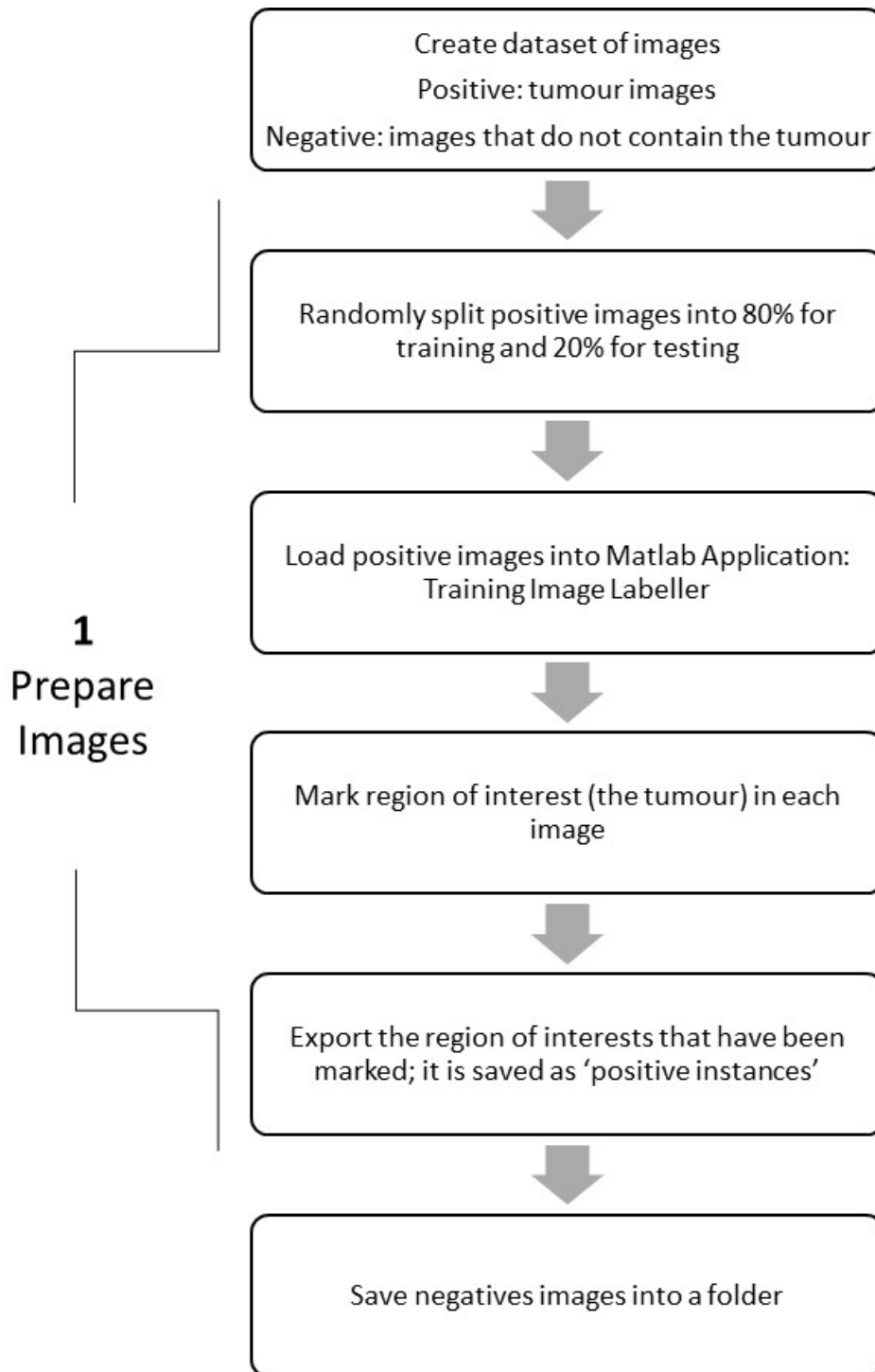
This following section describes machine learning techniques used to detect and isolate the tumour from the rest of the image. Previously, image processing was applied to a single image of a mouse in order to search for a specific feature, such as edges or contours. This section involves training the computer to detect tumours through inputting thousands of image. The machine learning methods presented are cascade classification and convolutional neural networks (CNN).

#### **2.2.4.8 Cascade Classification**

Cascade classification is a method of object detection that uses multiple cascades, or stages, to train classifiers, known as weak learners, with positive and negative images. The classifiers are boosted through learning with every cascade thus transforming the classifiers into strong learners with increased

predictive power. The two-step process used to perform cascade classification for tumour detection starts with preparing the images required, Figure 2.14. To facilitate the learning of the classifier, positive images, i.e. images displaying the tumour, and negative images, i.e. images not displaying the tumour, were used.

For cascade classification to work successfully, the classifiers require processing of large numbers of positive and negative images. The number of images was increased through rotating and flipping the images captured previously and altering the contrast. The following Figure 2.15 shows how one positive image of a tumour, highlighted with a red box, was transformed to give 36 positive images. Figure 2.16 shows some examples of negative images. Once the dataset was complete, 80% of both positive and negative images were used as training data and 20% for testing, to ensure the detectors were working successfully.



**Figure 2.14:** Pre-processing cascade classification – Prepare images

The first step in using cascade classification to detect tumours is to prepare the images for the process. This includes positive images containing the tumour and negative images not containing the tumour, to teach the classifier what it should and should not be detecting.



**Figure 2.15:** Sample of positive images

Positive images contain the tumour. To increase the data set, original subcutaneous xenograft images were transformed. Each row in the image grid represents a different contrast level applied to the original image marked with a red box. Each column represents a rotation of the image.



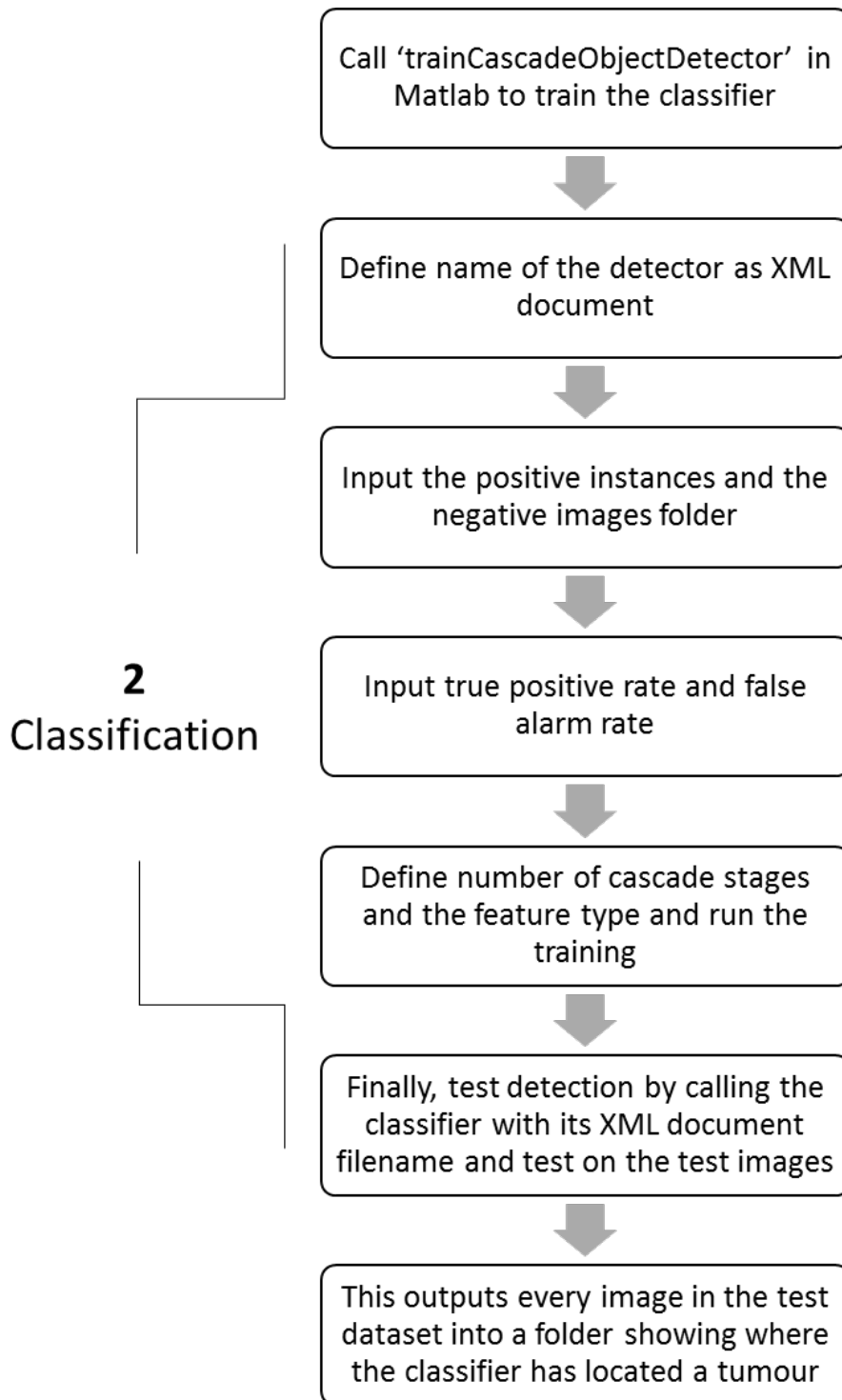
**Figure 2.16:** Sample of negative images

Negative images were used to teach the classifiers what not to detect. The sample of images contains the background of the image, including the blue glove, areas of the mouse body and the extremities of the mouse.

The positive images designated for training were loaded into the MATLAB Application '*Training Image Labeller*,' where the region of interest was marked. The region of interest is the area in the image which the detector must find thus the tumour was marked. The regions of interest for all the images were exported as '*positive instances*.' The negative images were simply saved in an alternative folder. The next step was to train and test the classifier, Figure 2.17. The classifier is a SVM, a supervised learning model used for categorizing data which in this case is a model designed to detect tumours.

The classifier was trained using an inbuilt MATLAB function, '*trainCascadeObjectDetector*'. The function required a number of parameters to be defined. The name of the classifier was defined and saved in a separate XML document to enable the classifier to be executed in later stages. The location of the '*positive instances*' file and negative images folders were added. Next, the '*true positive rate*' was defined; this is a number between 0 and 1 representing the fraction of positive images correctly classified. A higher true positive rate

means increased accuracy in the classifier identifying tumours but the process is more complex thus taking longer to execute.



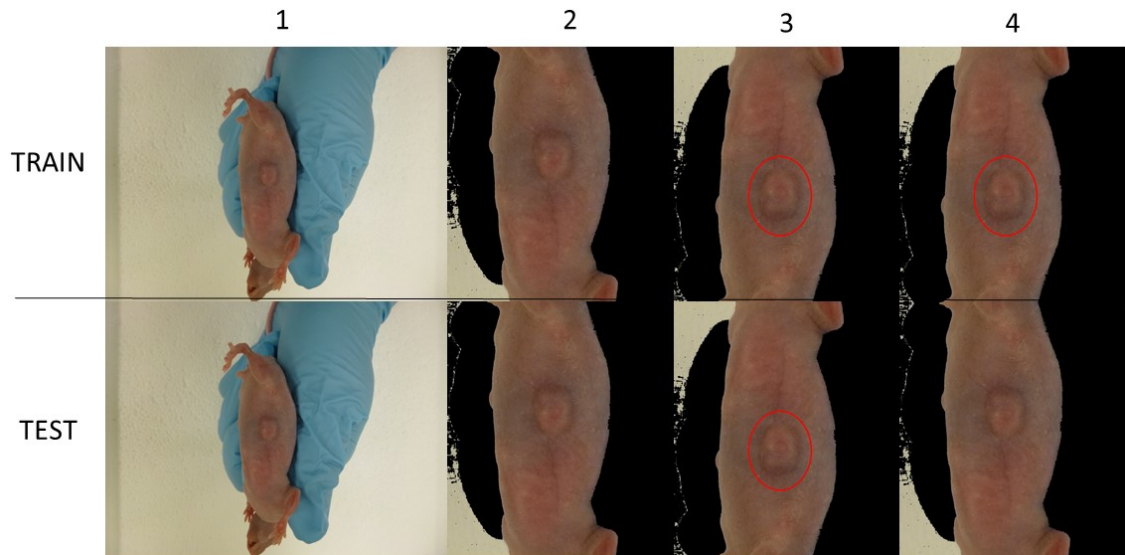
**Figure 2.17:** Cascade classification - classification

The second step in cascade classification, after preparing the images, is to complete the classification process. The classifier is trained with the dataset and then tested to find its accuracy.



The next parameter defined was the *'false alarm rate'*, a number between 0 and 1 representing the number of negative images incorrectly identified as tumours. A lower false alarm rate means increased accuracy but the process of detection requires more time and computational power as the classifier would keep training until it reached the required rate. This can be difficult to achieve with a small sample size. This was followed by defining the number of cascade stages as 5 due to the low sample size and low false alarm rate for the dataset.

The classifier searches for specific features within the images during its learning process. The three feature types are Haar-type features, histogram of oriented gradient (HoG), and local binary patterns (LBP). All three were tested and the best performing was used in further testing. With the required parameters defined, the *'trainCascadeObjectDetector'* function was implemented and the classifier was trained. To test the effectiveness of the classifier, it was applied to the test images. The classifier was executed using the XML document created in the previous step. It identified tumours through placing a yellow box around the region it believed to be a tumour and saving the resulting image in a folder.



**Figure 2.18:** Sample of images used for cascade classification

Cascade classification was run multiple times to find the optimum dataset for stronger classification. The columns show the four versions of cascade training with the rows showing the images used for training and testing the classifier. The red circles were added to the images to aid the classifier in detecting the tumour.

To ensure the classifier was trained effectively for detecting tumours in digital images, 4 versions of cascade classification were tested, each with different parameters and datasets of images. Figure 2.18 specifies the parameters used for each version. As a potential advancement on the learning parameters evaluated in this section, next, a more sophisticated form of machine learning, CNNs, were used to detect the tumour in the image. To increase the accuracy of tumour identification a more sophisticated feature learning technique was deployed (i.e. neural networks).

#### 2.2.4.9 Convolutional Neural Networks

The CNNs used in this work are deep neural networks that are pre-trained on huge ImageNet dataset. The idea here is to use the filters of a pretrained (learned) neural network to extract image features. Hence rather than using hand

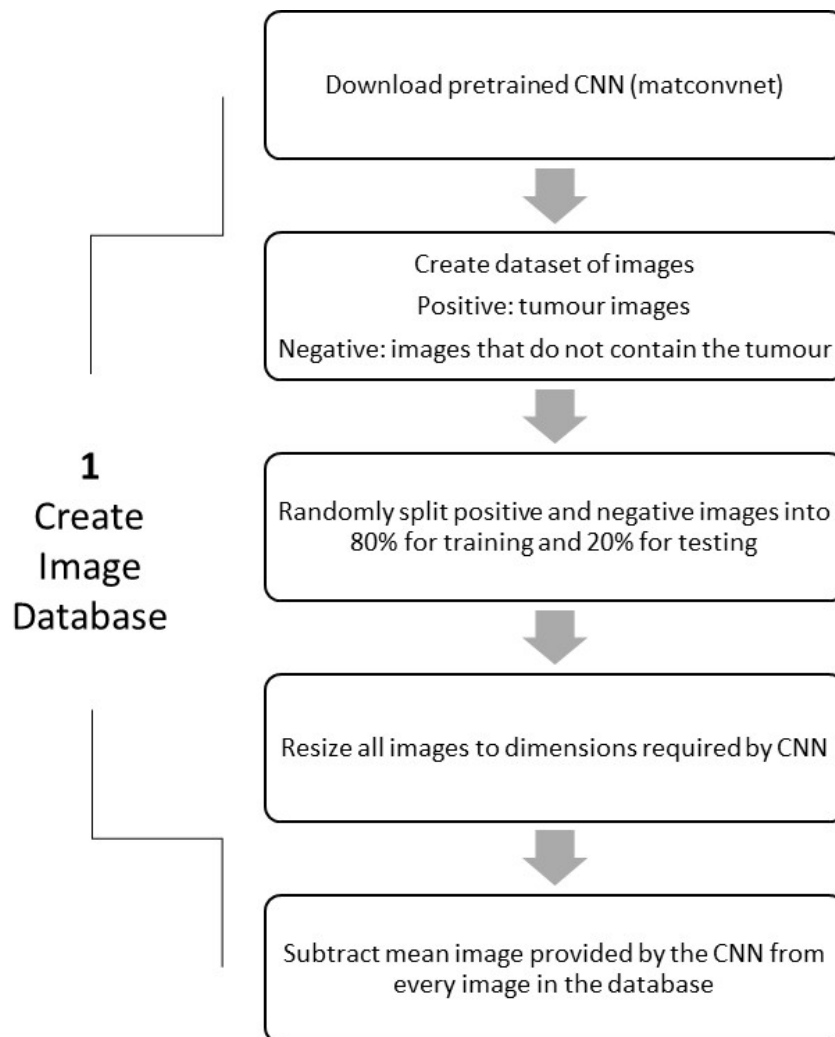
crafted classical features such as Haar, LBP and HOG we use a weights of the neural network to automatically interesting features from images. Actually, recent research has shown that off-the-shelf (pre-trained) neural networks are better image descriptors as compared to classical algorithms[130]. The CNNs used in this project to extract features that are used to train SVM classifiers to detect tumours were VGG-16, VGG-19 and VGG-face, created by the Oxford Visual Geometry Group (VGG)[126]. These were selected for use for this study, due to their accuracy as well as ease of application. See section 1.4.3.2 for details on these CNNs.

The CNNs have multiple layers, and each layer extracts different features from the images. To ensure the layer most suited to the dataset was used, the outputs of multiple layers were tested including the two fully connected layers (FC7 and FC6), and 2 last convolution layers. The layers are described in further detail in section 1.4.3. As with cascade classification, both positive and negative images were required but it was imperative that the number of positive and negative images were the same for both training and testing as this is a requirement for using CNNs. The negative images for training and testing were identical to the positive image numbers. The datasets were limited by the computational power available. The algorithm to use a CNN for tumour detection is presented in 3 stages; 'create image database', 'extract features' and 'classification'.

#### Create image database

Stage one is to create an image database. The algorithm can be seen in the following Figure 2.19. For CNNs to work successfully they require large numbers

of images. In order to do this, the number images was increased through rotating and flipping the images and altering the contrast. Both xenograft and pseudo-tumour images were used. Positive images and negative images were prepared as previously described.



**Figure 2.19:** Pre-processing CNN – create image database

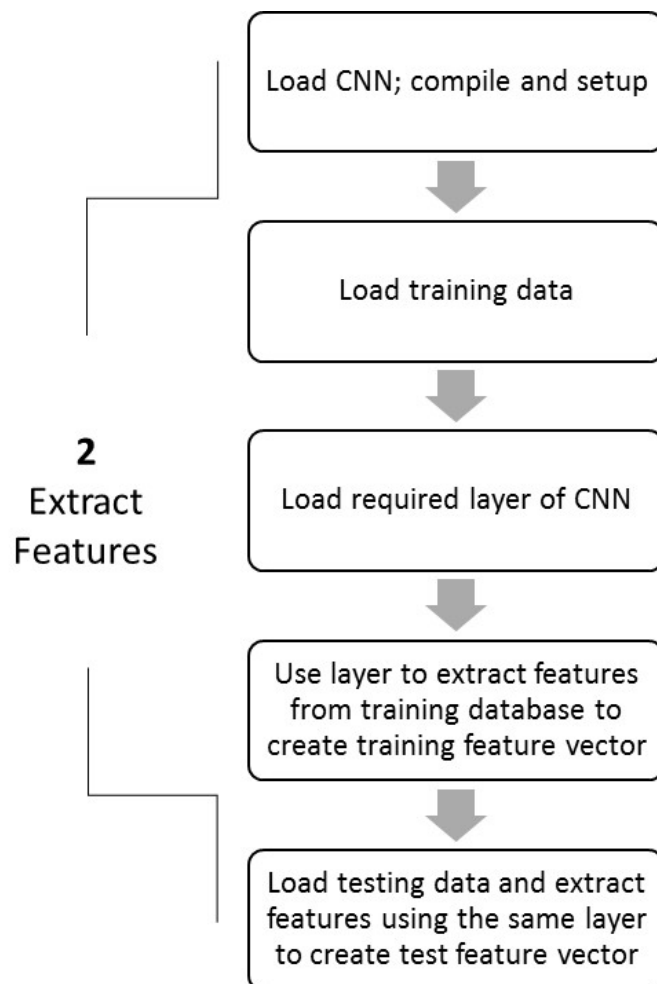
The first stage of the machine learning process was to create the image database, ensuring both positive and negative images were included and optimised to the right image size as required by the CNN.

With both datasets prepared, 80% of both positive and negative images were kept as training data and 20% were set aside for testing to ensure the detectors

were working successfully. Every image in the database was then resized to the dimensions required by CNN (see Table 3.2 for the image size required for each CNN). The final step in creating the image database was to subtract the mean image, provided by the CNN.

### Extract features

With the images prepared, the next step was to extract features. The algorithm for this is shown in Figure 2.20.



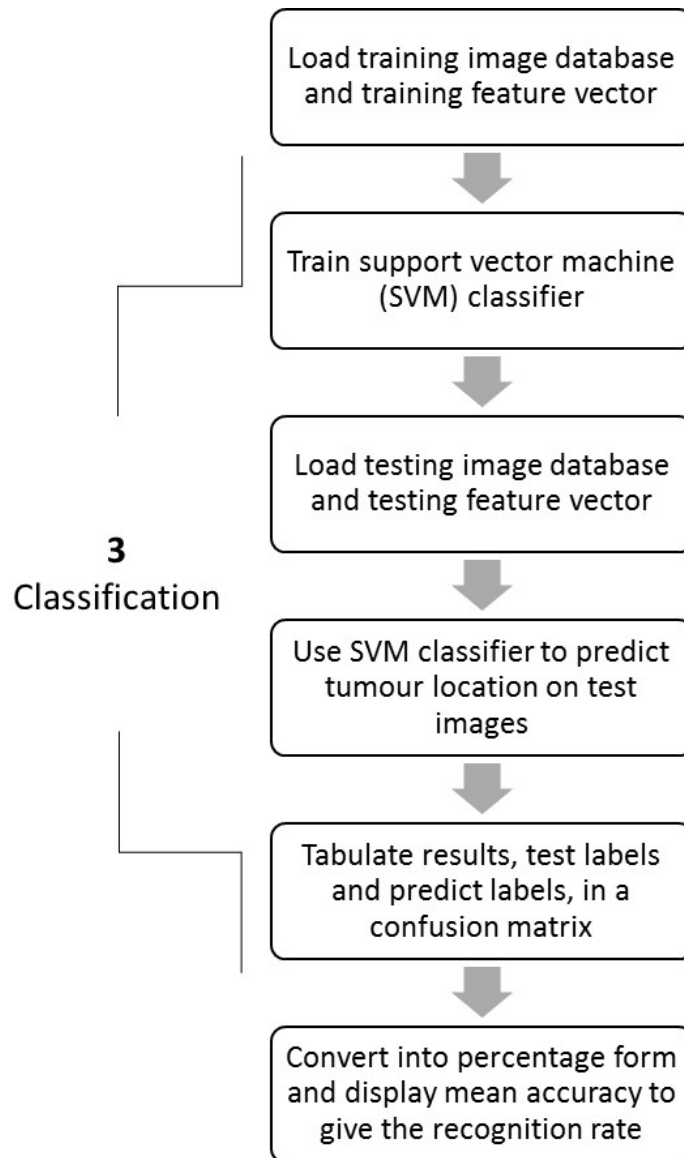
**Figure 2.20:** CNN – extract features

The second step in the machine learning process, after creating the image database, is the use a layer of the CNN to extract features form the images and save these in vectors.

After initially loading the CNN, compiling and setting it up, the compile and setup arrangements were saved within the CNN, and MATLAB was used to call them. The next step was to load both the training data and the required layer from the CNN which was then used to extract the image features and create a training feature vector. The testing data was then loaded and using the same layer, the features were extracted and saved as a test feature vector. Specifically, the framework used throughout our experiments is MatconvNet a deep learning framework developed by VGG that easily integrates with MATLAB.

### Classification

Now the features have been extracted, the next stage is to train a classifier and thus implement 'machine learning' to allow it to recognise tumours from non-tumours through studying the positive and negative images contained in the training feature vector. The algorithm for this is shown in the following Figure 2.21.



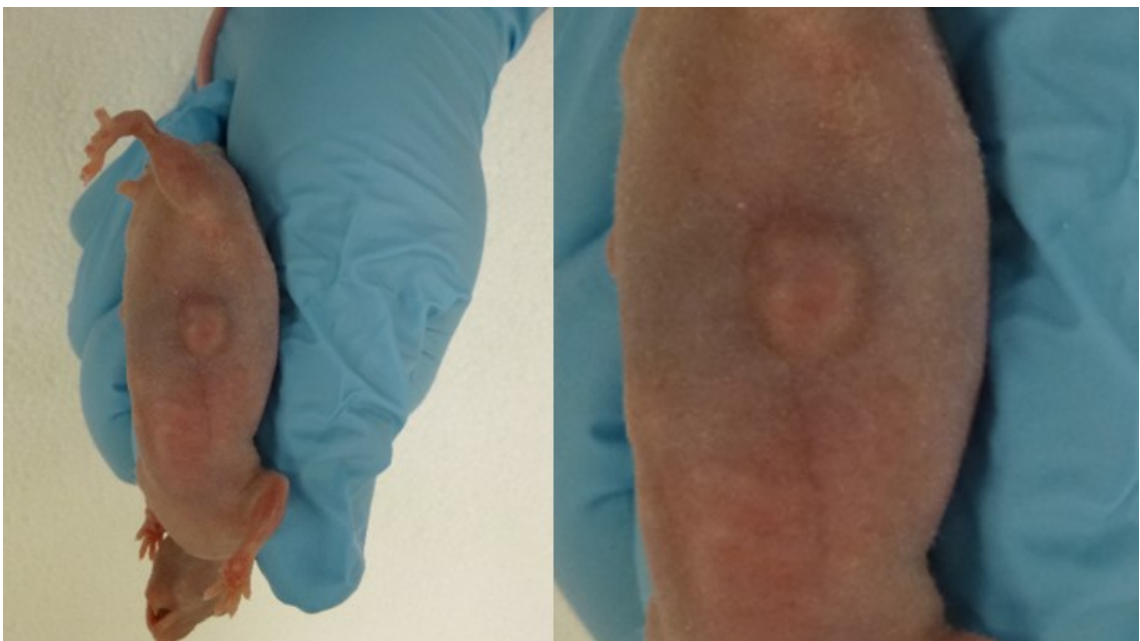
**Figure 2.21:** CNN - classification

The third and final step in the machine learning process after creating the image database and extracting features from the images was to complete the classification. The classifier is trained and then predicts location of the tumours on test images.

The SVM classifier was trained using the training feature vector. The next step was to test if the classifier has been trained successfully. To do this, the classifier was used to predict the location of tumours in the testing image database. Once the results were collected, the mean accuracy of the classifier in correctly identifying the tumours was calculated to give the recognition rate. To visualise

the location on the image that the detector identified as a tumour, a final step was added to the computer algorithm to put a red box around the detected area.

If the detection was unsuccessful, a manual step was introduced to ensure 100% tumour recognition. In this case, the image was cropped to remove outlying features and focus on the tumour itself, a representation of this is shown in the following Figure 2.22.



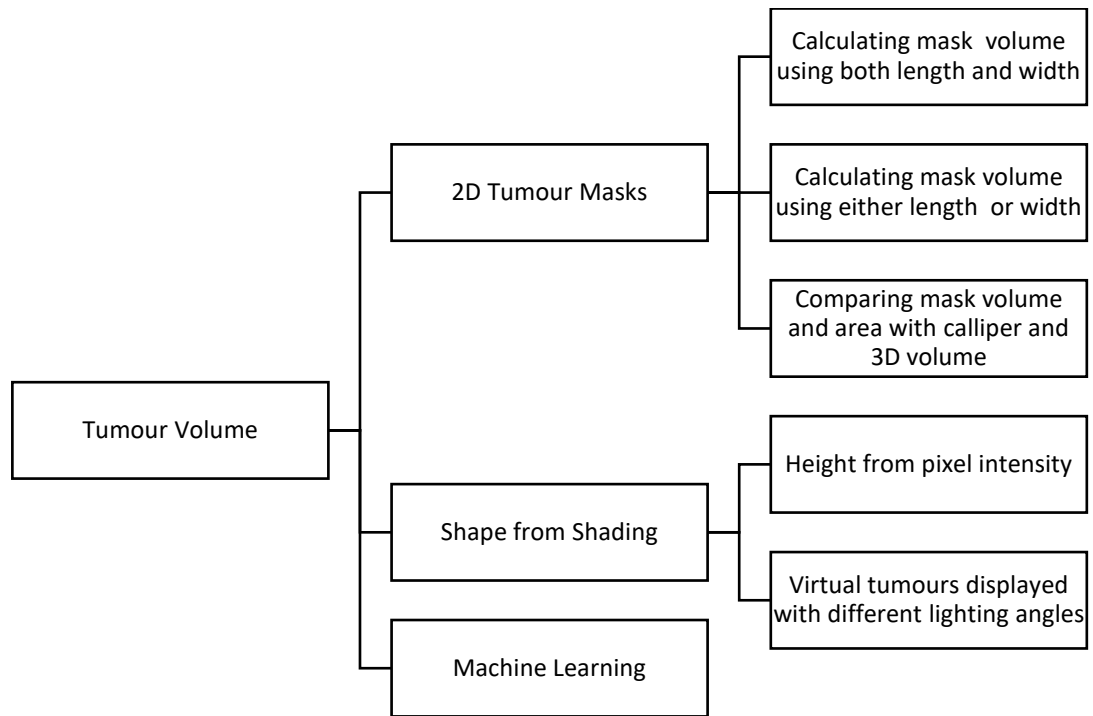
**Figure 2.22:** Original and cropped images used for detection by CNN

The image on the left hand side shows the original subcutaneous xenograft image used to test the classifier. The second image on the right hand side is a cropped version of the original image used if the classifier failed to correctly identify the original image.

Multiple techniques from image processing to machine learning have been applied to detect the tumour in the image automatically. The next stage of the project is to use the images to find a tumour volume that is comparable to, or better than, calliper measurements. The following flowchart, Figure 2.23, details the different methodologies used starting with 2D tumour masks, where volume



was calculated using length, width and mask area. Next, shape from shading techniques were applied where a relationship between light intensity and volume was investigated. Finally, machine learning techniques using pre-trained CNNs were used to find tumour volume.



**Figure 2.23:** A flowchart showing the different methodologies used to decipher tumour volume from s.c. xenograft images.

### 2.2.5 Tumour Volume – 2D Tumour Masks

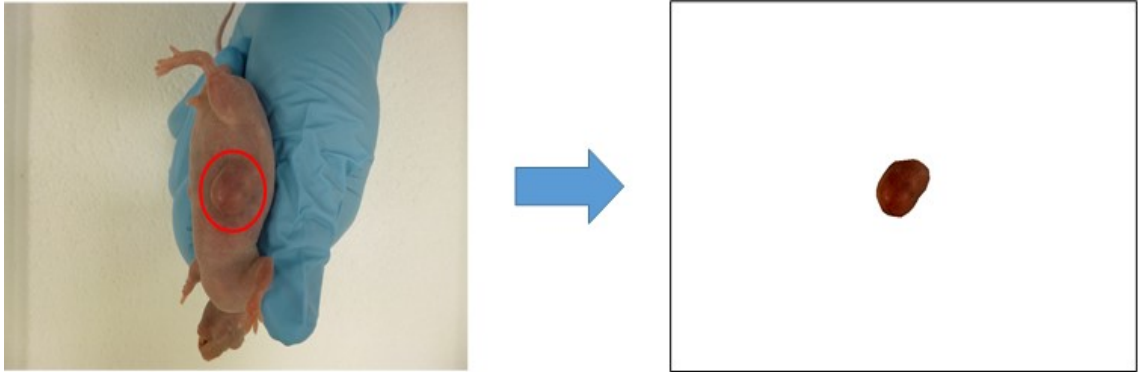
This section describes the methods used in calculating volume from the digital images by creating masks of the tumours and applying the modified ellipsoid formula, allowing for a direct comparison to calliper measurements. This was completed for both images of subcutaneous xenograft tumours and of pseudo-tumours. Furthermore, the pseudo-tumours were scanned in 3D thus giving a 3D

volume. This allowed not only evaluation of the accuracy of volumes generated from the digital image masks, but also of the calliper measurements themselves.

### 2.2.5.1 Digital Masks

Digital masks of the tumours are images that contain only the tumour with all background removed. These were created to isolate the pseudo-tumours and xenograft tumours from their surroundings in the images. This allowed for clear length and width measurements so the volume of the tumour could be calculated using the modified ellipsoid formula, thus making the measurement comparable to volume obtained from calliper measurements.

Masks were created using ImageJ. The ImageJ software was opened, 'File' and subsequently 'Open' were selected from the menu bar, and the image was chosen from the laptop's file directory. The tumour was zoomed onto and in some cases the image contrast was enhanced by 0.1% ('Process' in menu toolbar and then 'enhance contrast') to make the perimeter of the tumour clearer. The polygon tool was used to apply multiple points around the perimeter of the tumour, ensuring that the points corresponded with the tumour shape. When the perimeter was complete, 'Analyse' and 'Measure' on the menu bar were accessed to get the pixel number for the area selected. The mask was then created by selecting 'Edit' and 'Clear outside'. This means only the tumour selected was visible and the remainder of the image turned white. The mask was then uploaded for analysis in MATLAB. The same process was repeated for images of all tumours and models. The original image and its mask can be seen in Figure 2.24.

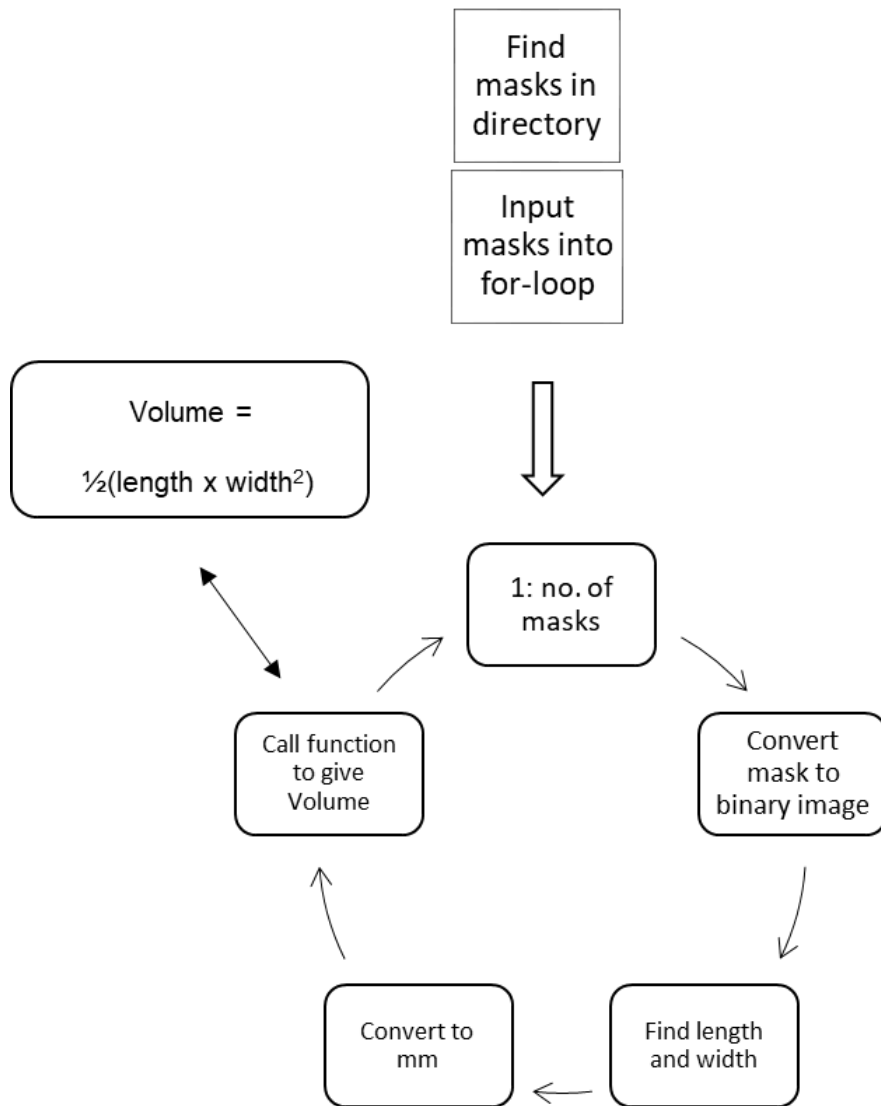


**Figure 2.24:** Digital mask of xenograft tumour

Original image of tumour xenograft is shown on the left, the mask generated is shown on the right. A red circle is placed around the tumour area in the original image.

### 2.2.5.2 Calculating Mask Volume using both length & width

Figure 2.25 highlights the steps taken in MATLAB to find the volume using the modified ellipsoid formula from the tumour masks. The masks were located in the file directory where they were stored and they were converted to binary black and white images, allowing for image processing. The maximum length and width of the mask were calculated through the function '*regionprops*' in MATLAB which fitted an ellipsoid around the mask shape and then found the major axis length, and the minor axis length perpendicular to this. Once the length and width were found, they were converted from pixels to mm. These measurements were inputted into the modified ellipsoid formula to find volume.



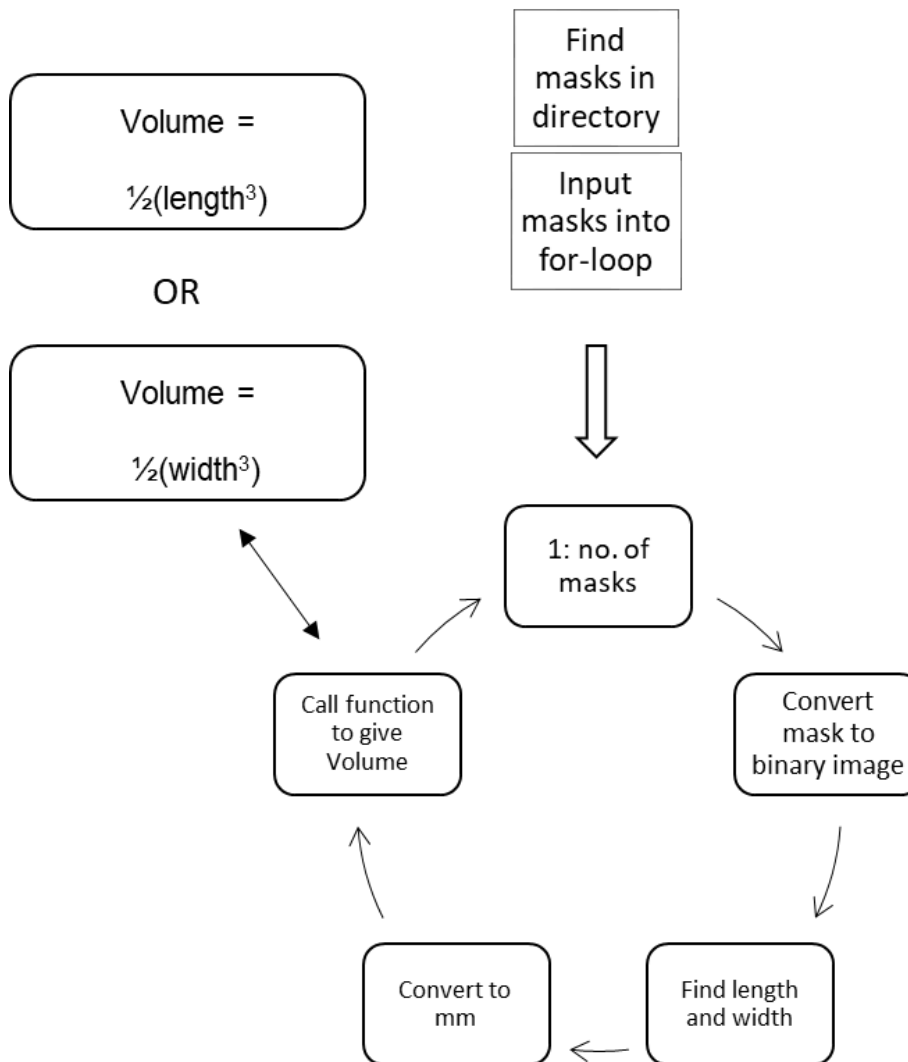
**Figure 2.25:** Mask volume using the modified ellipsoid formula

This figure shows the process of calculating the volume from the masks using the modified ellipsoid formula. The loop iterates through all the masks applying the volume onto each image.

### 2.2.5.3 Calculating Mask Volume Using Either Length or Width

The modified ellipsoid formula uses both length and width measurements to generate a volume. Volumes were generated using only the length measurement and using only the width measurement to test if both length and width parameters are essential for calculation. The algorithm, shown in Figure 2.26 followed the

same steps as the previous algorithm, Figure 2.25, to find the mask volume except that the function which calculates the volume is changed depending on which measurement is being used. Length and width were calculated using the same method, 'regionprops' in MATLAB.

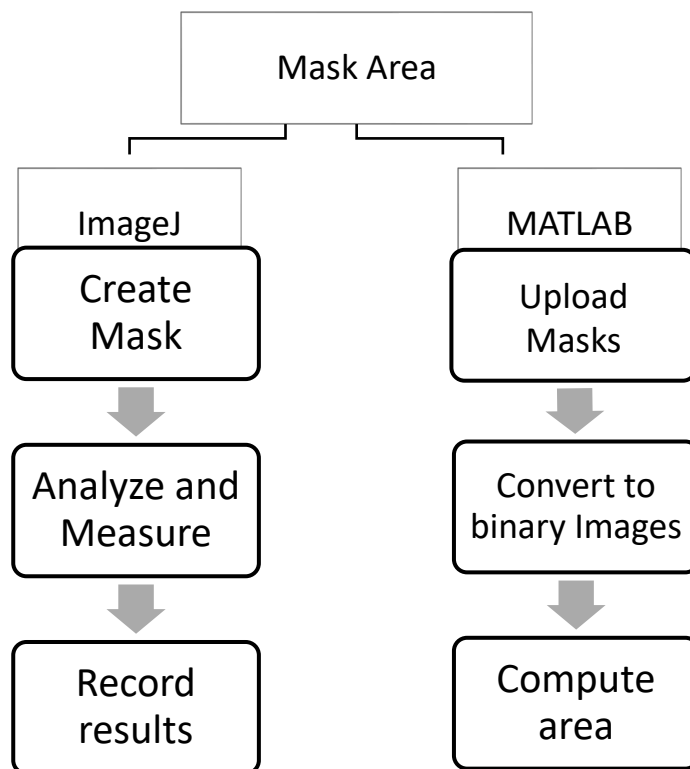


**Figure 2.26:** Algorithm: Mask volume using only length and width in modified ellipsoid formula

The process in finding the volume of the digital masks using either the length or width measurement is presented. The algorithm iterates through the masks contained in the directory to find the volume of each image.

#### 2.2.5.4 Finding Mask Area

When the masks were created, the area in terms of the number of pixels of the mask created were recorded in Excel. The area was also checked using both ImageJ and MATLAB and both returned identical values; the algorithm is shown in Figure 2.27.



**Figure 2.27:** Algorithm- mask area

Algorithm illustrated the two methods used to find the mask area in pixels, using ImageJ and MATLAB.

#### 2.2.5.5 Comparing Mask Volume and Area with Calliper and 3D volume

Once the volume measurements for the 2D tumour masks obtained from digital images had been calculated, these were then compared to calliper volume and volume obtained from 3D scanning. This was done to evaluate the ability of the

algorithm to compete with the existing gold standard, callipers. Mask area was also used to find a relationship between the area of the mask and the volume. This was achieved through using statistics.

The paired t-test was used to test if the two variables in question had means significantly different from each other. This test was chosen as all of the variables arise from the same group of tumours and so can be paired with each other. The paired t-test works by accepting or rejecting a null hypothesis at difference levels of significance. The null hypothesis states that there is no difference in the means of the two variables and so there is no significant difference between the data groups.

The test was run in MATLAB using the following command:

$$[h,p] = ttest (dataset1, dataset2);$$

The outputs of the test are h and p where h represents the hypothesis test results. When h=1, the null hypothesis is rejected and if h=0, the null hypothesis is accepted. Additionally, p provides the P value, ranging between 0 and 1.

Furthermore, the correlation co-efficient was found to establish if the variables had a strong or weak linear relationship, this is also known as *Pearson's r*. The statistic outputs a number between -1 and 1 showing the strength and direction of the relationship. -1 represents a negative relationship whereas 1 represents a positive influence. The purpose is to check if one variable influences another, this

is known as covariance and *Pearson's r* is worked out by dividing the covariance of the samples by the product of their standard deviations.

This test was run in MATLAB using the following command:

$$[R, P] = \text{corrcoef}(\text{dataset1}, \text{dataset2})$$

R is the correlation coefficient and P tests the hypothesis of no correlation; a value less than 0.05 indicates a significant difference between the two variable data sets. An alternative method for deciphering tumour volume through using light intensity was also used in these studies. This is known as shape from shading.

### **2.2.6 Tumour Volume - Shape from Shading**

Principles of shape from shading were investigated to establish a relationship between shading and volume. In an attempt to establish a relationship between pixel intensity and depth, artificial models of different height were created and imaged from the same height with and without a fixed light source in order to detect any changes in the pixel intensity. Therefore, models with increased height were closer to the source of the light thus receiving increased light intensity. Next, a relationship between the spread of pixels and volume was investigated by creating virtual models with different angles of lighting, to see if any of the resulting features correlated with volume.



### 2.2.6.1 Depth from Pixel Intensity

As previously described, 8 artificial models were created at different depths. These were then imaged, from the same height, both with and without flash. The flash is a light on the camera itself and by using it, a light source was directed at the model. To find the maximum pixel intensity of the image, the image was uploaded into MATLAB and the 'max' value was found. To establish if there was a relationship between the maximum pixel value and the depth of the model, a correlation co-efficient was calculated. This was done using MATLAB by calling the correlation function, '*corrcoef*'.

### 2.2.6.2 Virtual Tumours displayed with different lighting angles

As previously described, virtual models were created with known volumes and different lighting angles, allowing for multiple features regarding pixel intensity to be extracted from images of the models and regressed with volume. If there is a positive relationship, the shading on an image can be used to predict volume. The volumes of the models were found with the software Autodesk Maya through using the divergence theorem, as previously described (2.2.2.2 Volume Determination).

As the 2D images of the virtual models were created in Maya, the background of every model was created black. The model was extracted from the image with MATLAB by only showing pixels that were not black. Features were then extracted by calling the required statistical function in MATLAB. The features extracted were pixel intensity in the form of histograms (distribution of data), and

also skewness (asymmetry in data distribution), kurtosis (measure of the peaks in data distribution) and histograms of gradients. Model perimeter and area in pixels were also extracted.

Once the required feature was extracted, it was regressed with volume using multiple regression analysis. Multiple regression is a statistical method where the relationship between two or more variables is analysed. Regression was used to train a vector, which acts as a tester or predictor, using both datasets between which we are trying to find a relationship. Y is the observed response (volume) and X is the predictor. One value was then removed from each dataset. It was used to test the predictor this approach is known as leave-one-out cross validation.

Features from the two datasets acted as predictors, these were then use to perform regression as stated above. Subsequently, the volume (response variable) of held-out data (i.e. data not used in the training) is computed using the coefficients of the trained regressor. The new volume was then compared to the original (ground truth) value, given by,

$$abs(\text{ground truth volume} - \text{predicted volume})$$

In order to estimate the prediction accuracy the absolute error computed using equation above is averaged over all the test samples – this is known as the mean absolute error.

$$\frac{1}{n} \sum abs(\text{ground truth volume} - \text{predicted volume})$$

Furthermore, to ensure the relationship between pixel intensity and various statistical methods could be fully evaluated, principal component analysis (PCA) was also used. PCA is a statistical technique used to remove correlation that exist with data, it also helps to reduce the dimension of the dataset. Firstly, the principal component was determined from the dataset. This type of analysis breaks down the data set into paired Eigenvectors and Eigenvalues, where the Eigenvector is the direction and the Eigenvalue represents the variance in the data in that particular direction. Therefore, the Eigenvector/s with the highest Eigenvalue/s is/are the principal component/s of the data set. This allowed the removal of any unnecessary variables within the dataset with insignificant Eigenvalues, thus providing a more focused output of the dominant dimensions within the dataset.

Where the methods described focus only on pixel intensity and the effects of an external light source, the volume was subsequently determined through using machine learning, where multiple features could be evaluated.

### **2.2.7 Tumour Volume – Machine Learning**

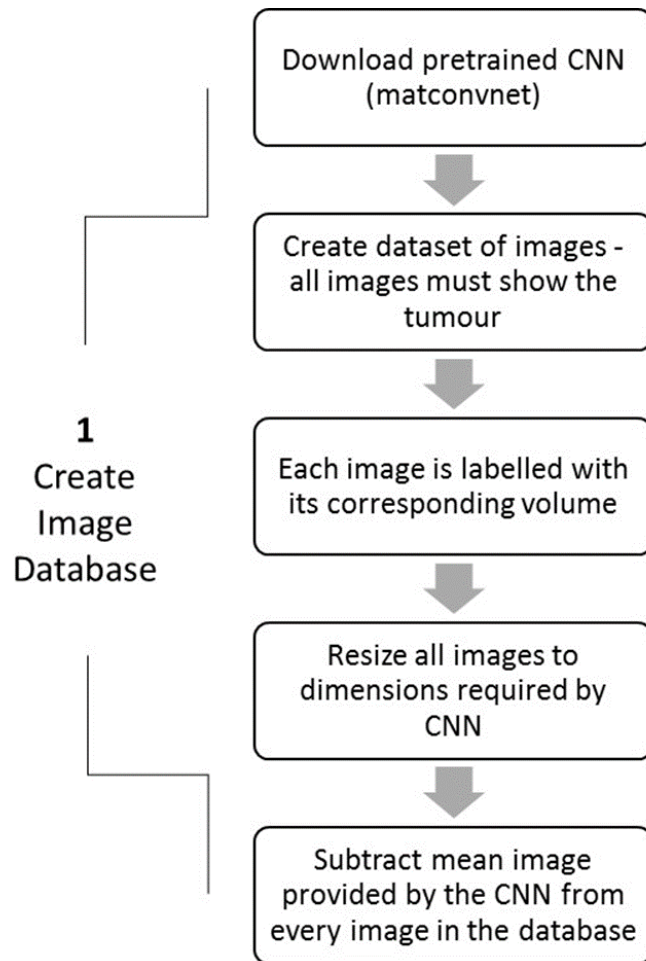
In an attempt to find volume from a 2D image, machine learning was used. CNN VGG-Face was selected, using layer 34 to extract features from the labelled tumour images. Machine learning algorithms (1.create image database, 2.extract features and 3.classification) were previously presented in section 2.2.5.2 Convolutional Neural Networks, and were used to detect tumours from images. The difference to the machine learning algorithms used here is that negative images were not required, and instead of classification, regression was used to

determine tumour volume. The three steps involved in the machine learning process were: create image database, extract features and regression as detailed below.

#### **2.2.7.1 Create image database**

When creating the image database, only positive images of tumours with known volume are required. Xenograft and pseudo-tumour images were used. To bulk up the dataset, the number of images was increased through rotating and flipping the images and altering the contrast, as shown in Figure 2.15.

Once the images were created, each image was labelled with its corresponding volume. To do this, an Excel document was created with the correct volume listed next to the name of the tumour. This information was then imported into MATLAB. Next, a loop was created so the script went through each individual image and found the volume value corresponding to the name of the image and thereby labelled the image with its volume. The images were then resized to fit the requirements of the CNN; the image size required for VGG-FACE is 244 x 244. As can be seen in Figure 2.28 the final step in preparing the image database was to subtract the mean image provided by the CNN from the images.



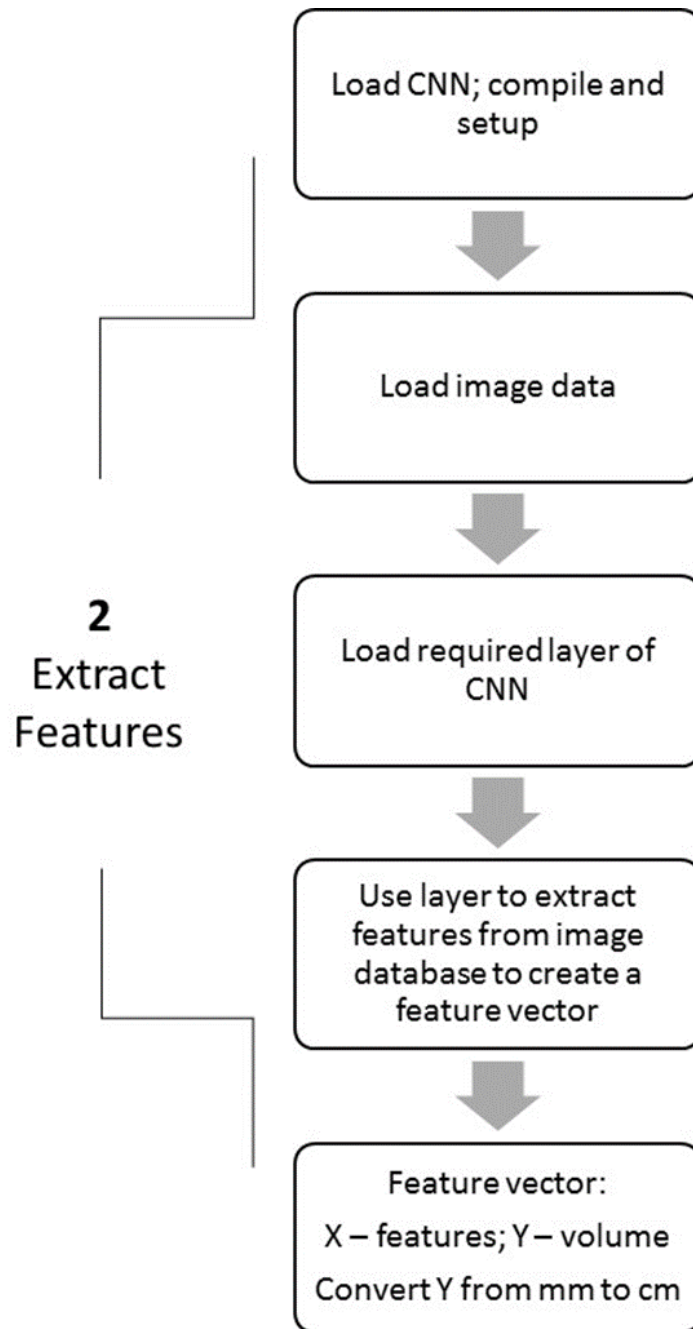
**Figure 2.28:** Algorithm 1 for machine learning to find volume showing how to create the image database

The first step in the machine learning process is to create the image database. Only positive images with tumours are required and must be labelled with their corresponding value to enable the classifier to find a connection between the two.

### 2.2.7.2 Extract features

Once the image database was complete, the next step was to extract features, as shown in Figure 2.29. The CNN and image database were loaded and the layer of the CNN to be used, 34, was specified. Layer 34 corresponds to the fully connected layer FC7 in VGG-16 and VGG-Face models. The features were then extracted using layer 34, resulting in a feature vector consisting of  $X$ , the features, and  $Y$ , the corresponding volumes. As a form of data preprocessing step, the  $Y$

covariable is rescaled by divided by 1000 to convert the volume from millimetres to centimetres.



**Figure 2.29:** Algorithm 2 to find volume with machine learning is to extract features to find volume

The second step in the machine learning process after creating the image database is to extract features from the images using the layer in the CNN.

### 2.2.7.3 Regression

Finally, partial least squares regression (PLS) was used to find a relationship between the features extracted from the images and their volumes. It was applied to the data in order to determine the linear relationship between the dependent variable (Y) and a set of predictor variables (Xs). PLS regression is a statistical technique which combines features from principal component analysis and multiple regression. In a nutshell it removes correlation between data samples and further performs regression.

In this model the prediction functions used are represented by factors extracted from the “ $Y'XX'Y$ ” matrix. From this matrix, the total number of prediction functions which can be extracted will typically exceed the maximum number of Y and X variables available, unlike other multivariate methods. PLS regression works by extracting the latent factors which explain the covariance between the independent and dependent variables as much as possible. Then the regression step predicts the values of dependent variables using the decomposition of the independent variables. This allows this type of regression model to be applied to datasets in which there are far fewer observations than predictor variables and can be used as an exploratory tool to select suitable predictor variables to identify any potential outliers before classical linear regression models are applied.

To evaluate the performance of the regression model i.e. its volume prediction accuracy, two cross validation techniques (leave-one-out, and k-folds) were employed. During cross validation, the computed mean absolute error (MAE) indicates the prediction accuracy, an MAE that is close to 0 indicates good

performance and a value close to infinity signifies poor prediction. Since the MAE computed during cross validation is parametric evaluation technique, a non-parametric evaluation method was also used in order to fully investigate the prediction accuracy. Here, the non-parametric testing tool used was kolmogorov smirnov's (KS) two sample test. KS test was performed to compare the predicted volume with the actual volume and this outputs a KS stat which describes the largest distance between the predicted and actual volumes. Furthermore the p-value returned by the KS test indicates whether the two samples were drawn from the same distribution or not. Ideally a good prediction accuracy is assumed when the two KS test signifies that the ground truth and predicted volumes come from the same distribution.

This section described the materials and methods used for tumour detection and tumour volume to facilitate achievement of the aims and objectives and answer the research question; can a non-invasive system replace the calliper? In the next section, Results, the findings from the experiments are presented.



### 3. Results

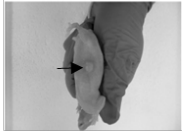
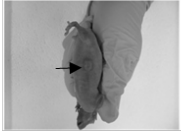
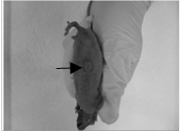

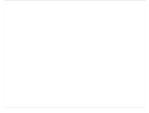
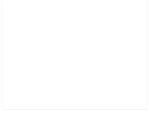









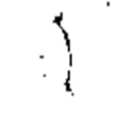


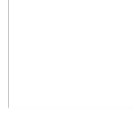


The results are sub-divided sequentially into 2 sections. In the first section, the results relating to detecting and isolating the tumour are presented, whilst in the second section the measurement of tumour volume is reported.

#### 3.1 Tumour Detection – Image Processing

In this section the results of applying imaging processing techniques to digital images of nude mice with s.c. tumours to detect the tumour are presented. The two step process comprises of isolating the region of the image containing the mouse, and then identifying the tumour. Techniques applied include colour segmentation, edge detection, noise reduction, contours, gradients and applying a threshold to the image, and these are presented in the sequence they were executed. The initial steps of this work were executed on a sample image, and once successful, were applied to the dataset of s.c. tumour images.

##### 3.1.1 Colour Segmentation

The first step towards detecting the tumour was to look to exploit any colour differentiation between the mouse and the remainder of the image. The image was captured in the RGB format and thus was segmented in the RGB colour space. The image was split into red, green and blue channels resulting in three binary images displaying pixels on a scale between 0 and 255, Figure 3.1.

Threshold Level	Channel Images		
	Red	Green	Blue
Original Image (no threshold)			
255			
200			
150			
100			
50			
0			

**Figure 3.1:** Original image split into red, green and blue channels, displayed at different pixel threshold levels to isolate the mouse from the remainder of the image.

The original red, green and blue channel images are shown in the first row, with a black arrow indicating the location of the tumour. Threshold level 255 and 200 show no image features, level 150 shows both mouse and hand. At level 100, whilst the blue image successfully isolates the mouse from the image, the red and green images fail to do so. Level 50 shows a weak mouse outline and level 0 shows no image features.

Multiple pixel thresholds were applied to the binary images to identify the pixels responsible for displaying the mouse. Higher threshold levels of 255 and 200 showed no image features. At level 150 the hand and mouse are visible as one feature, indiscernible from one another. Level 100 successfully displayed pixels responsible for the mouse. Decreasing the threshold level further was ineffective as level 50 weakly highlights the right hand side of the mouse body outline and level 0 displays no image features.

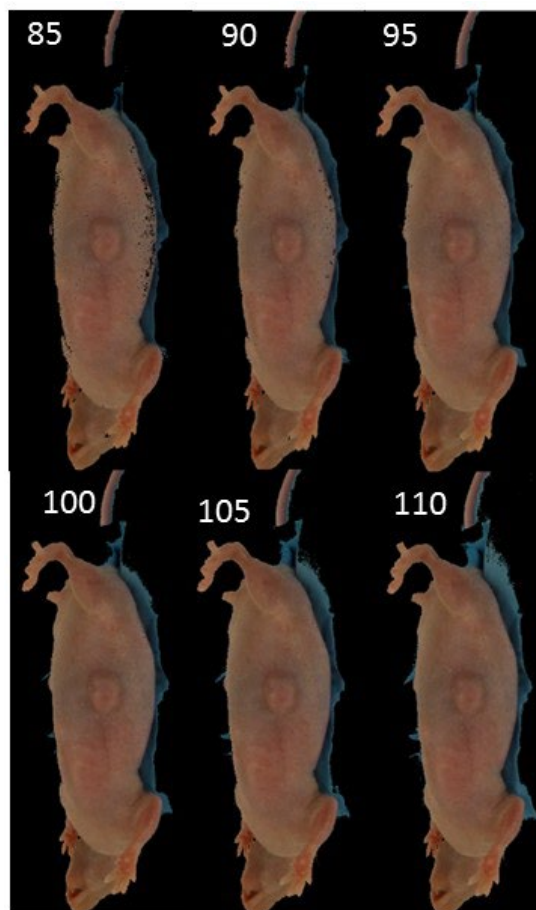
Whilst at level 100 the mouse outline is discernible, the red and green binary images are unsuitable as the body of the mouse is corrupted with contaminating pixels. Additionally, the red image would require further segmentation to remove the hand from the image. The blue binary image successfully displays the mouse in isolation and is void of corruption. Therefore, this image was applied to the original image as a mouse mask. Hence, only the pixels in the original image that correspond to the white pixels in the mask would be displayed thus ensuring that only the mouse is seen, Figure 3.2.

Although the resulting image successfully isolates the mouse from the hand and background of the image, blue remnants of the hand remain along the right hand side of the mouse body. To remove the blue hand whilst maintaining the integrity of the mouse body, the blue channel image was subjected to a tighter threshold selection between 85 and 110, Figure 3.3. Increasing the threshold resulted in a larger area of blue hand being visible. Decreasing the threshold reduced the visibility of the blue hand but compromised the mouse body as pixels along the right hand side of the mouse body were disrupted.



**Figure 3.2:** Applying the blue binary image mouse mask to the original image to isolate the mouse.

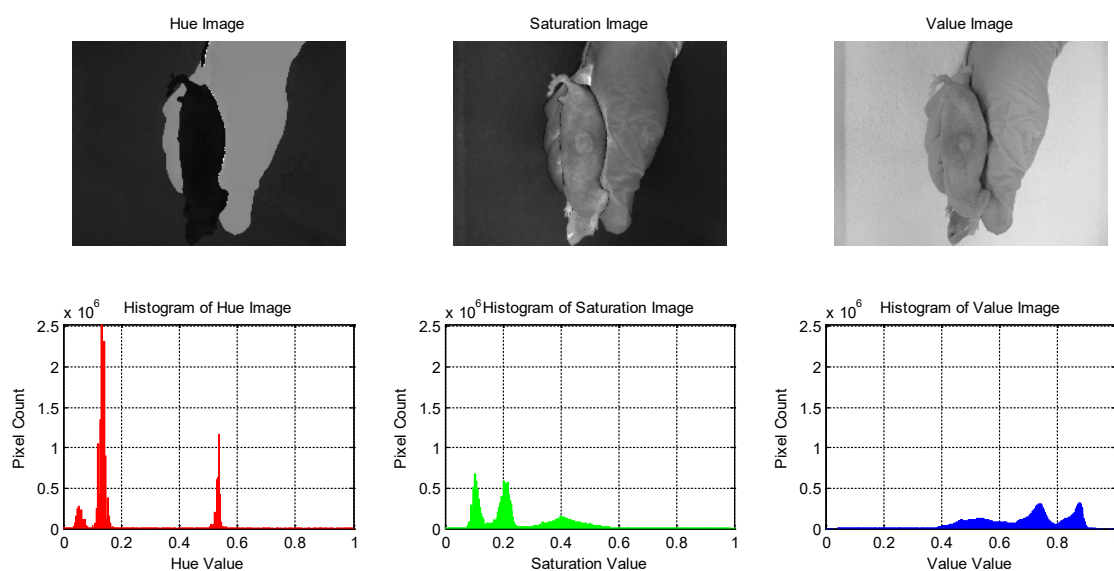
The mouse mask was applied to each layer of the original image, restricting the pixels on display. This resulted in an image where the mouse is clearly isolated but blue remnants of the hand are also visible down the right hand side of the mouse.



**Figure 3.3:** Blue mask thresholds between 85 and 110 applied to the original image to remove the blue glove from the image

Increasing threshold levels increased the amount of blue glove visible. Decreasing threshold levels reduces the amount of blue glove visible but introduced corruption in the mouse, specifically the right hand side of the mouse body.

The RGB model was unsuccessful in isolating the mouse body from the image therefore HSV colour segmentation was used. As the image was captured in RGB format, it was converted to the HSV colour space and split into hue, saturation and value channel binary images. Intensity histograms were created to visualise the distribution of pixels and pixel intensity, Figure 3.4.



**Figure 3.4:** HSV colour channels and corresponding pixel intensity histograms

Hue, saturation and value binary images are displayed with corresponding histograms. Saturation and value images offer no distinction between the hand and the mouse, with no defined peaks in the histograms. The hue image shows clear contrast between the mouse, hand and image background with three defined peaks in the histogram. The peak between 0 – 0.1 hue values represents the mouse and thus has been successful in identifying the mouse in the image.

The saturation and value images were unsuccessful in isolating the mouse from the image as there is minimal colour differentiation between the hand and the mouse. In the hue image, there is clear contrast between the mouse, the hand and the background. Furthermore, there are three clear peaks in the hue intensity histogram, representative of the mouse, hand and image background. The smallest peak, between 0 and 0.1, can be attributed to the mouse as it has the

smallest surface area. A threshold of pixels between 0 and 0.1 was applied to the hue image, resulting in a hue mask showing the mouse body. The hue mask was applied to each layer of the original image, Figure 3.5. This method was successful in isolating the mouse from the image.



**Figure 3.5:** The hue mouse mask was applied to the original image resulting in the mouse being isolated from the image.

The hue mouse mask was added to each layer of the original RGB image thus the only pixels visible on the final image corresponded to the white pixels from the hue mask. The resulting image shows the successful isolation of the mouse from the remainder of the image.

Following the successful isolation of the mouse from the sample image, this method was applied to every xenograft image collected from the study to ensure the results were reproducible. As Figure 3.6 shows, the method was inconsistent in isolating the mouse in its entirety and it often corrupted the tumour itself.

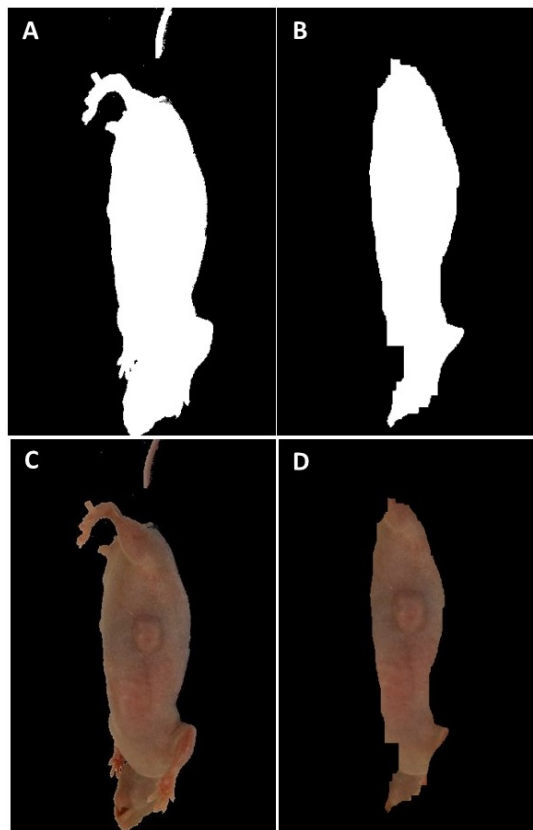


**Figure 3.6:** Sample of images with hue mask applied showing the inconsistency in isolating the mouse outline for each image

The hue mask, highlighting pixels corresponding to the mouse, was applied to the original xenograft images and this figure shows a sample of the results. The mouse body is corrupted in all the images including corruption of the tumour itself thus hsv segmentation was unsuccessful in isolating the mouse from the image.

#### 3.1.1.1 Morphological Operation - Erosion

As it is difficult to differentiate between the tumour and the surrounding mouse body, erosion was used to remove pixels along the outer boundary of the mouse where there was a significant difference in contrast than between the mouse body and the tumour thus increasing the focus on the tumour itself. Erosion, a method to remove pixels from object boundaries, was implemented on the hue mask generated in the previous section and the resultant image was applied to the original sample image of the mouse, Figure 3.7.



**Figure 3.7:** Effects of applying erosion on the hue mask and isolated mouse image

A is the modified hue mask with B showing the effect of applying erosion on A. C is the mouse mask created using A, and D is the effect of erosion on C. Erosion was successful in removing the outer boundary of the mouse body thus improving the process of isolating the tumour from the mouse body.

Evidently, erosion occurred on the outlying features of the mouse including the tail, limbs, head and the outer flank of the mouse body. The tumour and its surrounding area have not been compromised and instead, the tumour has increased in prominence in comparison to the mouse in its entirety. The eroded mouse image (D) was used in further testing, beginning with edge detection.



### 3.1.2 Edge Detection

Taking advantage of the reduced features in the eroded mouse image, Sobel edge detection was applied to highlight changes in pixel intensity at the interface between the tumour and the surrounding mouse body, Figure 3.8. The Sobel method defines an edge at the point where the image gradient is at maximum intensity. This method was unsuccessful in finding the tumour boundary and it was only able to detect the outline of the mouse body, Figure 3.8.



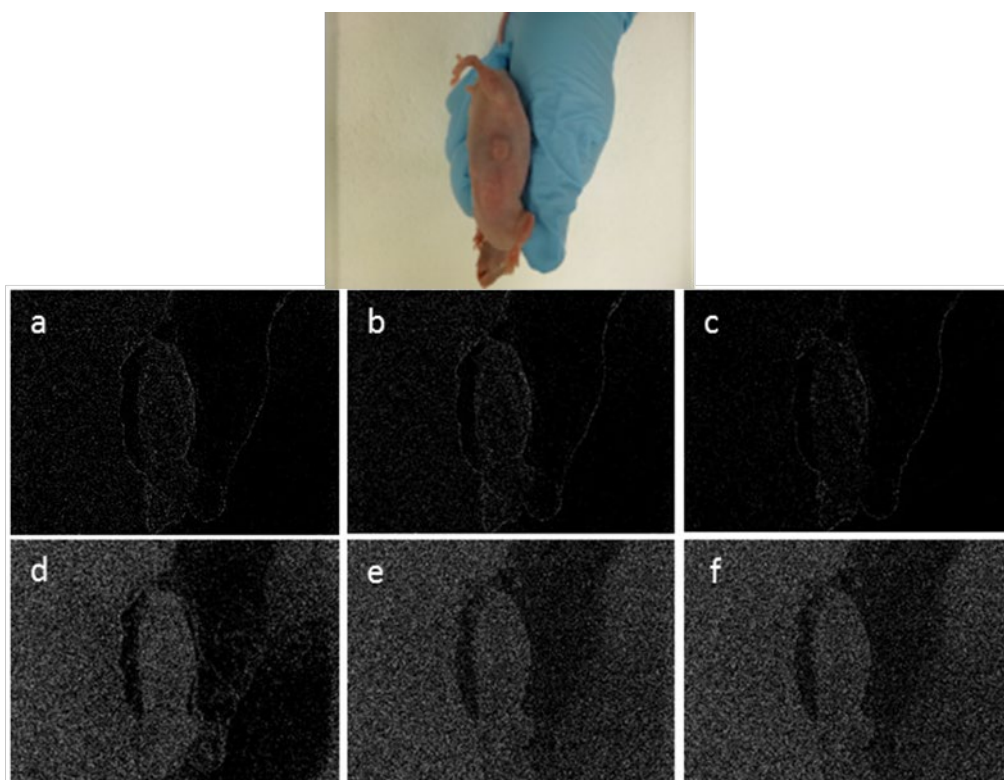
**Figure 3.8:** Right hand side image shows result of edge detection on eroded mouse mask image on left hand side.

Edge detection on the eroded image highlights the outline of the eroded mouse and does not identify the tumour boundary and thus is unsuccessful in isolating the tumour.

As the mouse image was modified through colour segmentation and erosion, this resulted in reduced information in the image. To rectify this, multiple edge detection techniques in addition to Sobel were applied to the original unmodified

image containing the subcutaneous xenograft tumour, Figure 3.9. None of these edge detectors were successful in outlining the tumour.

Figure sections a, b and c show the results for the Sobel, Prewitt and Roberts edge detectors, respectively. A weak and broken edge, representative of the mouse outline, is visible and there is no indication of a tumour outline. Figure sections d, e and f, show the results for the Canny, Laplacian of Gaussian and Zerocross edge detectors, respectively. Although the hand and mouse can be deciphered, there is no clear edge. Instead, multiple edges have been consistently detected across the mouse and image background resulting in a disrupted image.

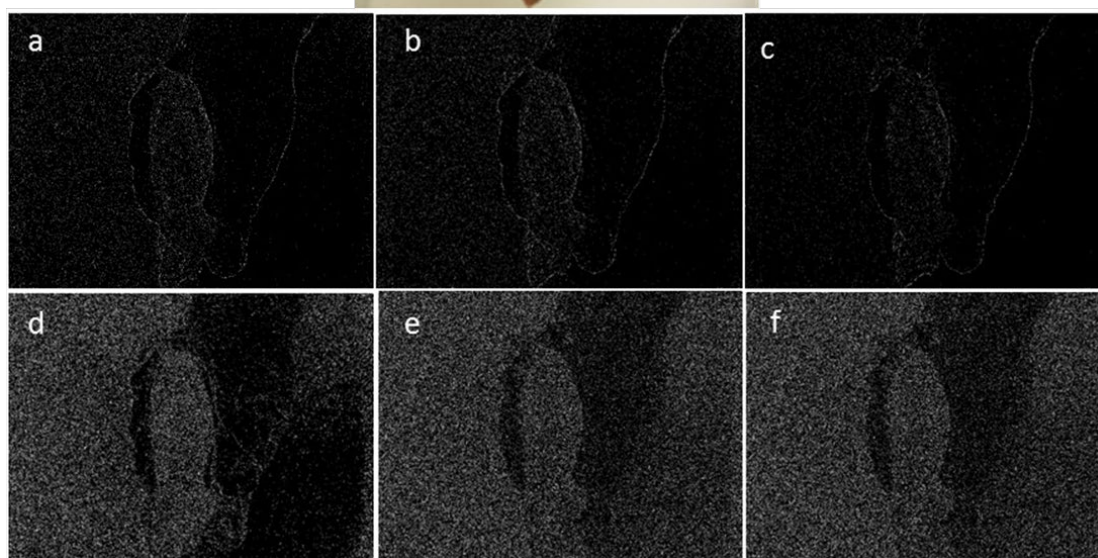


**Figure 3.9:** Edge detection applied on the original mouse image to identify the tumour boundary.

Various edge detectors applied on the image; a) Sobel b) Prewitt c) Roberts d) Canny e) Laplacian of Gaussian f) Zerocross. All edge detectors were unsuccessful in identifying the tumour boundary.

To improve upon the edge detection results seen in Figure 3.9, thresholds were applied to the image, ensuring the edge detectors would only identify edges within the threshold limits, Figure 3.10. Applying a threshold was unsuccessful in improving the edge detection or the quality of the resulting image.

Edge detection with thresholds applied to digital image:

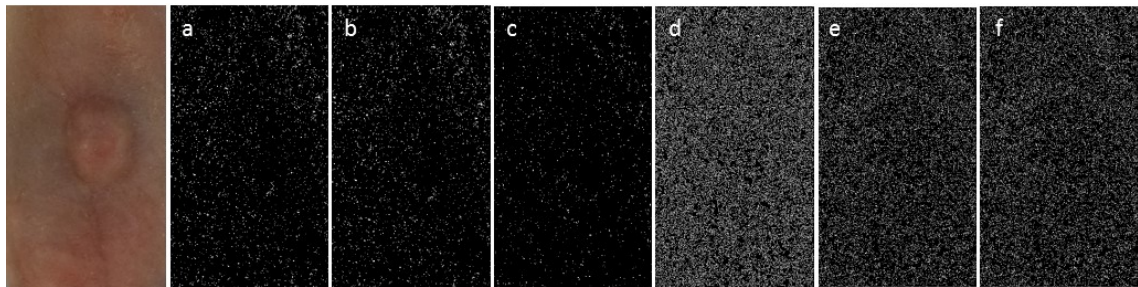


**Figure 3.10:** Edge detection with a threshold applied on original mouse image to identify the tumour boundary.

Various edge detectors applied on the image; a) Sobel b) Prewitt c) Roberts d) Canny e) Laplacian of Gaussian f) Zerocross. All edge detectors were unsuccessful in identifying the tumour boundary.

Edge detection on the original image has thus far been unsuccessful. To focus on the tumour boundary edges, background information, including the hand and mouse limbs, was removed by cropping the image to only display the tumour and its surroundings. Multiple methods of edge detection were applied on the cropped

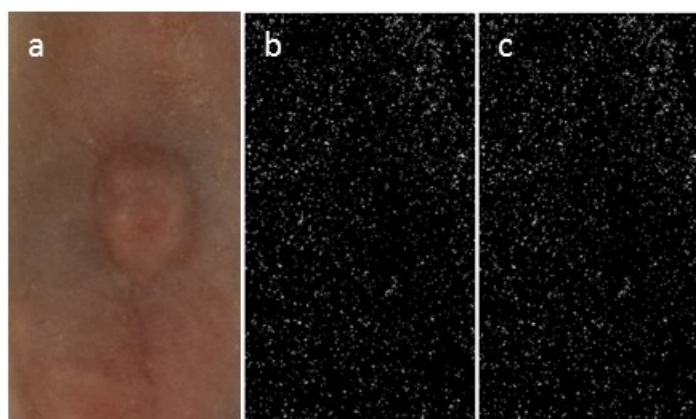
image, Figure 3.11. Similar to previous findings, there were no defined edges corresponding to the tumour outline.



**Figure 3.11:** Edge detection on the mouse flank showing the tumour and its surroundings was unsuccessful in detecting the tumour boundary

Various edge detectors applied on the image; a) Sobel b) Prewitt c) Roberts d) canny e) Laplacian of Gaussian f) Zerocross. All edge detectors were unsuccessful in identifying the tumour boundary.

To refine the process further, the image was subjected to noise removal using the Wiener method. Noise, disruptions in image edges caused at the time of image capture, can sabotage the edge detection methods. Noise was removed from the flank of the mouse before the Sobel edge detection technique was applied, Figure 3.12. Removing the noise had no effect on the edge detectors and the result remained unsuccessful in identifying the tumour boundary. Therefore, the next step was to try to find contours in the image which would permit isolation of the tumour from the mouse body.

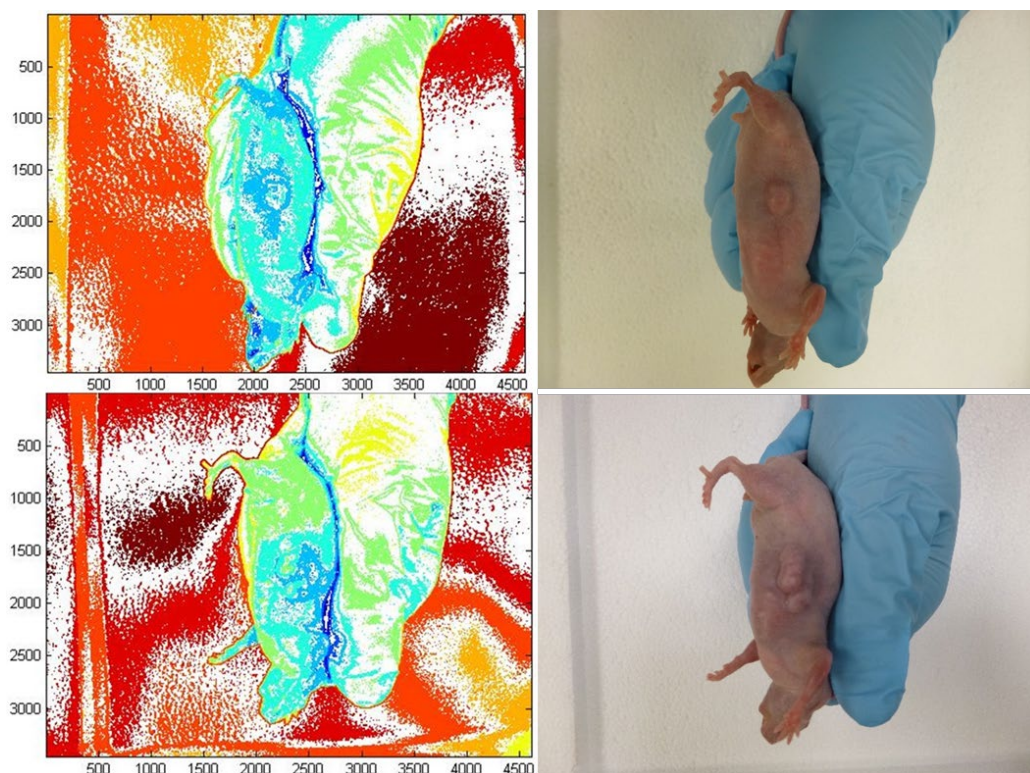


**Figure 3.12:** Noise reduction on image of tumour

Figure a) shows the image of the tumour on which edge detection was carried out; b) shows the result of edge detection using the Sobel method on the image from a and c) shows the result of the edge detector after noise removal using the Wiener method.

### 3.1.3 Contours

Where edge detection identifies changes in image intensity in local or neighbouring pixels, contours identify boundaries within groups of pixels. The following Figure 3.13 shows two xenograft images with their corresponding contour maps. The first contour map of a spherical tumour imperfectly outlines the tumour and the second contour map, of a multi-lobed tumour, has a partial tumour boundary outline. Although partial differentiation between the tumour and mouse boundary was achieved, tumours are highly irregular and as illustrated, this method is inconsistent for irregular tumour shapes. Furthermore, the mice on the contour maps have differing colour profiles, with the first mouse predominantly blue and the second mouse a mix of green and blue thus highlighting the inconsistency of features between images in the study.



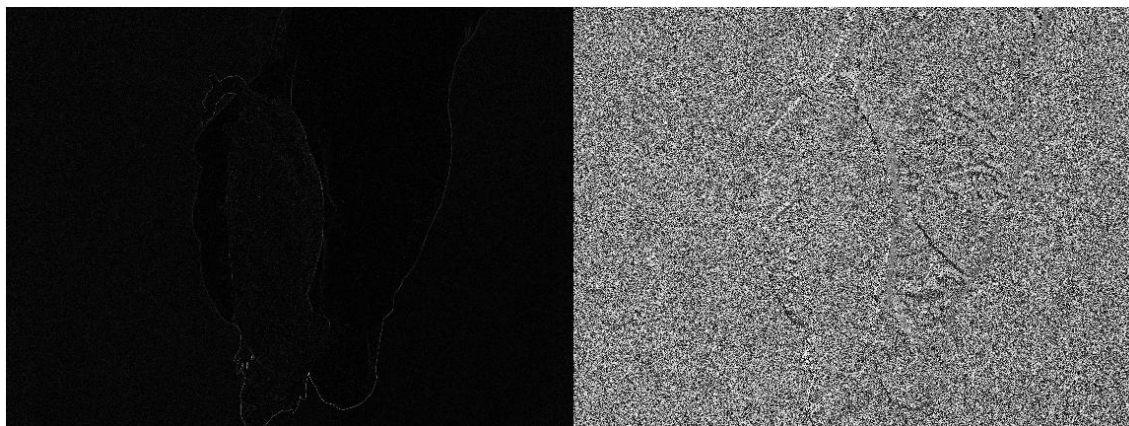
**Figure 3.13:** Xenograft images with their contour maps

The contour maps were created to identify the tumour from the mouse body. Although the first image with a spherical tumour shows an imperfect tumour boundary and the second image with a multi-lobed tumour shows a partial tumour boundary, the method is inconsistent in distinguish between the tumour and the flank for irregular tumour shapes.

### 3.1.4 Gradients

As contours were ineffective in isolating the tumour from the mouse body, the next logical step was to look at gradients since these allow evaluation of the magnitude and direction of pixels to identify boundaries. This method was used to isolate the tumour from the mouse body. As seen in Figure 3.14, this method was unsuccessful. The tumour and the mouse itself are not observable and the only image feature visible was a weak outline of the hand. Therefore, this technique was not useful for extracting information from the images. The next

step to isolate the tumour was to apply a threshold onto the sample image to ensure the tumour would be detected.



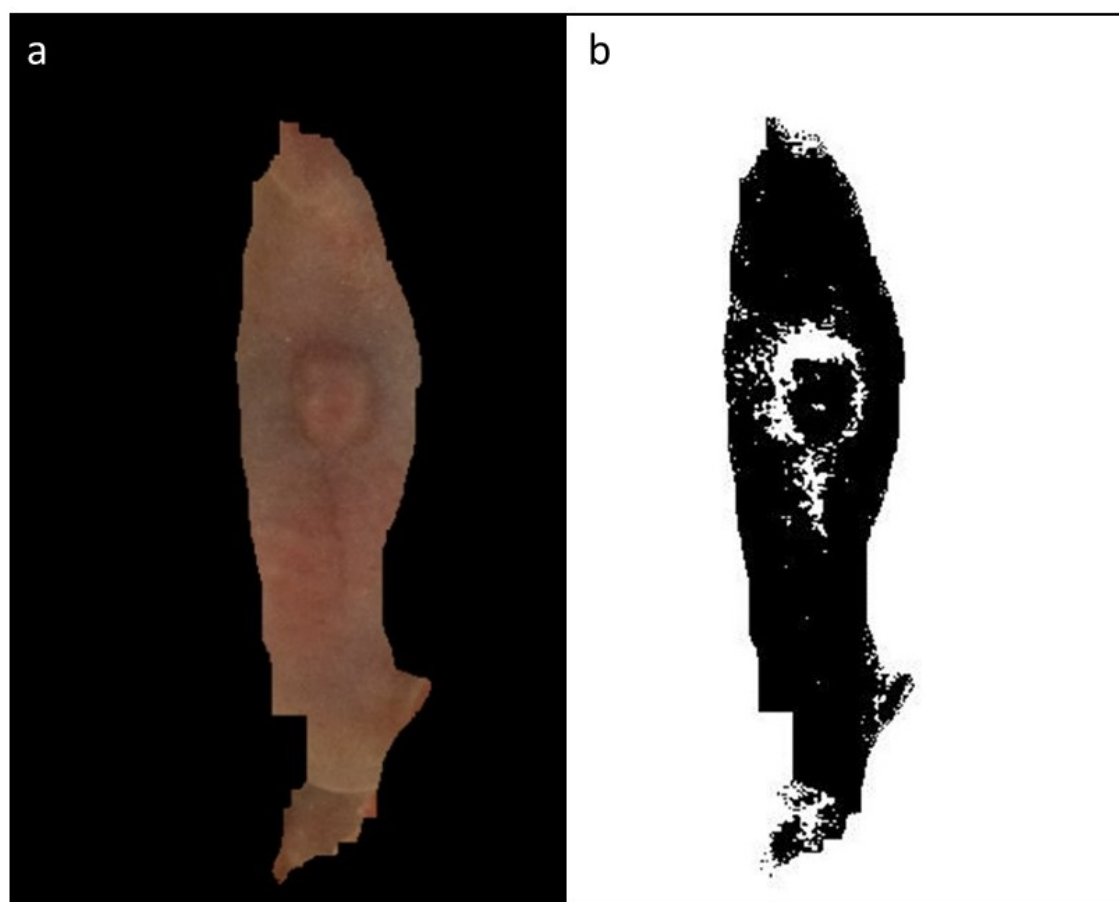
**Figure 3.14:** Gradient magnitude and gradient direction applied to the sample image. Gradient magnitude is shown on the left and gradient direction is on the right. Both images show a very weak boundary for the hand and mouse body with no indication of the tumour body thus this method was unsuccessful.

### 3.1.5 Threshold Image

Where previous techniques focussed on finding the tumour boundary in the original image, the following section focused on manually finding pixel intensity at the tumour boundary and using it to threshold the image.

The pixel intensity at the tumour boundary of the cropped RGB image produced through colour segmentation and erosion techniques (Figure 3.7) was used due to the reduced information in the image, thus ensuring the tumour boundary would be detected. The threshold value for the xenograft image shown in Figure 3.15 (a) was 70 and applying this as a threshold produced the image in Figure 3.15 (b). Although the tumour boundary has been roughly outlined, the boundary

extends down the flank and also captures the area surrounding the tumour thus rendering the tumour boundary itself indiscernible.



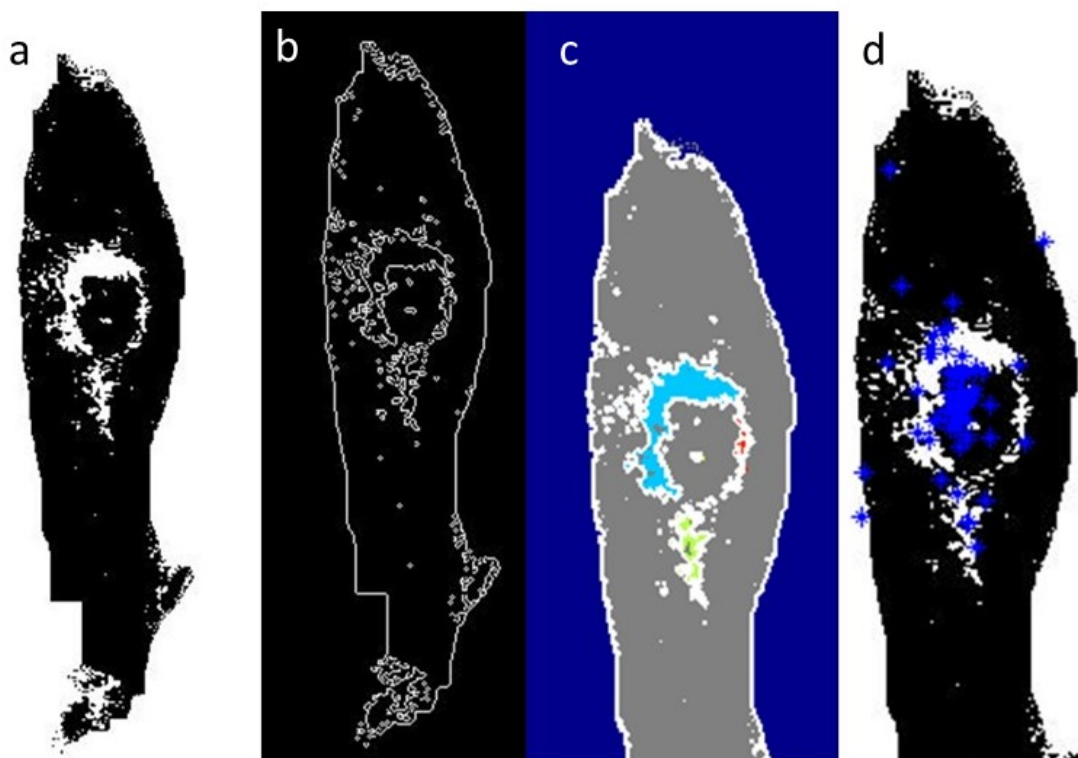
**Figure 3.15:** Original image with corresponding threshold image

Figure (a) showing the original cropped image from which the threshold was found from the tumour outline and (B) shows the result of using the threshold on the image. Although applying the threshold shows the tumour outline, the outline extends into the flank of the body thus the method is unreliable.

Multiple image processing procedures were then performed on the threshold image to identify the tumour boundary from the mouse body. The procedures included, edge detection, finding boundaries and finding centroids as described in methods section 2.2.4.7, Figure 3.16. Using edge detection and boundaries produced an incomplete and corrupted outline of the tumour, showing no consistency. Although the majority of centroids, represented with blue markers,



were based in the centre of the tumour, there were outlying centroids situated around the mouse of the body including the edge of the image indicating that this is not a reliable method to use for detecting the tumour.



**Figure 3.16:** Threshold image with image processing techniques

Figure (a) showing the original threshold image; (b) edge detection; (c) boundaries and (d) centroids (represented as blue stars). All three techniques were unsuccessful as the edges and boundaries imperfectly delineated the tumour outline but also included large areas of the flank that would need further segmentation. The majority of centroids were located in the tumour but there were numerous outliers on the edge of the mouse body indicating the unreliability of this method.

The image processing methods described thus far have been unsuccessful in locating the tumour in the original image displaying the nude mouse with a subcutaneous xenograft tumour. This indicates that the image does not contain the required information that can be exploited or extracted using these techniques, and thus other techniques were evaluated, focusing on machine learning. Where image processing focusses on specific features within a single

image such as the colour profile or the edges, machine learning involves extracting multiple features from multiple images, and thus it was thought that this may give a better chance of tumour detection given the heterogeneity of the samples and limited differential in features between tumour and background.

## **3.2 Tumour Detection – Machine Learning**

Machine learning methods were used to train classifiers with a view to differentiate tumours from non-tumours and therefore detect tumours in digital images of nude mice with subcutaneous xenograft tumours. Results obtained through cascade classification and pre-trained CNNs (methods described in Chapter 2.2.5) are presented below.

### **3.2.1 Cascade Classification**

Cascade classification was used to train a classifier to identify a tumour from a non-tumour through learning in multiple stages. Training was achieved through positive and negative images showing tumour and non-tumour components, respectively. Table 3.1 shows the dataset of images used for training and testing.

Finally, the classifier identified tumour regions on test images by enclosing the region it identified as tumour in a yellow box. Four versions of the test were executed.

**Table 3.1:** Experiments ran using cascade classification for tumour detection

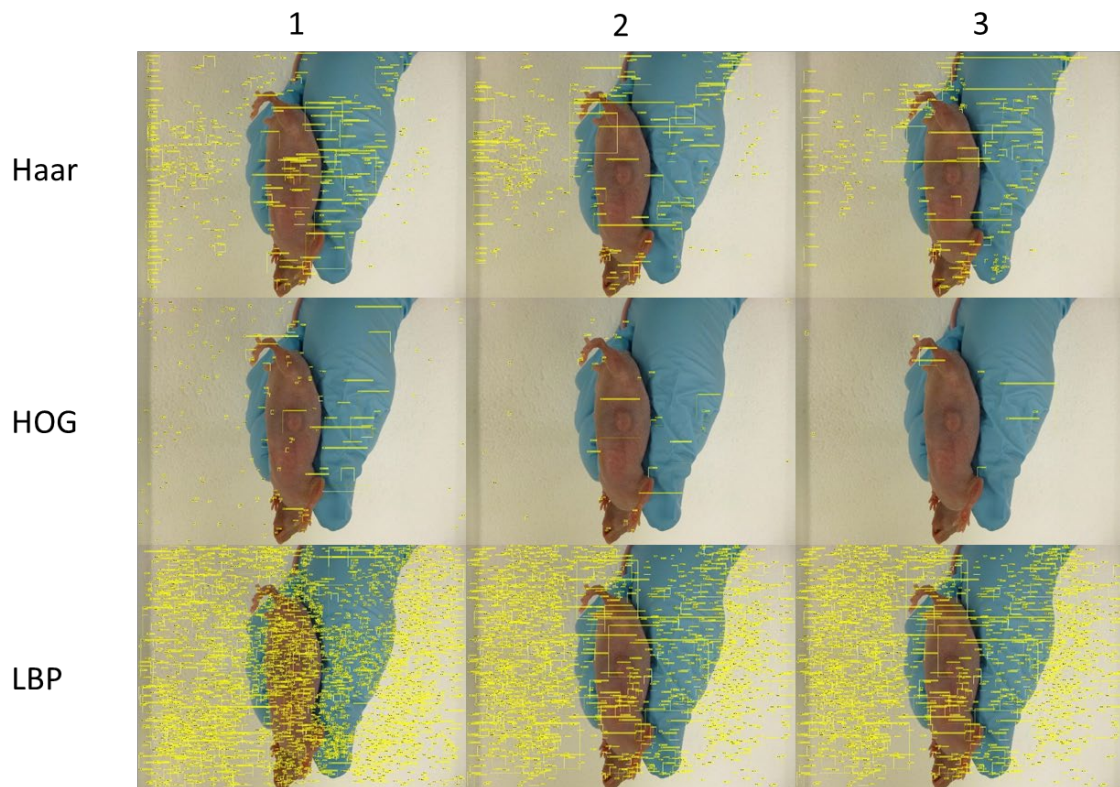
Each version of cascade classification had multiple sub-tests to allow the testing of different parameters including the feature type, true positive rate and false alarm rate. The difference in Tests 2, 3 and 4 are the training and testing images. \*The feature type that performed the best in version one was used in subsequent version.

Test	Sub-test	Feature type	True Positive Rate	False Alarm Rate	Number of Cascade stages
<b>1</b>	1	Haar	0.7	0.2	5
	2	Haar	0.9	0.4	5
	3	Haar	0.995	0.5	5
	4	HOG	0.7	0.2	5
	5	HOG	0.9	0.4	5
	6	HOG	0.995	0.5	5
	7	LBP	0.7	0.2	5
	8	LBP	0.9	0.4	5
	9	LBP	0.995	0.5	5
<b>2</b>	1	*	0.7	0.2	5
	2	*	0.9	0.4	5
	3	*	0.995	0.5	5
<b>3</b>	1	*	0.7	0.2	5
	2	*	0.9	0.4	5
	3	*	0.995	0.5	5
<b>4</b>	1	*	0.7	0.2	5
	2	*	0.9	0.4	5
	3	*	0.995	0.5	5

### 3.2.1.1 Version 1

The first version of cascade classification tested HOG, Haar and LBP feature types, Figure 3.17. Each feature type was tested three different times using different parameters to find the most successful combination in detecting tumours, details of which are specified in Table 3.1. The classifier was both trained and tested with the original xenograft images. Evidently, all three feature types had multiple negative classifications, Figure 3.17. LBP proved to be an ineffective method to identify the tumour due to the large volume of negative classifications concentrated around non-tumour components in the image,

particularly the image background. Both HOG and Haar were able to identify tumours albeit the classification also included 98-100% false classifications which varied from image to image.



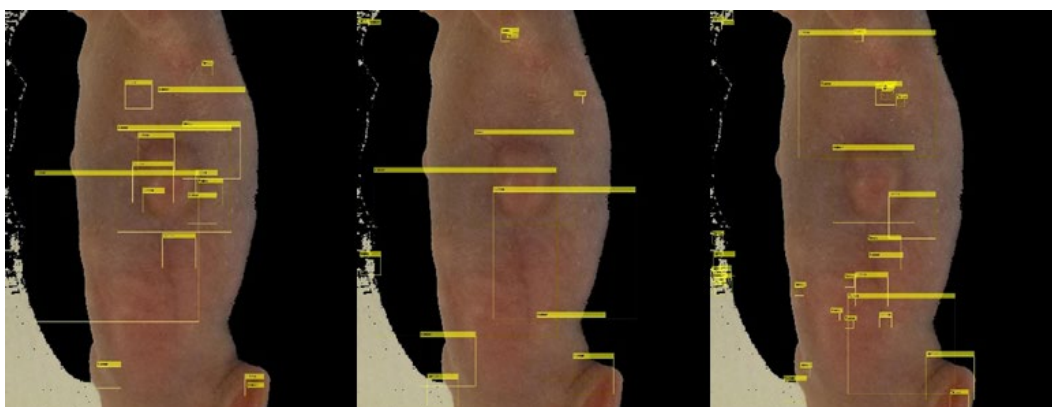
**Figure 3.17:** Cascade classification detection results to locate the tumour in the sample image

The rows show the three feature types tested, Haar, HOG and LBP. The columns represent the different tests run. Test run 1, 2 and 3 had a true positive rate of 0.7, 0.9 and 0.995, respectively, and a false alarm rate of 0.2, 0.4 and 0.5, respectively. All tests had 5 cascade stages. The classifier identifies the tumour by drawing a yellow box over the region. Evidently, all three feature types were unsuccessful in detecting the tumour but Haar features were the better of the three as test runs 1 and 2 both have a yellow box identifying the tumour.

Where the first version of classification included training and testing on the original xenograft images, the next three versions used a variety of cropped images, Figure 3.18.

### 3.2.1.2 Version 2

The second version of classification was conducted on cropped images, Figure 3.18. The images were cropped to focus on the mouse body and the image background was transformed to black to reduce the information in the image. Haar features were extracted and as previously, the test was executed three times with different parameters, as specified in Table 3.1. Although the tumour was correctly identified in experiments 2 and 3, numerous negative classifications indicated the process required further refinement.



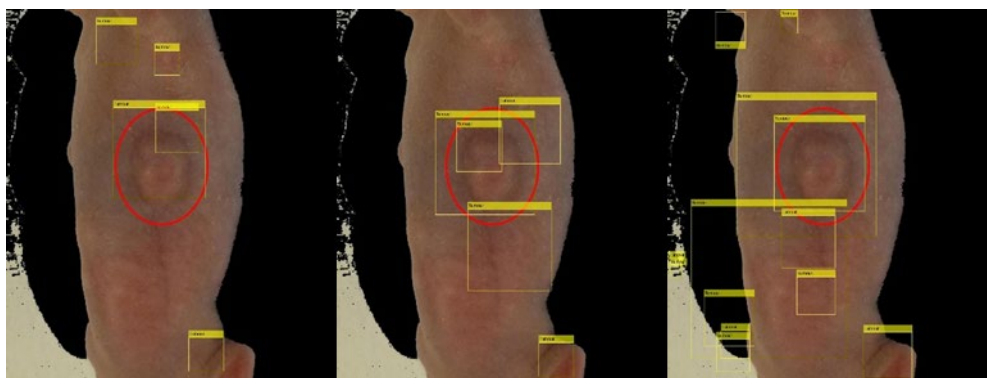
**Figure 3.18:** Version 2 of cascade classification testing to identify the tumour on the cropped image through enclosing the tumour in a yellow box.

Version 2 of the classification was trained and tested on the cropped image with a black background, Test run 1, 2 and 3 had a true positive rate of 0.7, 0.9 and 0.995, respectively, and a false alarm rate of 0.2, 0.4 and 0.5, respectively. All tests had 5 cascade stages. Version 2 was successful in detecting the tumour in experiments 2 and 3 but there were multiple negative classifications.

### 3.2.1.3 Version 3

In the third version of testing, the cropped image from the previous version of testing was supplemented with a red circle enclosing the tumour. These images were used for both training and testing the classifier. Figure 3.19 shows

successful classification for experiments 2 and 3. Furthermore, the instances of negative classifications decreased from the previous version of testing.

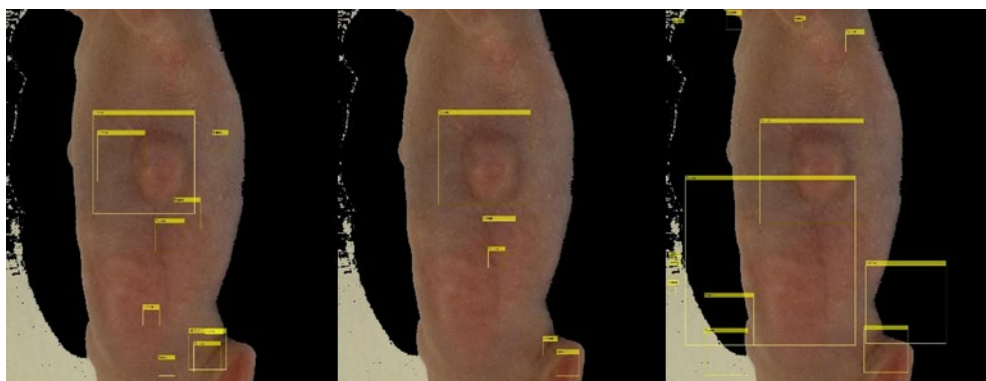


**Figure 3.19:** Version 3 of cascade classification testing to identify the tumour on the cropped image through enclosing the tumour in a yellow box.

Version 3 was trained and tested on the cropped image with a red circle highlighting the tumour. Test run 1, 2 and 3 had a true positive rate of 0.7, 0.9 and 0.995, respectively, and a false alarm rate of 0.2, 0.4 and 0.5, respectively. All tests had 5 cascade stages. Version 3 successfully detected all 3 tumours with reduced negative classifications as compared to version 2 but the method requires the observer to add a red circle to the image which is not feasible to replace callipers.

#### 3.2.1.4 Version 4

As this project aims to replace callipers using digital methods with minimal observer input, the final cascade classification test involved training the detectors with the red circle image as used in the previous version, but testing was completed using cropped images without the red circle. The process was executed three times using different parameters as specified in Table 3.1. Although the tumour was successfully detected in all three experiments, Figure 3.20, the number of negative classifications question the reliability and accuracy of the process. Therefore the next step was to test a more sophisticated method of machine learning, using convolutional neural networks.



**Figure 3.20:** Version 4 of cascade classification testing to identify the tumour on the cropped image through enclosing the tumour in a yellow box.

Version 4 was trained on the cropped image with the red circle but tested on the cropped image without a red circle. Test run 1, 2 and 3 had a true positive rate of 0.7, 0.9 and 0.995, respectively, and a false alarm rate of 0.2, 0.4 and 0.5, respectively. All tests had 5 cascade stages. Finally Version 4 successfully identified the three tumours but again there were negative classifications and the area enclosed by the yellow box contained the mouse body and thus would require further processing. Thus cascade classification was unsuccessful in isolating the tumour from the image.

### 3.2.2 Convolutional Neural Networks

This project aims to use digital imaging to replace callipers in deciphering subcutaneous xenograft tumour volume through calculating a 3D volume from a 2D image. The first step in achieving this goal is to detect the tumour in the image. Previous methods of image processing and cascade classification proved to be unsuccessful and so CNNs were utilised to achieve the goal. The CNNs used in this project to train classifiers to detect tumours were VGG-16, VGG-19 and VGG-face. The CNNs have multiple layers, and each layer extracts different features from the images. To ensure the layer most suited to the dataset was used, multiple layers were tested including the fully connected layers, before and after applying relu activation, Table 3.2.

**Table 3.2:** Convolutional neural networks tested in this project for tumour detection

Three versions of CNN testing and multiple layers were extracted to find the optimum CNN for detecting tumours. The mean image size is provided by and is a requirement of the CNN.

Test	CNN	Mean Image size	Layers extracted
1	VGG-16	224x224	FC7-after-relu FC7-before-relu FC6-after-relu
2	VGG-19	224x224	FC7-after-relu FC7-before-relu FC6-after-relu
3	VGG-Face	224x224	FC8 FC7-after-relu FC7-before-relu FC6-after-relu

**Table 3.3:** Datasets used to train convolutional neural network

Three datasets were used to train the classifiers. Dataset A contained only xenograft tumour images whereas B and C had the addition of pseudo-tumour images. Varying number of images were also tested, with the number limited by computational power.

Dataset	Images used	Total positive images	Positive training images	Positive testing images
A	Xenograft	5000	4000	1000
B	Pseudo-tumour and Xenograft	6000	4800	1200
C	Pseudo-tumour and Xenograft	7000	5600	2800

Three CNNs were tested with multiple layer extractions, Table 3.2. The process is described in Chapter 2.2.5.2. VGG-16, VGG-19 and VGG-Face were trained and tested with dataset A which consisted of 5000 Xenograft images. With a recognition rate of 1 representing 100% successful image detection, VGG-Face was the closest to this with the most successful layer, FC7-before-relu, yielding



a rate of 0.924. VGG-16 and VGG-19 were not as successful with the best rates of 0.899 with FC7-after-relu and 0.896 with FC7-before-relu, respectively.

To achieve an improved recognition rate, VGG-Face was trained with larger datasets B and C which consisted of both xenograft and pseudo-tumour images with 6000 and 7000 images respectively. Using FC7-before-relu for extraction, datasets B and C achieved recognition rates of 0.980 and 0.989, respectively. Therefore, the classifier trained with dataset C was the most successful.

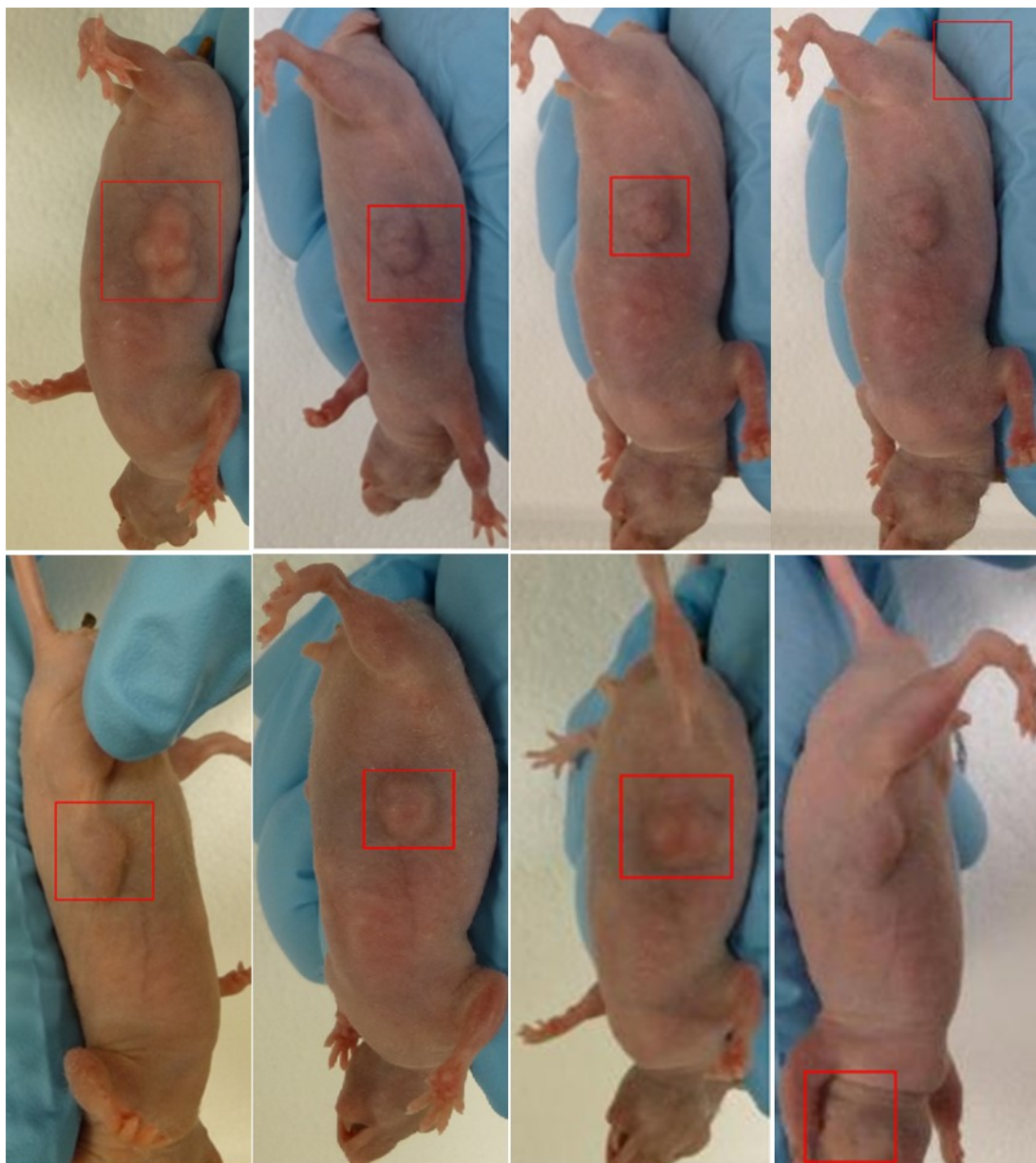
**Table 3.4:** Using CNNs to detect tumours

CNNs VGG-16, VGG-19 and VGG-Face were tested on dataset A, 5000 xenograft images and multiple layers were extracted. Fc7-before-relu extracted from CNN VGG-Face had the best recognition rate of 0.924 and so was further tested on larger datasets B and C consisting of xenograft and pseudo-tumour images with 6000 and 7000 images, respectively. Dataset C yielded the highest recognition rate of 0.989 and so was the most successful in identifying tumours.

CNN	Dataset used	Layer extracted	Recognition rate
<b>VGG-16</b>	A	FC7-after-relu	0.899
	A	Fc7-before-relu	0.898
	A	FC6-after-relu	0.895
<b>VGG-19</b>	A	FC7-after-relu7	0.882
	A	Fc7-before-relu	0.896
	A	FC6-after-relu	0.891
<b>VGG-Face</b>	A	FC8	0.907
	A	FC7-after-relu	0.904
	A	Fc7-before-relu	0.924
	A	FC6-after-relu	0.921
<b>VGG-Face</b>	B	FC7-before-relu	0.980
	C	Fc7-before-relu	0.989

To test if these results could translate to identifying the tumour on a digital image, the classifier was applied to cropped versions of the original xenograft images thus removing the background information from the image and centring on the

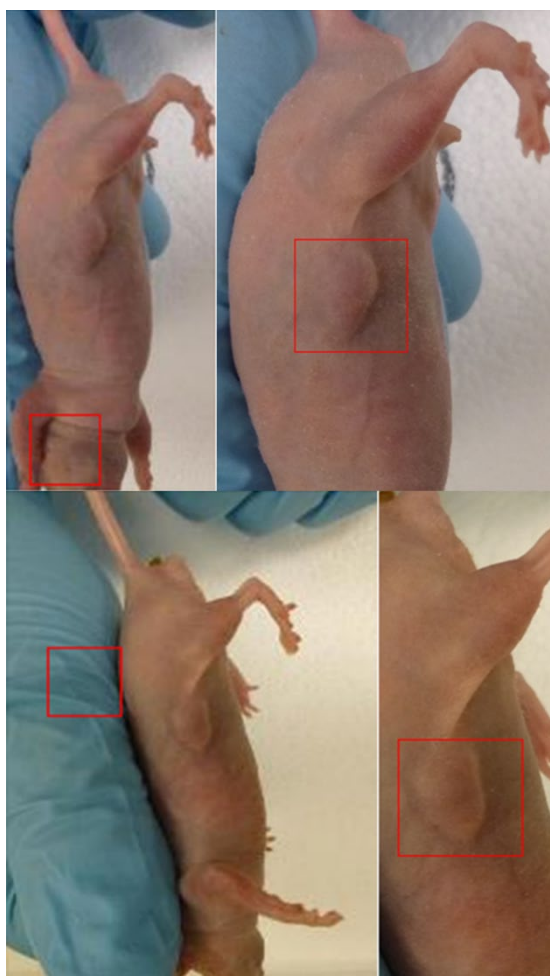
mouse body. The classifier enclosed the region it identified as a tumour within a red box. A sample of the resulting images shows 6 positive classifications of tumours and 2 negative classifications, Figure 3.21.



**Figure 3.21:** Detection results using classifier trained with VGG-Face with layer FC7-before-relu extracted.

A classifier with a recognition rate of 0.989 was applied to original xenograft images and a red box was placed around the area identified as a tumour. The classifier was trained with 7000 images consisting of tumour and pseudo-tumour images using CNN VGG-Face with layer FC7-before-relu extracted. A sample of images show 6 positive tumour classifications and 2 negative classifications.

To achieve a 100% recognition rate, a semi-automatic step was introduced. Negatively classified images were cropped to only show the flank of the body, removing all outlying information. The classifier was then reapplied giving positive classifications, and the detection was repeated; results are shown in the following Figure 3.22. This technique was successful in achieving a 100% recognition rate. Additionally, it is evident that the tumour has not been completely segmented, as the tumour has been identified with a square which does not match the tumour shape. This identification method will be taken into consideration when techniques to calculate tumour volume are applied.



**Figure 3.22:** Crop image and repeat detection for 100% classification

On images with failed detection, the image was cropped to focus onto the mouse body and reduce the information in the image. Tumour detection was then repeated, with the area identified as tumour enclosed in a red box. This method was successful in identifying the tumour.

This section on tumour detection showed the progression of techniques utilised to identify a tumour on a 2D digital image, as summarised in Table 3.5. From image processing techniques, including colour segmentation and edge detection, to machine learning using cascade classification and CNNs. Ultimately, the CNN VGG-face with layer FC7-before-relu successfully identified the tumour in the original s.c. xenograft images. The next step in achieving the aim to reproduce calliper volume measurements with digital imaging through finding a 3D volume from a 2D image is to find the volume itself.

**Table 3.5:** Tumour detection summary table

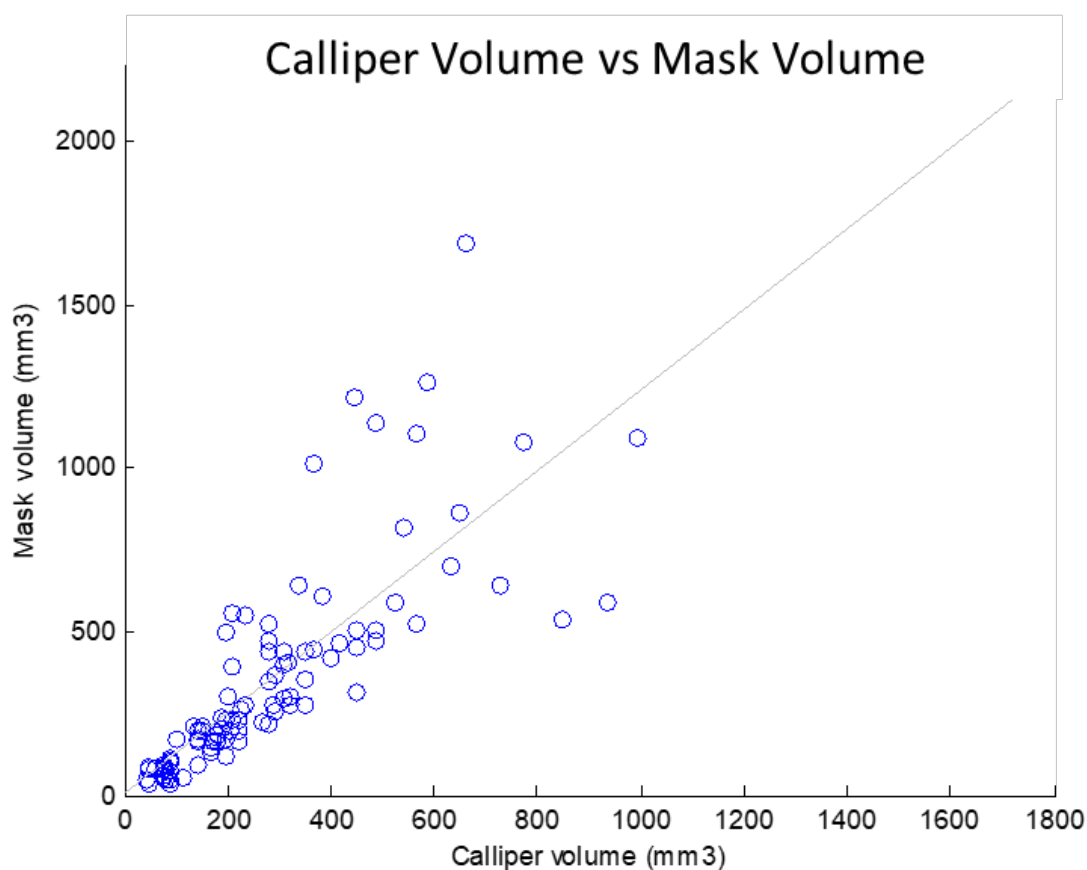
<b>Method</b>	<b>Comments</b>
<b>Colour Segmentation</b>	Using RGB colour-space was unsuccessful HSV colour-space was successful in isolating the mouse on the sample image but was unsuccessful on the full dataset of xenograft images.
<b>Morphological Operation Erosion</b>	Erosion was applied to the mouse mask created using colour segmentation and successfully removed the outlying features of the mouse including the tail, head and limbs. The resulting mask was used for edge detection and threshold imaging but the tumour boundary could not be identified.
<b>Edge Detection</b>	Edge detection was applied on the cropped image resulting from erosion in the previous step but only the mouse outline was visible. Edge detection was then applied to the original sample image using Sobel, Prewitt, Roberts, Canny, Laplacian of Gaussian and Zerocross detectors, both with and without a threshold. All detectors were unsuccessful in identifying any mouse features. The process was repeated on a cropped image focussed on the tumour and the surrounding body and again this yielded no results. Noise reduction techniques were applied to enable clear edges but this had no effect.
<b>Contours</b>	Contours were applied to a spherical tumour and a multi-lobed tumour, revealing an imperfect tumour boundary and partial tumour boundary, respectively. This method is inconsistent and would not be applicable to irregular tumour shapes.
<b>Gradients</b>	Unsuccessful, only a weak outline of the observer's hand was visible with no indication of the mouse or any of its features.
<b>Threshold Image</b>	The pixel intensity at the tumour boundary of the cropped image created using erosion was determined and used to threshold the image. The resulting image had an imperfect boundary spreading into the mouse body and down the mouse flank. Edge detection, finding boundaries and finding centroid but all were unsuccessful in recognising the tumour boundary.
<b>Cascade Classification</b>	Haar, HOG and LBP feature types were tested where Haar features were most successful in identifying the tumour albeit with multiple negative classifications. Further testing on cropped versions on the sample image with reduced information again identified the tumour but with multiple negative classifications thus rendering this an unreliable method of tumour detection.
<b>Convolutional Neural Networks</b>	VGG-Face, VGG-16 and VGG-19 were tested for tumour detection and VGG-Face with layer FC7-before-relu had the highest recognition rate. When tested on larger datasets, the classifier trained with VGG-Face layer FC7-before-relu had a recognition rate of 0.9886. To achieve a 100% recognition rate, the xenograft image was cropped and the classifier was executed again which successfully identified the tumour.

### 3.3 Tumour Volume – Image Processing

This section shows the initial attempts made to determine tumour volume from a 2D digital image. As previously described, masks of the xenograft tumours were generated from the original xenograft images and the volume was found using the modified ellipsoid formula. These are then compared to the volume found through calliper measurements, and the 3D volume, found using 3D scanning, to evaluate the accuracy of both methods.

#### 3.3.1 Using Linear Measurements to Determine Tumour Volume

Calliper tumour volume is calculated by taking length and width measurements which are then fed into the modified ellipsoid formula. First attempts at image processing of real tumour data involved making masks of digital images of the tumour to allow length and width measurements, as for callipers. This enabled a direct comparison of the modified ellipsoid formula for both callipers and the masks. Looking at Figure 3.23, there is a positive correlation (0.789) between calliper and mask volume data collected from s.c. xenograft tumours with the relationship weakening as the tumour size increases. At lower volumes under  $400\text{mm}^3$  there is a strong correlation (0.868), but with volumes over  $600\text{mm}^3$  the correlation drops to 0.081. It is not sufficient to conclude that mask volume calculated with the modified ellipsoid formula can replace calliper measurements. Using the same digital masks, different methods were used to find volume. These included using only length and width measurements as well as the mask area and are described in the following sections.



**Figure 3.23:** Relationship between calliper and mask volumes found using the modified ellipsoid formula.

Graph showing the relationship between calliper and mask volume, both calculated using the modified ellipsoid formula. There is a significant positive correlation (0.789) between calliper and mask volume data collected from subcutaneous xenograft tumours with the relationship weakening as the tumour size increases.

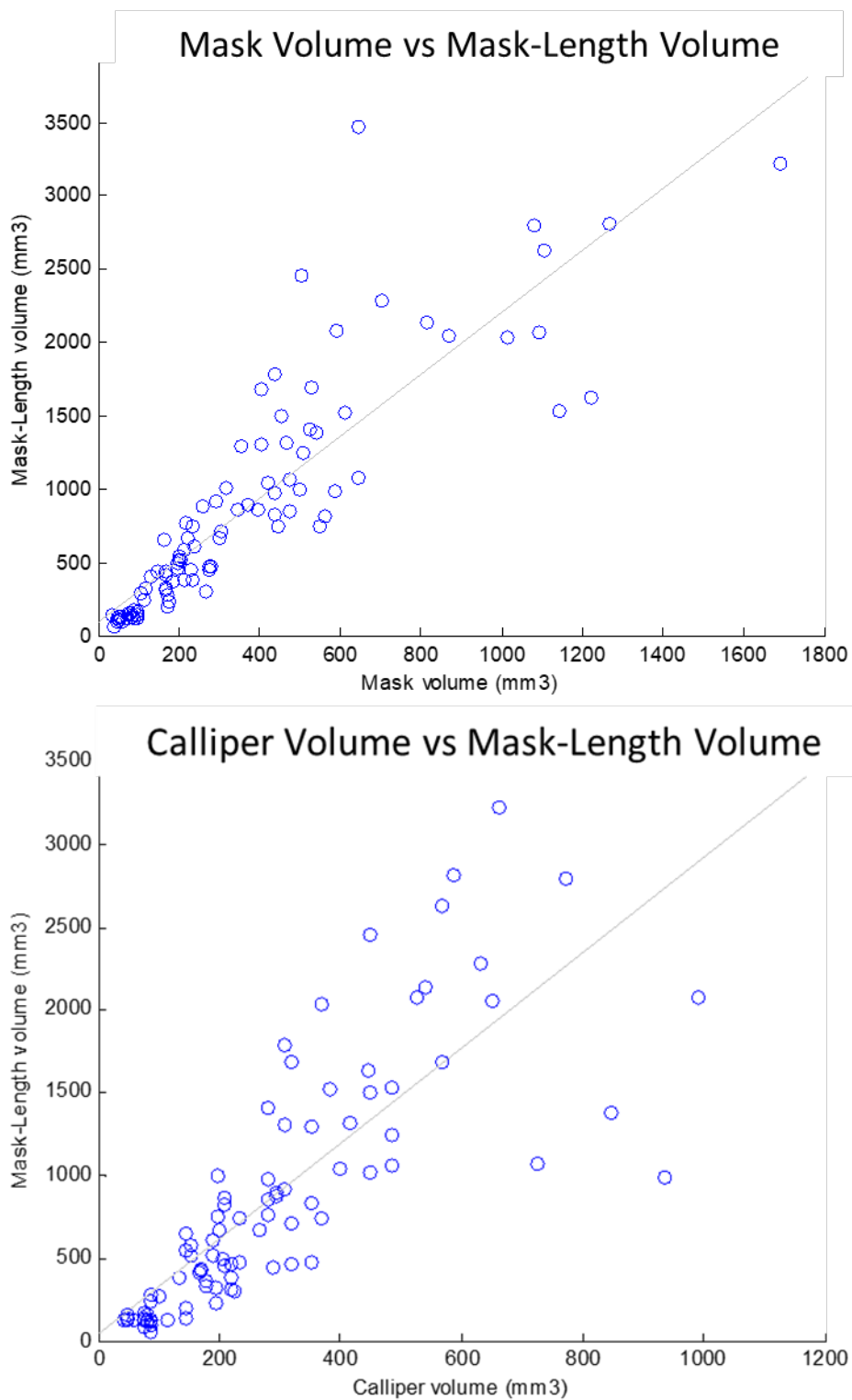
### 3.3.2 Mask Volume Created With Length and Width

The modified ellipsoid formula involves using both length and width. The formula was adapted further to include either only length or only width measurements. The purpose of this was to see if fewer parameters could be used to replicate the calliper measurements and to see the effects only considering one measurement would have on the image processing data. Figure 3.24 shows that mask volume

created using length has a positive correlation with the original mask volume (0.872) but has a poorer relationship with the calliper volume (0.760).

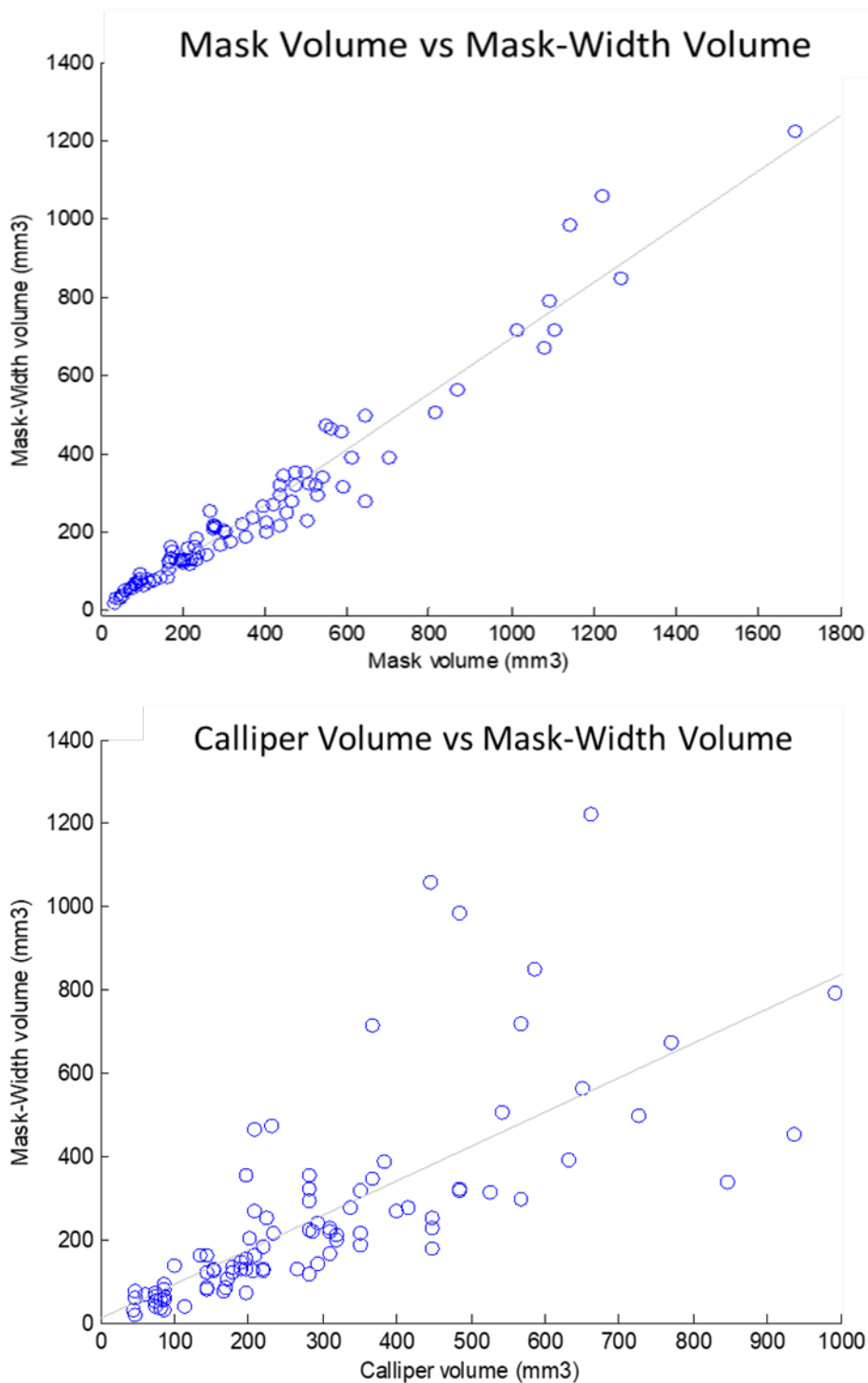
Mask volume created using width has a very strong relationship with original mask volume (0.978) showing that the original mask volume could be created using just one parameter but the relationship with the calliper (0.730) is weaker than the calliper relationship with the original mask volume (0.789). The length and width need to be evaluated further to note the difference between callipers and mask measurements.





**Figure 3.24:** Mask volume created with length Vs original mask volume and calliper volume

Graphs to show the relationships between the mask volumes created using only length, and the original mask and calliper volume. Length has a positive correlation with the original mask volume (0.872) but has a poorer relationship with the calliper volume (0.760).

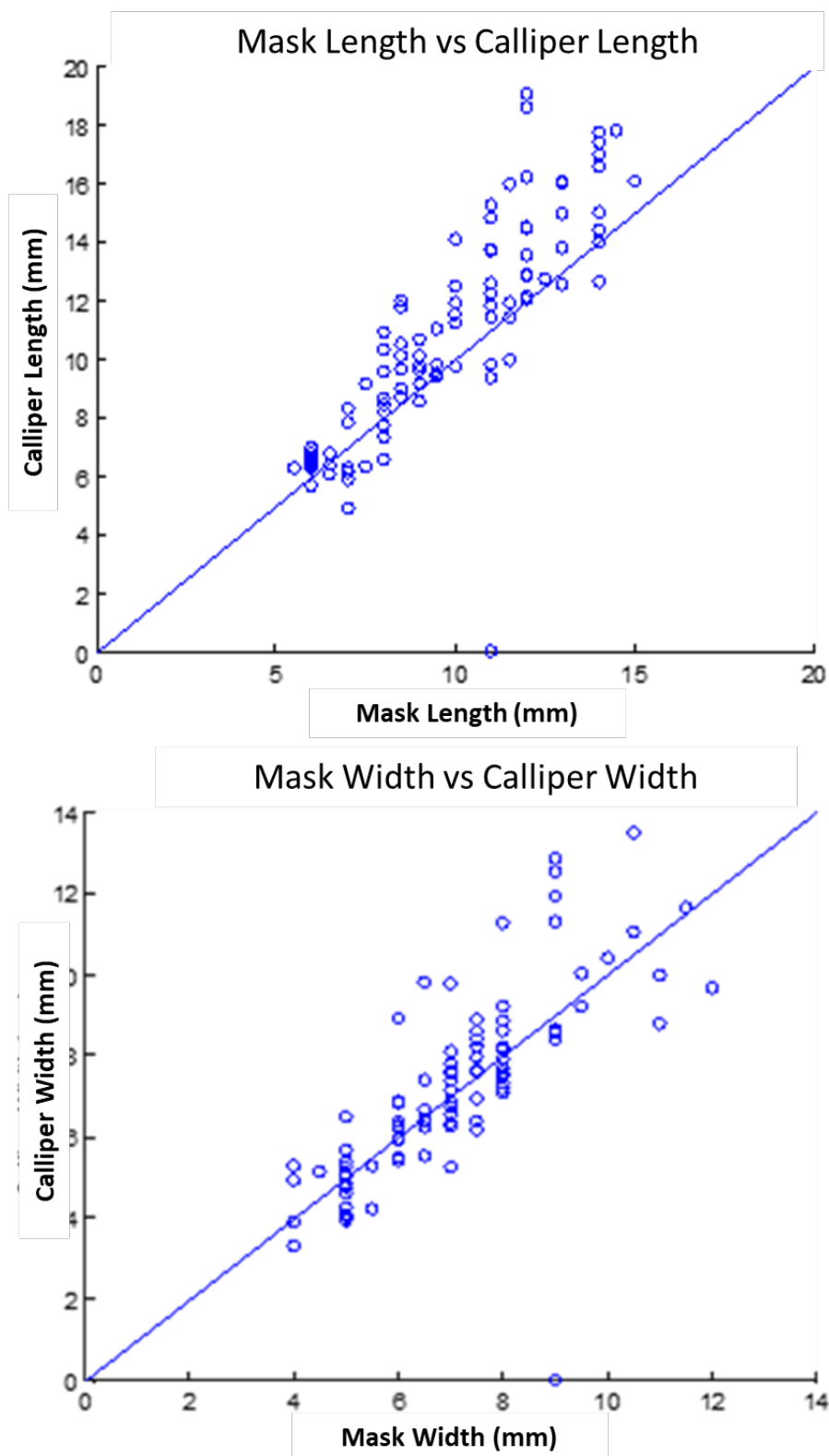


**Figure 3.25:** Mask volume created with width Vs original mask volume and calliper volume

Minor length has a very strong relationship with original mask volume (0.978) showing that the original mask volume could be created using just one parameter but the relationship with the calliper (0.730) is weaker than the calliper relationship with the original mask volume (0.789).

### 3.3.3 Length and Width comparison between mask and calliper

Looking at length and width more closely facilitates the understanding of the difference between mask and calliper volume data. The formula to calculate volume is the same for both, the modified ellipsoid formula, but the measurements inputted are different. The mask length has a strong correlation with calliper length (0.833) and there is a weaker relationship between the mask width and the calliper width (0.750). The 45° lines in the graphs (Figure 3.26) are not regression lines to fit the data, instead they demonstrate what an absolute linear relationship between the parameters would look like. It can be seen the calliper is overestimating the length in width in comparison to the length and width found from the mask. The discrepancies between length and width for both callipers and mask show that other techniques should be explored to calculate volume from the masks instead of the using modified ellipsoid formula.

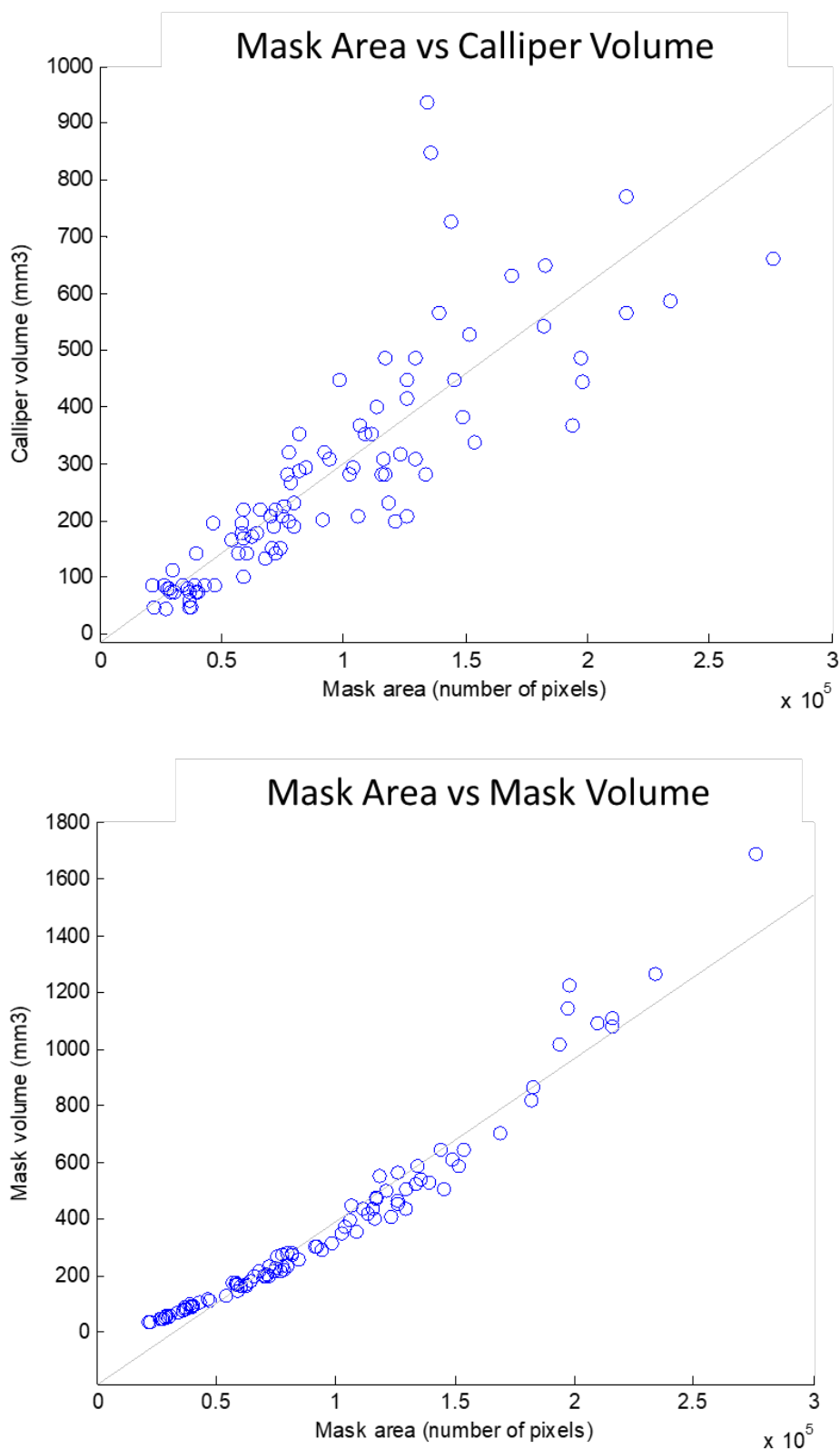


**Figure 3.26:** Comparing length and width for mask and calliper measurements

Graphs to show the relationship between length and width determined in the mask compared to calliper measurements. The mask length has a strong correlation with calliper length (0.833) and there is a weaker relationship between the mask width and the calliper width (0.750).

### 3.3.4 Mask Volume Represented by Mask Area

The mask area, i.e. number of pixels in the mask image, was seen to have a positive relationship with calliper volume (0.838), as can be seen in Figure 3.27. A much stronger relationship was noted with mask volume (0.976), determined through finding length and width and inputting these into the modified ellipsoid formula. Thus showing how the mask volume could be readily recreated without having to use the modified ellipsoid formula. Therefore, features such as the area of the tumour could be exploited to obtain volume information that is not specific to ellipsoids.



**Figure 3.27:** Mask area compared to mask and calliper volume

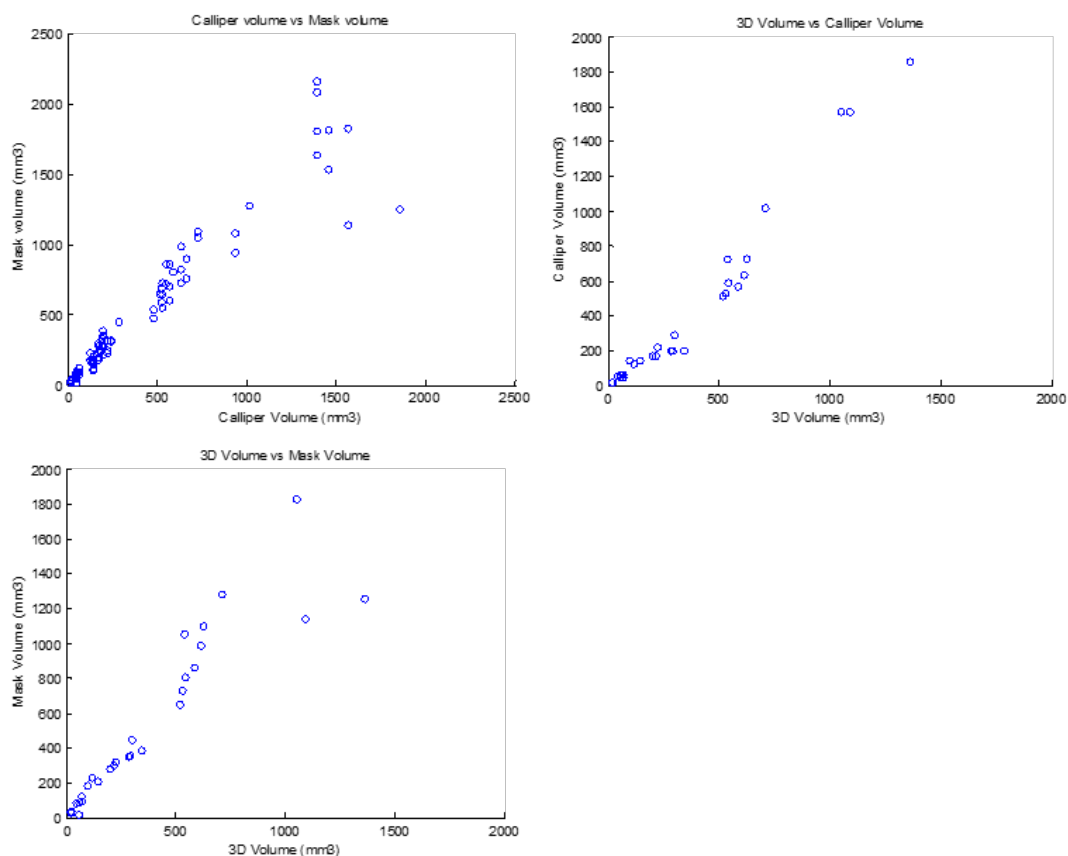
Graphs to show the relationships between mask areas, with calliper and mask volume. The mask area was seen to have a positive relationship with calliper volume (0.838). A stronger relationship was noted with mask volume (0.976).

Calculations so far have been based around the modified ellipsoid formula. In order to gain further insight, pseudo-tumour models of known volume were constructed and then compared to the volumes generated from mask and callipers.

### 3.3.5 Pseudo-tumour Model Volume

Pseudo-tumour models were created to allow 3D measurements to be taken, as well as calliper measurements and 2D masks from digital images. These measurements could then be compared to the calliper measurements and digital image processing enabling the evaluation of the accuracy of these methods in predicting true volume, as calculated through 3D scanning. The pseudo-tumour models were made in both regular and irregular shapes to mimic tumours.

Figure 3.28 shows that there is a high correlation (0.952) between calliper measurements and mask volumes obtained from the pseudo-tumour models. Although both have a high correlation, the calliper volume is closer to the true volume for the artificial model (0.981), as the mask volume has a weaker correlation (0.929). The models were constructed to mimic tumour xenografts. Calliper measurements of the artificial models were not consistent for all three shape types with a stronger correlation for the abstract shape and a weaker correlation for the kidney bean shape, showing that the modified ellipsoid formula can't predict true volume for a range of different tumours.



**Figure 3.28:** Pseudo tumour volumes from 2D masks, 3D scanning and callipers

Graphs to show the relationship between pseudo-tumour volumes determined through 2D masks, 3D scanning and callipers. There is a high correlation (0.952) between calliper measurements and mask volumes obtained from the pseudo-tumour models. Although both have a high correlation, the calliper volume is closer to the true volume for the artificial model (0.981), as the mask volume has a weaker correlation (0.929).

Ultimately, the modified ellipsoid formula is a poor predictor of true volume for 2D measurements in both callipers and masks. In further work, the modified ellipsoid formula can be replaced and 2D to 3D conversion algorithms can be applied to the digital images to calculate volume which can be checked against the true volume found from 3D imaging. Therefore, the next step to determine volume was by using shape from shading.

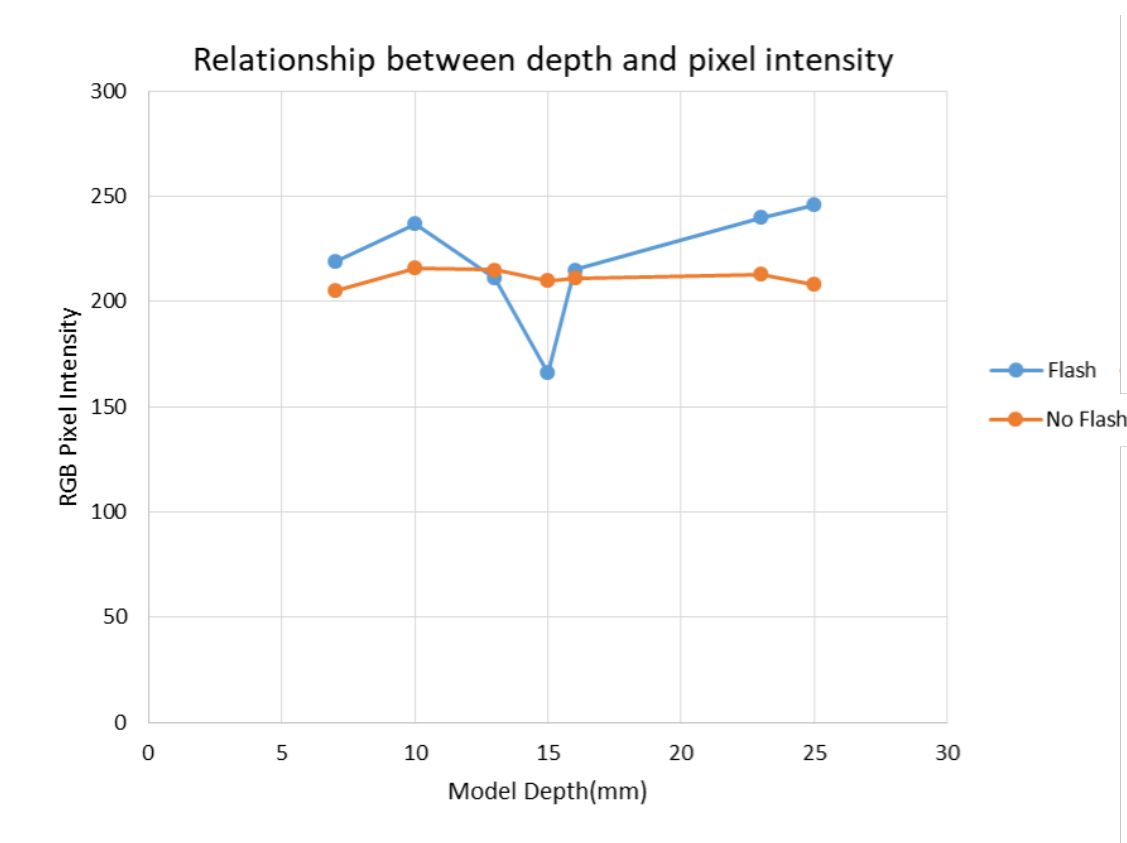


### **3.4 Tumour Volume – Shape from Shading**

In an attempt to establish a relationship between pixel intensity and depth, artificial models of different height were created and imaged from the same height both with and without a fixed light source in order to detect any changes in the pixel intensity. Next, a relationship between the spread of pixels and volume was investigated through creating virtual models with different angles of lighting.

#### **3.4.1 Depth and Pixel Intensity**

Images were captured of artificial models, with the depth varying between 7 and 23mm, from a fixed position both with and without a light source, in the form of a flash. Evidently, as seen from Figure 3.29, there is minimal correlation between pixel intensity and height (0.114). Pixel intensity remained consistent as the model depth increased whilst there was no flash. However, when flash was added, the pixel intensity became highly varied and did not follow a pattern of increase or decrease, which would suggest that addition of a light source is unlikely to assist in determining tumour depth in this instance. This is discussed further in the discussion section (see page 177).



**Figure 3.29:** Depth vs pixel intensity

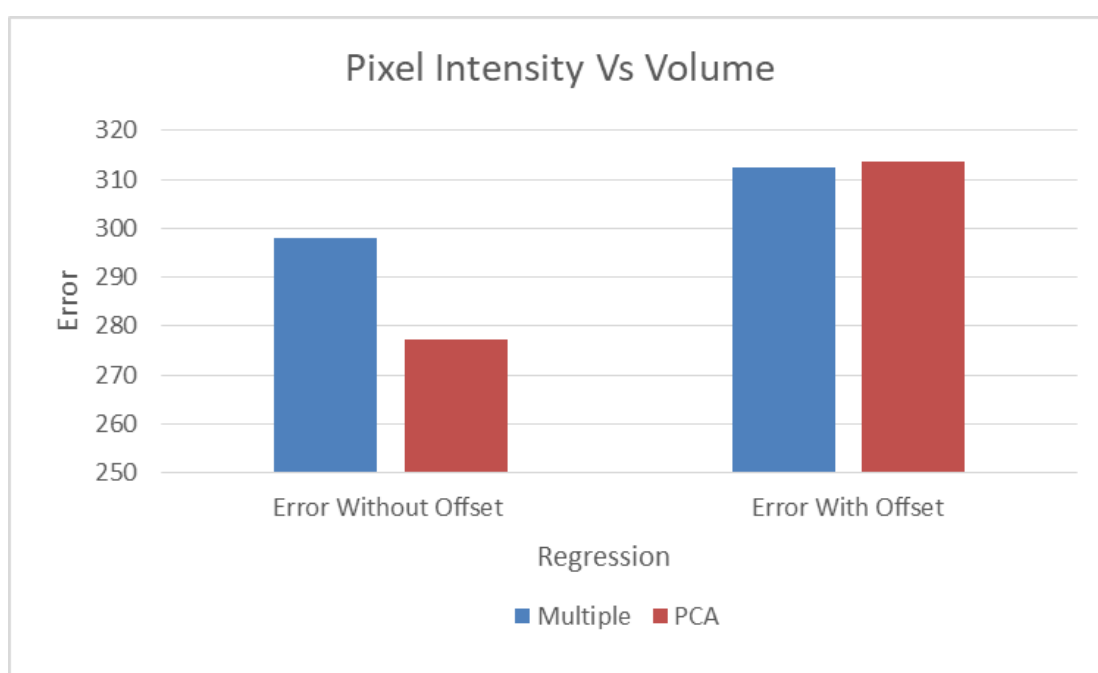
Graph to show the relationship between the depth of a model and its pixel intensity. There is minimal change in pixel intensity for models imaged with no flash. However, when subjected to flash, the pixel intensity varies and does not follow a pattern as is demonstrated by the low correlation coefficient (0.114).

### 3.4.2 Virtual tumours and different lighting angles

Virtual models resembling tumours, of different shapes and sizes, were created and different lighting angles were applied to create varying degrees of shading (Figure 2.1). The purpose of this was to determine the tumour volume by finding a relationship between shading and the volume. The pixel intensity information of the model was exported as a histogram and multiple regression analysis and PCA was conducted with volume. Figure 3.30 shows the results, both with and without offset. Offset is the point where the regression line hits the y axis. With offset

means the regression algorithm isn't forced to go through origin whereas without offset, the line is assumed to pass through origin. With offset can give better results at times as the line is not forced to go through the origin. The ideal range for this project is to have an error as close to 0 as possible.

Although error without offset was less than the error with offset, the high errors found from regressing pixel intensity with volume suggested no relationship between the two variables, Figure 3.30.

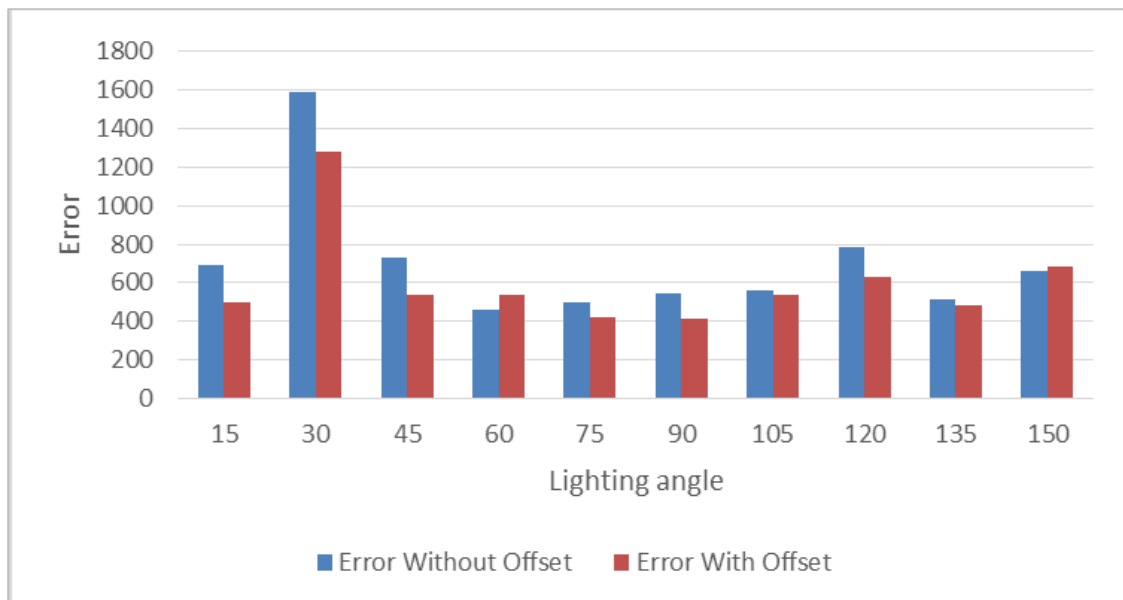


**Figure 3.30:** Pixel intensity vs. volume

Graph to show the relationship between pixel intensity and volume. The first graph shows the overall error using Multiple and PCA regression. The errors without offset are 298.05 and 277.12 for multiple and PCA regression, respectively. Errors with offset are 312.3 and 313.7 for multiple and PCA regression, respectively.

Multiple light angles were applied to the models to create shading, and multiple regression was used to find the error for each light angle separately (Figure 3.31). Although the errors vary between 414.86 (-90 X axis) and 1593.4 (-30 X axis),

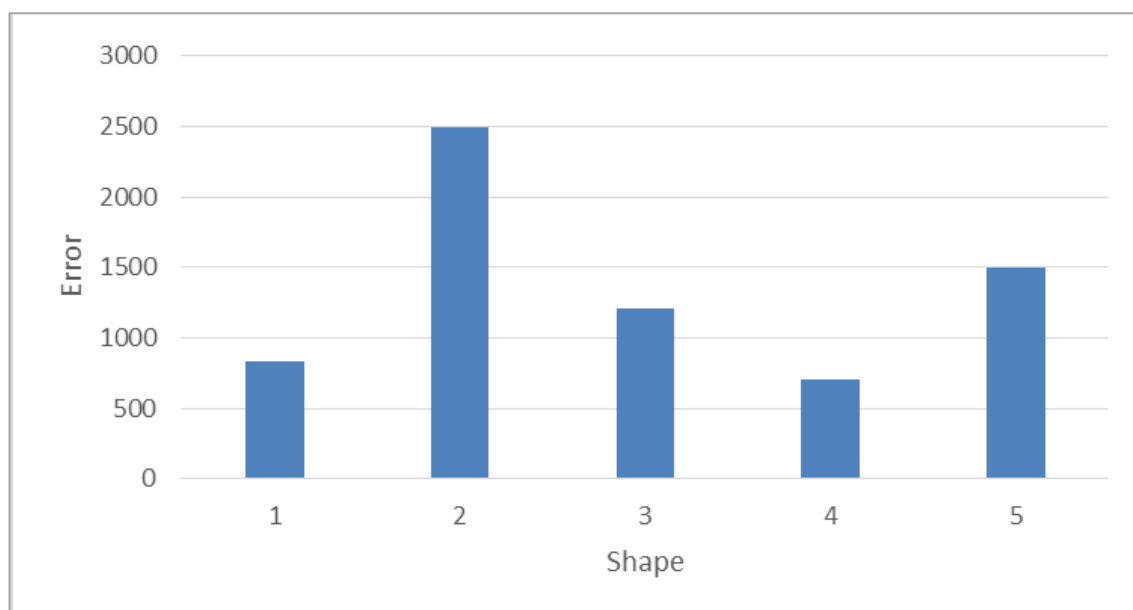
they are still very high and therefore there is a poor relationship between the effect of the angle of lighting on pixel intensity and the volume of the model.



**Figure 3.31:** Effect of lighting angle on determining tumour volume

Graph to show the error for lighting angles varying between 15 and 150 with volume. The error varies from 414 to 1593.4 and therefore, there is a poor relationship between the lighting angle and volume.

Furthermore, as the virtual tumour models were fashioned in a variety of shapes and sizes, the error was found for each individual shape to investigate if exposing a light source onto certain shapes gives an indication of the volume. The results can be seen in Figure 3.32. Although shape 4 had the best error, the error was over 500 and therefore clearly showed that there was no relationship between the shapes on pixel intensity vs. volume.



**Figure 3.32:** Effect of shape on determining tumour volume

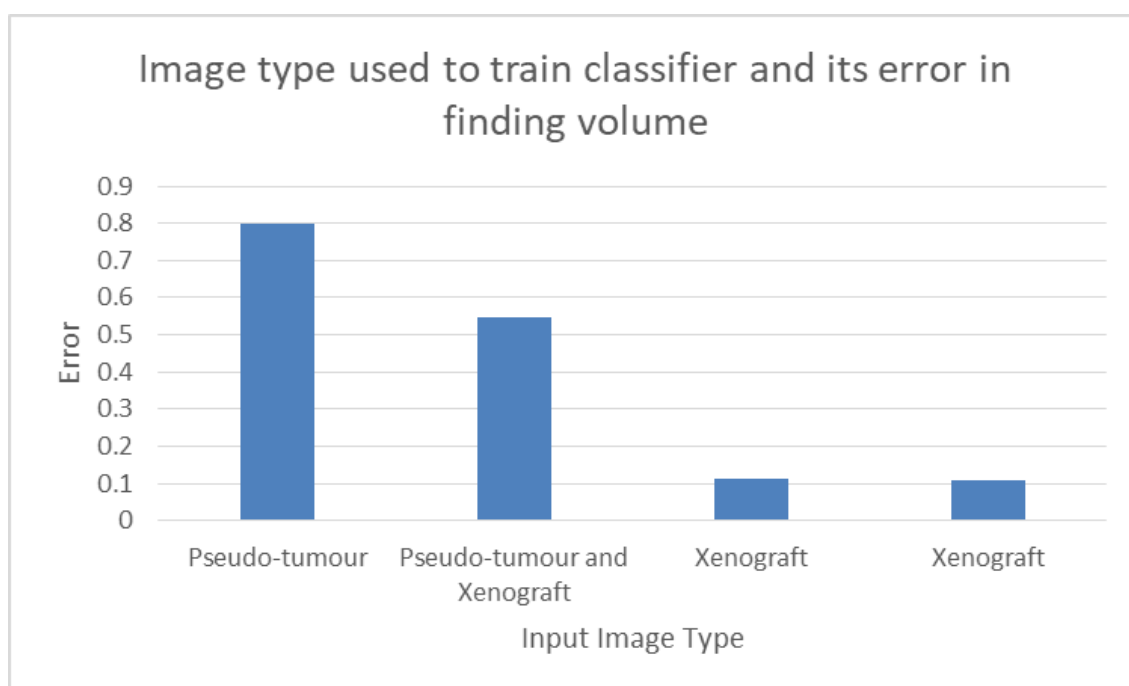
Graph showing high errors when finding a relationship between shape and volume.

The results show that using shape from shading principles to find tumour volume has been unsuccessful. Where this technique focussed on the differences in pixel intensity on the model created from an external light source, next, machine learning is used where it extracts multiple features from the models in order to find a relationship with the volume.

### 3.5 Tumour Volume – Machine Learning

The ultimate aim of this project is replace callipers with a digital imaging system in deciphering the 3D volume of s.c. tumours from a 2D image. With previous methods of image processing and shape from shading failing to replicate volume measurements found using callipers, a CNN was tested. The pre-trained CNN VGG-Face trained classifiers with tumour images and their corresponding volumes, using layer FC7-before-relu to extract features.

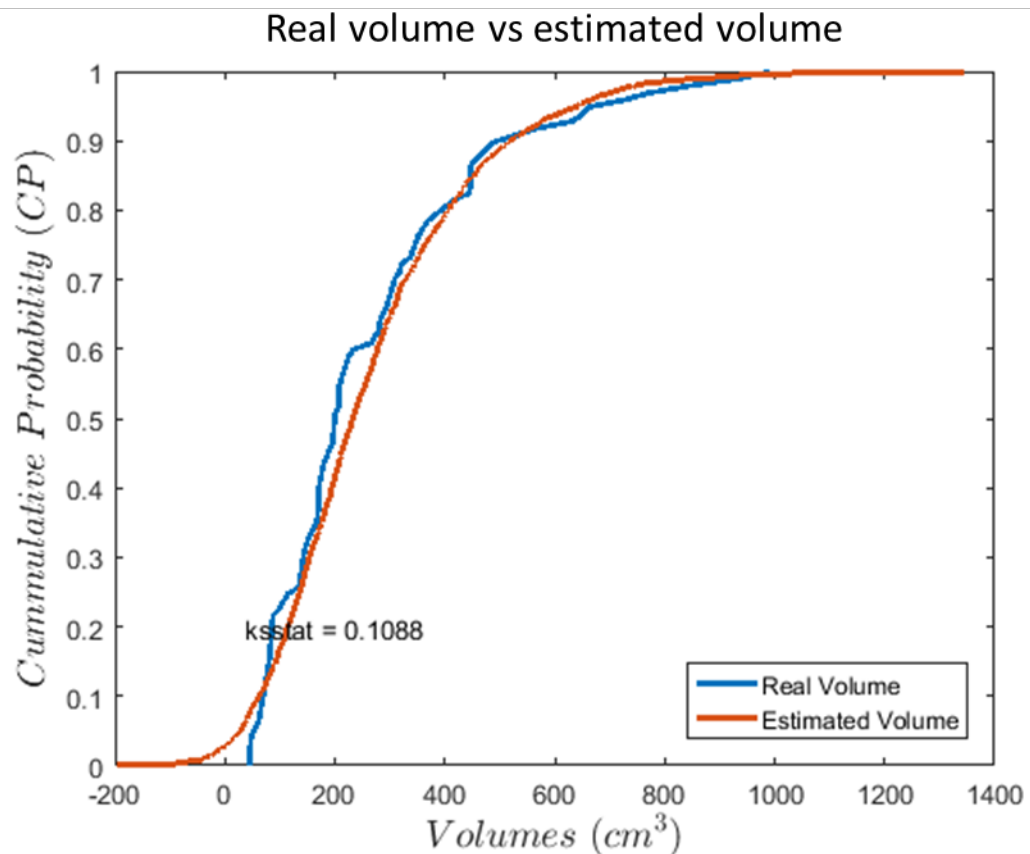
Figure 3.32 shows the datasets used to train the classifiers, specifying image type and the number of images in each dataset. Pseudo-tumour images, both exclusively and in combination with xenograft images, yielded the largest errors of 0.7970 and 0.5465, respectively. Training the classifier with only xenograft images significantly reduced the error to 0.1135 and using an increased number of images reduced the error further to 0.1081. Furthermore, using an increased number of xenograft images. Due to limited computational power and datasets this was the best error that could be achieved, and this could be improved with more extensive capabilities in both.



**Figure 3.33:** Error seen when finding volume for different datasets used to train classifiers.

Starting from the left, the data sets were trained with 3400, 6892, 2977 and 3499 images, respectively. Using exclusively tumour images obtained from xenografts reduced the error.

The trained classifier was then tested on the original xenograft images. The real volume was plotted with the estimated volume obtained from the callipers, Figure 3.34. The graph indicates a positive correlation between both real and estimated volume. The null hypothesis is accepted at 5% significance level, this says that the two samples statistically come from the same distribution. The test returned 3 values: h, p and kstat. H is a 1 or a 0, p is the probability, and kstat is the distance between the graphs. The h was 0 and the p value is 0.031 making the null hypothesis accepted. The kstat was 0.1 showing the close proximity between the two graphs.



**Figure 3.34:** Real Vs estimated tumour volume

Graph showing the relationship between the real volume and the estimated volume provided by callipers. P value of 0.031 indicates both volumes are from the same distribution and a kstat value of 0.109 show their close proximity.

#### 4. Discussion

Motivated by the inadequacies of the industry standard use of callipers to measure tumour growth in s.c. tumour transplantation models, this project aimed to replace the methodology with non-invasive digital imaging. These inadequacies include physically squeezing the tumour, yielding highly subjective results with poor accuracy and reproducibility. Furthermore, by only using 2D measurements to work out a 3D volume there is no differentiation between a flat shaped tumour and a tumour growing outwards. Through capturing 2D digital images and calculating the 3D volume measurement directly from the image, physical contact on the tumour was removed thus providing 3Rs refinement benefit to the mice. Furthermore, determining tumour volume through executing computer algorithms instead of the biased opinion of the person carrying out the calliper reading reduced subjectivity resulting in increased reproducibility and accuracy.

Current research applying digital imaging to tumour detection and measurement is focused on melanoma detection in which the boundary between skin and tumour is evident and easily identifiable due to the high contrast in colour. For example, Pereira et al enhanced their algorithm for melanoma classification using features found on the tumour border[131]. Whereas early research focussed on image processing methods to detect melanomas, the limitations of this have now been recognised and current research explores advanced computer algorithms. Johansen et al[132] use hyperspectral imaging to enhance the information obtained from standard RGB images. RGB images capture information on 3



bands, red green and blue whereas hyperspectral images capture hundreds of bands.

In this thesis, the lack of colour differentiation between the tumour and surrounding body in the nude mice used in these studies was challenging and resulted in the application of sequential methodologies ranging from image processing to machine learning to attempt to differentiate between the two. Nude mice were used in this thesis as the challenges in differentiating the tumour from the surrounding skin in hairy mice would be increased due to the overlying fur causing further disruption of the contour of the tumour thus increasing the difficulty of extracting tumour measurements. Initial work using 2D digital images of tumour xenografts were found to be insufficient to evaluate sophisticated methodologies on their own.

To investigate shape of shading principles and to satisfy the number of images required for machine learning methods for training a convolutional neural network, more images were required. Therefore, this data was augmented by virtual and pseudo-tumour models. Simple image processing methods such as edge detection were unable to initially identify the tumour or extract a volume from the 2D images. However, machine learning techniques using pretrained convolutional networks proved successful in both detecting the tumour and calculating its volume. This novel work has offered an inexpensive and simple replacement to callipers through eliminating the need to physically make contact with the tumours to obtain tumour measurements.

To satisfy the aims of this thesis to provide a non-invasive digital imaging system as an alternative to calliper measurements to obtain tumour volume measurements, the primary form of data collection involved capturing 2D digital images of s.c. tumours. Although this thesis focuses on s.c. xenografts in nude hairless mice, in future work the methods would be applied to other s.c. xenografts including those in immunodeficient hairy mice (e.g. SCID mice) or s.c. allografts in immunocompetent hairy mice. To achieve 3Rs refinement benefits by replacing callipers and therefore reducing animal handling, when capturing digital images, it was vital to remove physical contact on the tumour whilst ensuring images were consistent and reproducible. Although the tunnelling handling method was used in this study to cause less anxiety to the mouse, there is still an element of stress as the animal is restrained and moved from its environment in order to be imaged. Therefore, a home-cage detection method would be most beneficial by offering increased refinement benefits to the mouse as this method would eliminate any handling as images would be captured within the cage itself.

Fixing the camera on a copy stand and marking the surgical table at a predetermined location, clearly highlighting where the mouse would be held, enabled the observer to focus on ensuring the tumour could be seen from the camera screen. Digital images of squash balls, a regular fixed shape of known dimensions, were taken at the start and end of every session for image calibration to ensure the mouse would be placed in the correct position under the camera. Data used in this study focussed on the number and pattern of pixels in the 2D digital image captured. Although the copy stand and squash ball were in fixed position, the distance between the two was not included as a measurement of

length or breadth. This method eliminates any physical contact with the tumour as well as reducing observer variation due to fewer decisions the observer must make in comparison to the number of decisions made when taking calliper measurements. With calliper measurements the observer must hold the mouse and determine the longest length of the tumour and its perpendicular width whilst judging how much to squeeze the tumour, thus resulting in a higher number of decisions and higher rates of observer variation.

In regards to the image itself, the background of the image was purposefully kept neutral to ensure a smooth transition with minimal barriers for different labs when replicating the image format and adopting this method instead of using standard calliper measurements. The blue colour of the glove contrasted to the nude mouse, this was advantageous as it allowed for clear differentiation between the two objects in the images. One of the aims of this work is to detect the tumour and one of the first methods attempted to achieve this was colour segmentation, thus taking advantage of the different colour profiles found in the image. The challenge was to find a colour threshold separating the tumour from the surrounding nude mouse body, which proved difficult as both had the same colour profile.

Image processing techniques used to detect the tumour in the digital image included isolated methods as well as sequential methods including colour segmentation, edge detection, contours, gradients and threshold. Ultimately all were unsuccessful primarily due to the lack of colour differentiation between the tumour and surrounding mouse skin. This project relied on commonly used image processing algorithms to extract information regarding the tumour boundary itself

or to pre-process the image to make the boundary clearer or pave the way for other techniques to differentiate the boundary.

Colour segmentation has been highly successful in detecting skin melanomas and is often used for diagnostic purposes. However, the use of colour segmentation has not been limited to melanomas and has been used in various applications, including identifying white blood cells in peripheral blood and bone marrow images for the diagnosis of leukaemia. When looking for published models relevant to this work, studies on facial recognition were useful as the majority of facial recognition systems use Caucasian males which share similar colour-spaces to that of the nude mice used in this study. Both RGB and HSV colour spaces have commonly been used in this field and so were used in this project[133].

Colour segmentation in both RGB and HSV colour spaces were ultimately unsuccessful. The lack of colour differentiation between the tumour and surrounding skin meant that even if a threshold was found, other areas of the mouse body and tumour itself became corrupted as they share the same colour profile. There was also a discrepancy of the colour profile of the different mice in the same study which meant the segmentation methods which focussed on specific pixel thresholds were not applicable to every mouse in the study.

As colour segmentation was unsuccessful at identifying thresholds, edge detection methods were tested and yielded no positive results. Tested methods included Sobel, Prewitt, Roberts, Canny, Laplacian of Gaussian and Zerocross detectors. Edge detection was applied in combination with pre-processing steps

including noise reduction methods to reduce the information in the images as well as morphological operations to erode and eliminate outlying mouse features.

Contour maps were created of the sample images in attempt to group pixels of a similar colour profile together. Contours were applied to a spherical tumour and a multi-lobed tumour, revealing an imperfect tumour boundary and partial tumour boundary, respectively. This method is inconsistent and would not be applicable to irregular tumour shapes. Gradient methods were applied to the sample image to visualise any patterns or associations between the pixels and this again gave no results, indicating the uniformity of pixels in the results.

As commonly used methods were unsuccessful, the next step was for manual intervention. The pixel value at the tumour boundary of the cropped image created using erosion was determined and used to threshold the image. The resulting image had an imperfect boundary spreading into the mouse body and down the mouse flank clearly highlighting that the colour of the tumour boundary is also present in the mouse body. The border around the tumour was thick and therefore not showing a singular boundary line, boundaries and centroids were applied to the threshold image to extract more information but they were unsuccessful.

Using these image processing methods both in isolation and in combinations were unable to detect the tumour due to the lack of colour differentiation. Other strategies have been reported for detection of tumours such as Lattoofi et al[134] and Sengupta et al[135].

Lattoofi et al used as combination of image processing techniques to identify melanoma skin lesions, producing an algorithm with 93% accuracy[134]. Morphological operations and thresholding were used as preprocessing steps before the ABCD rule was applied. A represents asymmetry where the image was split in half and if both sides were identical, the lesion is classified as benign. Next is B for border, where benign lesions have smooth borders and melanomas have irregularities. C is for the colour variations present in melanomas as D is for dermoscopic structure such as the diameter as melanomas grow faster than benign lesions and therefore have larger diameters. Each factor is scored and combined to give a total dermatoscopy score, which a score higher than 5.45 giving a melanoma classification.

Melanomas already have a set of rules which can classify them, such as irregular border, asymmetrical and colour variation, and then image processing techniques are found to satisfy the rules. This work has no rules and needs to find identifying features in order to capitalise on them. The techniques Lattoofi implemented rely on high visibility of the tumour and the ability to detect border, colour and shape[134]. This is in contrast with this thesis which needs techniques that don't rely on these image processing features due to lack of visibility and a lack of colour differentiation which renders the image processing techniques unsuccessful in identifying the tumour from the surrounding mouse skin.

Sengupta et al used image processing techniques to improve edge detection of the skin lesions neurofibromatosis and nodular melanoma using multiple pre-processing steps and Ant Colony Optimisation (ACO)[135]. ACO is a method inspired by the behaviour of ants, specifically the communication between them

where ants release pheromones to signal information to other ants. The pre-processing steps used by Sengupta included converting the image from RGB to grayscale, enhancing the contrast of the image using 1% of the high and low pixel intensities in the image and median filtering, where every pixel is compared to its neighbouring pixels and given the median value. Next, the edge detectors Sobel, Prewitt and Canny were applied to give the edges of the skin lesions, as an edge map, and ACO was used to enhance the edges found.

The first stage of ACO was initialisation comprised of transforming the edge maps into a matrix and split into 8-pixel grids known as cliques. Next is the pixel transition phase where the ants are able to move along their 8 pixel clique, looking for changes in pixel intensity. The third step is pheromone updating, ensuring pixels the ants have moved locally within their clique and globally, which looks at total ant movement. The process is repeated multiple times. Stop criteria halts the ant movement if it touches a pixel path covered by a neighbouring ant and once all pixels have been covered, the process will stop.

When reviewing the results of this paper, it is evident that the skin lesion already has a colour differentiation between the lesion and the surrounding skin[135]. The pre-processing steps as well as edge detection and ACO were successful as the colour differentiation already existed and these techniques enhanced the differences thus allowing clearer edges. The difficulty in this thesis is the lack of colour differentiation, there is no base for image processing techniques to find and enhance as there is no change in pixel intensity at the tumour boundary. The resulting images from thresholding and the contour maps in this thesis clearly indicate that the tumour boundary is not defined. Although ACO had a positive

impact on the edge maps found by Sengupta[135], it wouldn't work in this research as the edge maps created on the sample images were unsuccessful in locating any tumour features.

Unlike this study, in these studies colour differentiation facilitates tumour detection, therefore other disciplines need to be studied to see if a viable solution to the differentiation issue is available. An area where there is poor differentiation is in the assessment of cracks in concrete. Mohan et al published a critical review allowing a broader view of the industry and the challenges faced as well as solutions to overcome these[136].

Cracks in concrete are warning signs of the possible failure of the structure and so it is important to be able to identify and quantify the damage early[136]. Early methods involved manual inspection which was reliant on the expertise and experience of the observer and so automated methods were developed to reduce bias and increase efficiency and reproducibility. Mohan published a critical review of crack detection methods based on image processing[136]. Although the paper covers various image types including infrared, ultrasonic and laser imaging, to name but a few, the focus for the purposes of this thesis was on digital imaging processing to translate the techniques and methods for tumour detection.

Interestingly, many of the techniques used in crack detection digital image processing are similar to those applied in skin lesion detection, such as morphological approach, median filtering, threshold method and centroid method[136]. This highlighted the simplicity of image processing methods as a colour differentiation must be present for these techniques to be successful.



A study by Liu used image processing techniques to identify bugholes[137]. Bugholes are imperfections in the surface of concrete resulting from trapped air or water thus resulting in an unsatisfactory appearance. Automated processes provide efficiency in identifying and quantifying the bugholes and Liu 2017 used image processing techniques to detect bugholes and therefore grade the quality of the surface of the concrete. The paper firstly converted the images from RGB to grayscale to allow investigation of the tone of the bugholes by looking at surface reflectivity. Bugholes have decreased reflectivity in comparison to normal concrete and therefore this was a distinguishing factor[137].

Next, Liu enhanced the contrast of the image, specifically grey linear enhancement, to increase the difference in pixel intensity between bugholes and the normal concrete surface[137]. The image was then segmented by pixel intensity, with the results showing a two value histogram representing both bugholes and normal concrete. The results were inconclusive as it was difficult to find a single threshold and so an automated thresholding method was applied, the OTSU method. This method iterates through all threshold values, analysing the spread of pixels at each threshold. This method was found to be successful in comparison to edge detection methods[137].

When analysing the images of bugholes, it is clear to observe their location as they are darker in colour in comparison to normal concrete. This paper works on increasing this difference in pixel intensity to allow detection to be easier and were able to do so as there was a distinguishable difference present. The challenge this thesis faced was the lack of distinguishable features as a clear

tumour boundary could not even be found due to the tumour and surrounding skin having the same pixel intensity. Edge detection did provide positive results for Liu but also gave false positives whereas the OTSU method yielded a higher accuracy and therefore was chosen for the detection of bugholes[137]. Edge detection did not give any accurate results in this thesis and so enhancement of contrast would have little effect as the tumour boundaries were incomplete in terms of differences in pixel intensities between the tumour and the surrounding skin.

Another discipline using image processing for object detection is agriculture. The uses of agriculture range from feeding the population to providing materials for textiles as well as medicines, hence it is vital to identify pests and disease on crop plant leaves. These defects are visible by eye and are usually detected by trained experts[138]. However this is expensive and impractical, especially due to the dwindling workforce. Furthermore, trained experts can be specialists in specific areas and therefore have difficulty recognising anomalies for non-native pests and diseases. Ngugi 2021 presents a review highlighting the uses of computing algorithms for automatic disease and pest detection, thus reducing the need for trained observers[138]. Initially, RGB images were utilised as RGB cameras are readily available, easy to operate as well as cost effective. As well as using image processing, machine learning methods are highly utilised in this field due to accuracy in extracting image features facilitating better detection. As was the case with bugholes and previous research discussed, the object being detected has a colour differentiation from its surroundings whereas the challenge this thesis faces is a lack of colour differentiation.

As image processing methods were evidently not sufficient to satisfy the aims of this thesis due to the lack of colour differentiation between the tumour and its surrounding skin, and there do not appear to be solutions to this, even from other disciplines, the next step was to introduce machine learning methods as a way to detect the tumour from the 2D digital images.

Cascade classification in this thesis applied three commonly used object detection methods[139], Haar, HOG and LBP feature type classifiers. Whilst these classifiers started as object detectors in facial recognition, and are still highly used in face recognition (Adeshina 2021[140]), they are now commonly used across various disciplines such as vehicle classification (Sanjana 2021[141]), and detection of kidney abnormalities (Kaur 2020).

Haar features use pattern recognition to classify an object allowing for the detection of structures regardless of the scale, illumination or colouring of the image or object. It consists of grey and white rectangles, focussing on pixel intensity, which are placed around common features of the object. In facial recognition, common features are the eyes, nose, cheeks, these are attributes belonging to the object being detected. Haar features can determine that the nose area is brighter than the eyes, thus aiding the detection process[140].

LBP is a texture descriptor where local changes in pixel intensity show edges and points and can be combined globally to show patterns. Ahonen used LBP in facial recognition and found that different features of the face were weighted differently, such as the eyes, nose and mouth thus allowing them to be distinguished from each other[142].

HOG use gradients vectors for object classification, this looks at the direction of gradients and edges in an image to identify shapes. This is useful in detecting silhouettes and contours[143]. As well established detectors, Haar, HOG and LBP have easy to use frameworks customisable for the given task. They are often utilised in combination with a variety of methodology, for example, HOG is often combined with SVM to provide powerful facial recognition[144].

In this thesis, Haar features were most successful in identifying the tumour albeit with multiple negative classifications. Further testing on cropped versions on the sample image with reduced information again identified the tumour but with multiple negative classifications thus rendering this an unreliable method of tumour detection.

Li 2017 used Haar classifier for facial recognition but combined this with three weaker classifiers to boost detection results[145]. The three weaker classifiers focused on different features of the face. Translating these findings in terms of this thesis, instead of focussing on classifiers to identify the tumour, classifiers could be trained on all features of the mouse including the head and limbs thus allowing these features to be eliminated. Looking at the results of cascade classification on the sample images, it wasn't a case of a mouse limbs being mistaken for tumours, as cascade classification was also run on cropped images that did not include any limbs or features outside the flank of the mouse, and still gave false positives. Therefore, this is not a viable method to improve cascade classification for tumour detection.

In this thesis, pretrained convolutional neural networks VGG-Face, VGG-16 and VGG-19 were tested for tumour detection and VGG-Face with layer Relu6 had the highest recognition rate and therefore were the most successful in detecting the tumour boundary. When tested on larger datasets, the classifier trained with VGG-Face layer Relu6 had a recognition rate of 0.9886. To achieve a 100% recognition rate, the image which did not have the tumour correctly identified was cropped and the classifier was executed again which successfully identified the tumour. The cropping of the image does not introduce an element of observer bias as the cropped image is not exclusively of the tumour, it also includes the immediate surrounding mouse body and so the observer does not need to attempt to decipher the tumour boundary. Only images with false detection need to be cropped, with a 99% recognition rate, this leaves 1% of images which may require intervention. Whilst a significantly high tumour detection rate without requiring user intervention was achieved with this method, it has to be remembered that the competing calliper technology offers a 100% detection rate for the tumour outline, and therefore further optimisation needs to be considered.

Where this thesis focussed on image processing and machine learning as two separate entities, Khan 2019 combined the two thus optimising the process[146]. This study achieved classification accuracy rates ranging between 85-96.5% through implementing a hybrid model incorporating both segmentation and classification[146]. This work allowed Khan to go beyond classifying skin lesions as melanoma to differentiating between benign and malignant melanomas. The image processing techniques initially applied to the image of the melanoma include contrast enhancement, weighted Gaussian and median filtering. The resulting image was segmented in the YCbCr colour space and a threshold was

applied resulting in a binary image. The next step was to extract features for classification, using a DCNN-9 model which is based on VGG and AlexNet models[146].

In this thesis. Images of tumours and pseudo-tumours were used to train the CNN classifiers. Contrast was altered on images to give a larger sample size of images but no attempts were made to segment the images initially as the image processing methodology applied initially was unsuccessful. To include processed images in the sample of images used to train the classifier, they would have to be included in place of an image of a whole tumour as the maximum number of images used to train a classifier was restricted. The biggest limiting factor for CNN in this thesis was the lack of computational power thus the classifier could only be trained on a maximum number of images. This work was able to identify 99% of tumour boundaries (0.989 recognition rate), even though the classifiers were trained on a limited number of images. It is highly probable that with increased computational power and a larger sample of images, the recognition rate could improve further.

With a high recognition rate, and a method to increase this to 100%, the first aim of this thesis to detect the tumour was satisfied and the next aim of this thesis was to calculate tumour volume. With the first thesis aim to detect the tumour boundaries achieved, the next aim was to automate calculation of tumour volume. As reviewed in the introduction section, calculation of tumour volume from 2D calliper measurements has several issues in terms of accuracy and reproducibility. This is why the gold standard formula provided by Tomayko finds ellipsoidal volume without needing the height measurement.

Feldman 2009 looked at the relationship between length, width and height in an attempt to improve on the current gold standard and found  $(3 \times \text{height}^2)$  to have similar results to  $(\text{length} \times \text{width})$ . They provided an alternative formula to the current gold standard, reducing the overall error from 10% to 4%. Although this work potentially reduces the error found using the standard methods, it still requires calliper measurements. The work reduced observer variation by ensuring one observer carried out all measurements but this is not feasible in real life scenarios. Furthermore, this study transplanted tumours on the leg instead of the mouse flank, therefore there was minimal interference from surrounding skin. Again, this is not feasible in real life scenarios and growing tumours on a leg for longer studies would be detrimental to the health of the mouse as this would affect mobility.

In the first attempt to determine tumour volume in this thesis using computing and bypassing callipers, the maximum length of the tumour and perpendicular width of tumours were automatically determined by a computer algorithm thus removing physical contact on the tumour as well as observer bias. The computer algorithm processed a mask of the tumour, which was obtained through manual separation of the tumour from the remainder of the digital image at the tumour boundary as decided by the observer, and found the longest length of pixels to represent the maximum length and its perpendicular width. The modified ellipsoid formula was used to determine the tumour volume thus enabling a direct comparison of results obtained from using digital and calliper methods. A positive correlation (0.789) was found between tumour volume determined from digital methods and callipers. Smaller volumes under  $400\text{mm}^3$  had a strong correlation (0.868) but as

volume increased the ability to correlate was lost as volumes above 600m<sup>3</sup> had a weak correlation (0.081).

Looking singularly at length and width measurements, length obtained from both digital and calliper methods had a higher correlation of 0.833 whereas width had a lower correlation 0.750 thus suggesting that observers operating callipers were better at determining the maximum length of a tumour in comparison to its perpendicular width. As it is a manual process for the observer, determining a perpendicular value whilst holding a moving animal explains the weakness in the width measurement. Interestingly, comparing the area of the digital tumour mask with the calliper volume and digital volume gave correlations of 0.838 and 0.976, respectively. The results show a relationship between the number of pixels in the tumour mask and the tumour volume, suggesting that the volume obtained by callipers could be reproduced by studying the area of the tumour and thus avoiding length and width measurements.

These results were found through comparing volumes found from digital masks of the tumour with volumes found from callipers. As discussed in the introduction chapter, callipers are themselves flawed. To gain more insight into the reliability of using masks as well as callipers, pseudo-tumours were constructed to enable testing calliper measurements, mask measurements and a third measurement of volume found from 3D scanning, believed to be closer to true volume.

Calliper volume and mask volume from the pseudo-tumours had a high correlation of 0.952. Calliper had a very high correlation of (0.952) with 3D volume and masks had a lower (0.929) correlation. The methodology to use mask volume



has been unsuccessful in superseding volume results found from callipers. However, taking images and calliper measurements of a pseudo-tumour constructed with blu-tack is far removed from reality. For the pseudo-tumours, it was easy to take calliper measurements as the tumours were not sitting subcutaneously on a moving animal, and therefore there was no pressure of time and a surgical environment. Furthermore, as the pseudo-tumours were constructed from blu-tack, applying callipers with too much strength would cut into the model and reveal a dent, and so it was easy to judge when to stop squeezing, which would not have been as easy on a moving animal.

Additionally, one of the flaws of callipers is the observer variability. This digital method of obtaining longitudinal measurements removes observer bias in the maximum length and width measurement but these measurements are derived from a digital mask created by an observer.

When assessing MRI scans, the observer must take two-dimensional measurements to define the area on the MRI showing tumour lesions. As with calliper measurements, this introduces inaccuracies and are not always reliable. Kickingreder 2019 automated this process using artificial neural networks, teaching the software to find the tumour area without any observer involvement[147]. The observers manually created masks of the MRI images, delineating the lesions on the image and removing all interfering background information and used these to train the artificial neural network.

Kickingreder found that the reliability values of the software and standard practice were 87% and 51%, respectively. This highlights that the masks were

successful at training the neural network. The study ensured reliability of their masks by ensuring all masks were created by experienced radiologists and then checked again by an experienced neuroradiologist[147].

Although the same observer created all of the masks in this thesis, this method is not realistic for future lab work involving multiple observers who may not be familiar with the tumours. This has potential to introduce error as the observer must visually find the tumour volume and extract the mask accordingly. Evidently, this method is flawed and perhaps too simple to contribute towards determining tumour volume from a 2D digital image. The next method to find tumour volume was shape from shading, where the link between pixel intensity and depth was investigated.

Ciaccio[123] used shape from shading to capture images of the small intestine focusing on detection of protruding areas in order to diagnose coeliac disease. The study found a direct relationship between reflectance map and image brightness. Brighter pixels were closer to the camera in comparison to darker pixels which were further away, giving an indication of depth. This work incorporated shade from shading methods through finding a link between pixel intensity and pseudo-tumour height. It was expected that a taller tumour with greater depth have a higher pixel intensity in comparison to a shorter tumour with decreased depth. Seven pseudo-tumours were imaged both with and without using the built-in flash light source on the camera. Those without flash had no external light source and the pixel intensity remained almost linear and consistent as the pseudo-tumour height increased indicating no relationship between pixel intensity and model depth. Adding flash resulted in varying pixel intensity, with

the pixel intensity dropping between 10 and 15mm and then increasing between 15-25mm models.

To obtain more concise results, virtual tumours were constructed. These allowed the applications of a variety of different light angles and tumour shapes whilst removing human error as the process was computerised. Pixel intensity was then investigated against the volume of the tumour itself instead of its height. Pixel intensity and volume were regressed together. The ideal error would be 0, and the errors found were very high using both PCA and multiple regression methods, within the range of 277 – 299 without offset, and within the range of 312 – 314 with offset. Furthermore, Figure 3.30 in this thesis clearly shows that increasing volume measurements caused a big increase in the error found.

Evidently, this work found no connection between pixel intensity and depth. Looking at the sample of images of s.c. tumours used in this thesis, earlier work for tumour detection highlighted that the images had no distinguishable features and so image processing techniques performed poorly. To find tumour volume, neither longitudinal measurements nor the addition of a light source to use shape from shading principles were individually successful in determining the tumour. This indicates the lack of colour differentiation and change in pixel intensity at the tumour boundary in 2D digital images.

A paper by Delgado-San Martin 2019 presents a method to obtain tumour measurements from an image with a 3D scanning technique, BioVolume™ (Fuel3D, Oxford, UK; [www.fuel3d.com](http://www.fuel3d.com)) incorporating both digital and infrared imaging[148]. The BioVolume device contains two RGB cameras, three white

light flashes and an infrared camera. Each image capture contains six RGB images and one infrared image. The RGB images are reconstructed on a surface while the infrared image, which indicates the regions containing tumour due to poor vascularisation, is mapped onto the surface. As the work in this thesis has shown, digital images are unable to provide accurate segmentation results and so this study purposed the thermal image for segmentation to ensure only the tumour was mapped onto the 3D reconstruction. Instead of using the gold standard modified ellipsoid formula, the study used the spheroid volume to compare tumour volumes obtained through BioVolume and callipers and also used BioVolume to determine a cylindrical volume.

When comparing length and width measurements obtained from both BioVolume and calliper measurements, the BioVolume measurements were within 3mm of calliper measurements in 88.86% and 90.99% of measurements, respectively. Furthermore, there was a high correlation between tumour volume measurements found via callipers and BioVolume although BioVolume estimated lower volumes than the callipers overall. Interestingly, using the spheroid formula for volume, BioVolume had lower rates of inter-operator variability but when the cylindrical volume was determined, BioVolume had similar rates of inter-operator variability to the callipers. When capturing BioVolume measurements, the back of the mouse is held up to the BioVolume device for imaging purposes. The position the mouse is held in may vary between observers.

Although this paper found a high correlation between their method to calculate tumour volume, BioVolume, and calliper volume, the work is only successful if their specific equipment is purchased and utilised, deviating away from the

principle of using low-cost standard imaging equipment as aimed for in this thesis. Furthermore, the mouse must be held against the equipment which negatively impacts the mice as they are not only being removed from their environment, they are also held alongside a machine they are unfamiliar with which can be stressful.

The aim of this thesis was to both detect the tumour and determine tumour volume from a 2D image, thus allowing a cheap and efficient method to improve accuracy in comparison to calliper measurements whilst providing refinement benefits for the mouse. With previous methods of image processing and shape from shading failing, machine learning methods were used to teach classifiers to find an association between the xenograft images and their corresponding volume measurements. Multiple datasets were used to find a classifier with the smallest error.

When using machine learning in tumour detection, it was found that combining pseudo tumour images with xenograft images improved tumour recognition rates. Whereas when training classifiers for volume measurements, using pseudo-tumour measurements caused much higher errors than using xenografts themselves. Therefore, the number of xenograft images we had access to became a limiting factor. Two datasets consisted of only xenograft images, one with 2977 images and another with 3499 images, giving errors of 0.1135 and 0.1081 respectively. The biggest limiting factor for CNN in this thesis was the lack of computational power thus the classifier could only train on a maximum number of images. An increased number of xenograft images would have resulted in a reduced error.

Finally, the classifier was tested on original xenograft images, and the real volume was plotted with estimated volume obtained from the callipers. It was found that the two samples came from the same distribution and therefore there was no significant difference between the two. This work has successfully found tumour volume measurements using 2D digital images. Although this provides refinement benefits to the mice as it avoids the squeezing of the tumour by the callipers, the mouse must still be held in order to capture images of the tumour and the angling of the mouse could introduce bias. The next step in this work would be to a home cage set up for the mouse where images could be captured from within the cage, eliminating the need for the mouse to be held. This would also ensure the mouse would stay in its own environment and not suffer the stress of being transferred out of the cage into a surgical room setup.

#### **4.1 Conclusion & next steps**

This thesis set out to replace calliper use for tumour measurement with a non-invasive programming solution applied to 2D digital images, to not only eliminate the inaccuracies found with callipers, but to reduce physical contact with the tumour, thus offering refinement benefits to the mouse. This would also be a relatively cheap methodology which could be adopted routinely in most animal facilities and require no specialist training. Commercially available approaches which carry out similar measurements have become available since this work began, e.g. BioVolume[148] which uses a combination of digital imaging and infrared, but this requires expensive machinery and increases animal handling.

This novel work has successfully achieved this. The attempts at using image processing techniques to find the tumour boundary quickly indicated that the biggest challenge in this work was a lack of colour differentiation between the tumour and surrounding mouse body. Research was consulted from edge detection in melanomas to cracks in concrete and pests in agriculture but none offered a solution to solve this challenge, and thus a different approach, machine learning, was taken.

Machine learning methods were utilised and pre-trained convolutional neural networks were able to decipher a tumour volume from a digital image with 99% accuracy and a semi-automatic step was introduced to achieve a 100% accuracy level. There is a clear pathway for future work as this work was limited by computational power. A larger dataset would increase accuracy levels further.

Furthermore, although this work removes physical contact on the tumour, it still requires animal handling. Therefore the next refinement to progressing this work would be that instead of having to restrain the mouse to capture its image, cameras would be placed around the home cage so the mouse could remain in its own environment. Either multiple images could be captured from multiple angles, or a video could be captured where the mouse and the tumour would be tracked, and the tumour subsequently measured. The initial groundwork achieved through this thesis will facilitate further development to progress towards these key refinements in experimental tumour graft measurement.





## 5. References

- [1] Cancer Research UK, “Cancer Statistics for the UK.” [Online]. Available: <https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk>. [Accessed: 04-Jan-2017].
- [2] Cancer Research UK, “Cancer risk statistics | Cancer Research UK.” [Online]. Available: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/risk>. [Accessed: 22-Feb-2023].
- [3] D. Hanahan and R. A. Weinberg, “Hallmarks of cancer: the next generation,” *Cell*, vol. 144, no. 5, pp. 646–674, 2011.
- [4] A. S. Ahmad, N. Ormiston-Smith, and P. D. Sasieni, “Trends in the lifetime risk of developing cancer in Great Britain: comparison of risk for those born from 1930 to 1960.,” *Br. J. Cancer*, vol. 112, no. 5, pp. 943–7, Mar. 2015.
- [5] A. Urruticoechea, R. Alemany, J. Balart, A. Villanueva, F. Vinals, and G. Capella, “Recent advances in cancer therapy: an overview,” *Curr. Pharm. Des.*, vol. 16, no. 1, pp. 3–10, 2010.
- [6] J. M. Jessup, L. S. McGinnis, G. D. Steele, H. R. Menck, and D. P. Winchester, “The National Cancer Data Base. Report on colon cancer.,” *Cancer*, vol. 78, no. 4, pp. 918–26, Aug. 1996.
- [7] R. W. Beart, G. D. Steele, H. R. Menck, J. S. Chmiel, K. E. Ocwieja, and D. P. Winchester, “Management and survival of patients with adenocarcinoma of the colon and rectum: a national survey of the Commission on Cancer.,” *J. Am. Coll. Surg.*, vol. 181, no. 3, pp. 225–36, Sep. 1995.
- [8] J. Figueras *et al.*, “Surgical resection of colorectal liver metastases in patients with expanded indications: a single-center experience with 501

- patients.," *Dis. Colon Rectum*, vol. 50, no. 4, pp. 478–88, Apr. 2007.
- [9] J. S. D. Mieog *et al.*, "Fundamentals and developments in fluorescence-guided cancer surgery," *Nat. Rev. Clin. Oncol.*, vol. 19, no. 1, pp. 9–22, 2022.
- [10] G. Delaney, S. Jacob, C. Featherstone, and M. Barton, "The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines.," *Cancer*, vol. 104, no. 6, pp. 1129–37, Sep. 2005.
- [11] E. Kapiteijn *et al.*, "Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer.," *N. Engl. J. Med.*, vol. 345, no. 9, pp. 638–46, Aug. 2001.
- [12] K. A. Vallis and I. F. Tannock, "Postoperative Radiotherapy for Breast Cancer: Growing Evidence for an Impact on Survival," *JNCI J. Natl. Cancer Inst.*, vol. 96, no. 2, pp. 88–89, Jan. 2004.
- [13] U. Veronesi *et al.*, "A preliminary report of intraoperative radiotherapy (IORT) in limited-stage breast cancers that are conservatively treated," *Eur. J. Cancer*, vol. 37, no. 17, pp. 2178–2183, 2001.
- [14] D. K. Bahn *et al.*, "Salvage Cryosurgery for Recurrent Prostate Cancer After Radiation Therapy: A Seven-Year Follow-up," *Clin. Prostate Cancer*, vol. 2, no. 2, pp. 111–114, Sep. 2003.
- [15] Z. Zhang, X. Liu, D. Chen, and J. Yu, "Radiotherapy combined with immunotherapy: the dawn of cancer treatment," *Signal Transduct. Target. Ther.*, vol. 7, no. 1, p. 258, 2022.
- [16] B. A. Chabner and T. G. Roberts, "Chemotherapy and the war on cancer," *Nat. Rev. Cancer*, vol. 5, no. 1, pp. 65–72, 2005.
- [17] T. L. Jackson, "Intracellular accumulation and mechanism of action of

- doxorubicin in a spatio-temporal tumor model.," *J. Theor. Biol.*, vol. 220, no. 2, pp. 201–213, Jan. 2003.
- [18] E. K. Rowinsky, E. A. Eisenhauer, V. Chaudhry, S. G. Arbuck, and R. C. Donehower, "Clinical toxicities encountered with paclitaxel (Taxol).," *Semin. Oncol.*, vol. 20, no. 4 Suppl 3, pp. 1–15, Aug. 1993.
- [19] G. W. Sledge, "Phase III Trial of Doxorubicin, Paclitaxel, and the Combination of Doxorubicin and Paclitaxel as Front-Line Chemotherapy for Metastatic Breast Cancer: An Intergroup Trial (E1193)," *J. Clin. Oncol.*, vol. 21, no. 4, pp. 588–592, Feb. 2003.
- [20] T. Helleday, E. Petermann, C. Lundin, B. Hodgson, and R. A. Sharma, "DNA repair pathways as targets for cancer therapy.," *Nat. Rev. Cancer*, vol. 8, no. 3, pp. 193–204, Mar. 2008.
- [21] R. M. Elledge and C. K. Osborne, "Oestrogen receptors and breast cancer," *BMJ*, vol. 314, no. 7098, pp. 1843–1843, Jun. 1997.
- [22] "Tamoxifen for early breast cancer.," *Cochrane database Syst. Rev.*, no. 1, p. CD000486, Jan. 2001.
- [23] I. E. Smith and M. Dowsett, "Aromatase inhibitors in breast cancer.," *N. Engl. J. Med.*, vol. 348, no. 24, pp. 2431–42, Jun. 2003.
- [24] W. A. See *et al.*, "Bicalutamide as Immediate Therapy Either Alone or as Adjuvant to Standard Care of Patients with Localized or Locally Advanced Prostate Cancer: First Analysis of the Early Prostate Cancer Program," *J. Urol.*, vol. 168, no. 2, pp. 429–435, Aug. 2002.
- [25] L. M. Weiner, R. Surana, and S. Wang, "Monoclonal antibodies: versatile platforms for cancer immunotherapy.," *Nat. Rev. Immunol.*, vol. 10, no. 5, pp. 317–27, May 2010.
- [26] Y. Hu *et al.*, "Investigation of the mechanism of action of alemtuzumab in a

- human CD52 transgenic mouse model.,” *Immunology*, vol. 128, no. 2, pp. 260–70, Oct. 2009.
- [27] L. Kraehenbuehl, C.-H. Weng, S. Eghbali, J. D. Wolchok, and T. Merghoub, “Enhancing immunotherapy in cancer by targeting emerging immunomodulatory pathways,” *Nat. Rev. Clin. Oncol.*, vol. 19, no. 1, pp. 37–50, 2022.
- [28] W. N. Hait, “Targeted cancer therapeutics,” *Cancer Res.*, vol. 69, no. 4, pp. 1263–1267, 2009.
- [29] G. Daley, R. Van Etten, and D. Baltimore, “Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome,” *Science (80-. )*, vol. 247, no. 4944, pp. 824–830, Feb. 1990.
- [30] D. Cortez, L. Kadlec, and A. M. Pendergast, “Structural and signaling requirements for BCR-ABL-mediated transformation and inhibition of apoptosis.,” *Mol. Cell. Biol.*, vol. 15, no. 10, pp. 5531–41, Oct. 1995.
- [31] L. A. Kujawski and M. Talpaz, “The role of interferon-alpha in the treatment of chronic myeloid leukemia.,” *Cytokine Growth Factor Rev.*, vol. 18, no. 5–6, pp. 459–71, Jan. .
- [32] S. O’Brien *et al.*, “Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia,” *N. Engl. J. Med.*, vol. 348, no. 11, pp. 994–1004, Mar. 2003.
- [33] Z. Fu, S. Li, S. Han, C. Shi, and Y. Zhang, “Antibody drug conjugate: the ‘biological missile’ for targeted cancer therapy,” *Signal Transduct. Target. Ther.*, vol. 7, no. 1, p. 93, 2022.
- [34] Bio-ITWorld.com, “FDA Approved ADC Drugs list up to 2022.” [Online]. Available: <https://www.bio-itworld.com/pressreleases/2022/11/28/fda-approved-adc-drugs-list-up-to-2022>. [Accessed: 28-Feb-2023].

- [35] A. Kamb, S. Wee, and C. Lengauer, "Why is cancer drug discovery so difficult?," *Nat. Rev. Drug Discov.*, vol. 6, no. 2, pp. 115–120, 2006.
- [36] R. Fisher, L. Pusztai, and C. Swanton, "Cancer heterogeneity: implications for targeted therapeutics," *Br. J. Cancer*, vol. 108, no. 3, pp. 479–485, Feb. 2013.
- [37] C. L. Vogel, "Efficacy and Safety of Trastuzumab as a Single Agent in First-Line Treatment of HER2-Overexpressing Metastatic Breast Cancer," *J. Clin. Oncol.*, vol. 20, no. 3, pp. 719–726, Feb. 2002.
- [38] G. R. Oxnard, "The cellular origins of drug resistance in cancer," *Nat. Med.*, vol. 22, no. 3, pp. 232–234, Mar. 2016.
- [39] M. M. Gottesman, "Mechanisms of Cancer Drug Resistance," *Annu. Rev. Med.*, vol. 53, no. 1, pp. 615–627, Feb. 2002.
- [40] C. Holohan, S. Van Schaeybroeck, D. B. Longley, and P. G. Johnston, "Cancer drug resistance: an evolving paradigm," *Nat. Rev. Cancer*, vol. 13, no. 10, pp. 714–726, Sep. 2013.
- [41] A. F. Gazdar, "Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors," *Oncogene*, vol. 28, pp. S24–S31, Aug. 2009.
- [42] P. A. Jänne *et al.*, "AZD9291 in EGFR Inhibitor-Resistant Non-Small-Cell Lung Cancer," *N. Engl. J. Med.*, vol. 372, no. 18, pp. 1689–1699, Apr. 2015.
- [43] N. E. Davidson *et al.*, "AACR Cancer Progress Report 2016," *Clin. Cancer Res.*, vol. 22, no. 19\_Supplement, pp. S1–S137, Oct. 2016.
- [44] T. Klabunde and A. Evers, "GPCR Antitarget Modeling: Pharmacophore Models for Biogenic Amine Binding GPCRs to Avoid GPCR-Mediated Side Effects," *ChemBioChem*, vol. 6, no. 5, pp. 876–889, May 2005.
- [45] "ChEMBL." [Online]. Available: <https://www.ebi.ac.uk/chembl/>. [Accessed:

- 20-Mar-2017].
- [46] A. Koutsoukas *et al.*, “From in silico target prediction to multi-target drug design: Current databases, methods and applications,” *J. Proteomics*, vol. 74, no. 12, pp. 2554–2574, Nov. 2011.
- [47] J. Mestres *et al.*, “The topology of drug–target interaction networks: implicit dependence on drug properties and target families,” *Mol. Biosyst.*, vol. 5, no. 9, p. 1051, 2009.
- [48] T. Voskoglou-Nomikos, J. L. Pater, and L. Seymour, “Clinical predictive value of the in vitro cell line, human xenograft, and mouse allograft preclinical cancer models,” *Clin. Cancer Res.*, vol. 9, no. 11, pp. 4227–4239, 2003.
- [49] J. P. Hughes, S. Rees, S. B. Kalindjian, and K. L. Philpott, “Principles of early drug discovery,” *Br. J. Pharmacol.*, vol. 162, no. 6, pp. 1239–49, Mar. 2011.
- [50] R. M. Phillips, M. C. Bibby, and J. A. Double, “A critical appraisal of the predictive value of in vitro chemosensitivity assays,” *J. Natl. Cancer Inst.*, vol. 82, no. 18, pp. 1457–1468, 1990.
- [51] W. M. S. Russell and R. L. Burch, *The principles of humane experimental technique*. Methuen, 1959.
- [52] L. U. Sneddon, L. G. Halsey, and N. R. Bury, “Considering aspects of the 3Rs principles within experimental animal biology,” *J. Exp. Biol.*, vol. 220, no. Pt 17, pp. 3007–3016, Sep. 2017.
- [53] NC3Rs, “The 3Rs.” [Online]. Available: <https://www.nc3rs.org.uk/the-3rs>. [Accessed: 26-Mar-2018].
- [54] J. Richmond, “Refinement, Reduction, and Replacement of Animal Use for Regulatory Testing: Future Improvements and Implementation Within the

- Regulatory Framework," *ILAR J*, vol. 43, no. Suppl\_1, pp. S63-68, Jan. 2002.
- [55] M. Festing, "Doing better animal experiments; together with notes on genetic nomenclature of laboratory animals.," vol. 13, no. 3, p. Insert 1-Insert 8, 2000.
- [56] T. Poole, "Happy animals make good science," *Lab. Anim.*, vol. 31, no. 2, pp. 116–124, Apr. 1997.
- [57] J. L. Hurst and R. S. West, "Taming anxiety in laboratory mice," *Nat. Methods*, vol. 7, no. 10, pp. 825–826, Oct. 2010.
- [58] P. McGonigle and B. Ruggeri, "Animal models of human disease: Challenges in enabling translation," *Biochem. Pharmacol.*, vol. 87, no. 1, pp. 162–171, 2014.
- [59] H. Office, "Animal research and testing," 2015. [Online]. Available: <https://www.gov.uk/guidance/research-and-testing-using-animals#publications>. [Accessed: 14-Sep-2015].
- [60] K. Strömberg, V. Nordin, M. Miller, B. Akerström, and C. Gillberg, "AUTISM IN THALIDOMIDE EMBRYOPATHY: A POPULATION STUDY," *Dev. Med. Child Neurol.*, vol. 36, no. 4, pp. 351–356, Nov. 2008.
- [61] N. Chernoff, J. M. Rogers, and R. J. Kavlock, "An overview of maternal toxicity and prenatal development: considerations for developmental toxicity hazard assessments," *Toxicology*, vol. 59, no. 2, pp. 111–125, Dec. 1989.
- [62] L. R. Kelland, "Of mice and men: values and liabilities of the athymic nude mouse model in anticancer drug development," *Eur J Cancer*, vol. 40, no. 6, pp. 827–836, 2004.
- [63] B. A. Ruggeri, F. Camp, and S. Miknyoczki, "Animal models of disease:

- Pre-clinical animal models of cancer and their applications and utility in drug discovery,” *Biochem. Pharmacol.*, vol. 87, no. 1, pp. 150–161, 2014.
- [64] T. Van Dyke and T. Jacks, “Cancer modeling in the modern era: progress and challenges,” *Cell*, vol. 108, no. 2, pp. 135–144, 2002.
- [65] M. Sho *et al.*, “Requirements for induction and maintenance of peripheral tolerance in stringent allograft models,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 37, pp. 13230–5, Sep. 2005.
- [66] E. A. Sausville and A. M. Burger, “Contributions of human tumor xenografts to anticancer drug development,” *Cancer Res.*, vol. 66, no. 7, pp. 3351–3354, 2006.
- [67] J. J. Tentler *et al.*, “Patient-derived tumour xenografts as models for oncology drug development,” *Nat. Rev. Clin. Oncol.*, vol. 9, no. 6, pp. 338–350, 2012.
- [68] M. C. Villarroel *et al.*, “Personalizing cancer treatment in the age of global genomic analyses: PALB2 gene mutations and the response to DNA damaging agents in pancreatic cancer,” *Mol. Cancer Ther.*, vol. 10, no. 1, pp. 3–8, Jan. 2011.
- [69] H.-Y. Shin *et al.*, “Identification of prognostic markers of gynecologic cancers utilizing patient-derived xenograft mouse models,” *Cancers (Basel)*, vol. 14, no. 3, p. 829, 2022.
- [70] M. C. Bibby, “Orthotopic models of cancer for preclinical drug evaluation: advantages and disadvantages,” *Eur. J. Cancer*, vol. 40, no. 6, pp. 852–857, 2004.
- [71] A. Rahman and M. I. Choudhary, *Frontiers in anti-cancer drug discovery. Volume 4.* .
- [72] J. Linxweiler *et al.*, “Primary tumor resection decelerates disease



- progression in an orthotopic mouse model of metastatic prostate cancer,” *Cancers (Basel)*, vol. 14, no. 3, p. 737, 2022.
- [73] M. Singh and L. Johnson, “Using genetically engineered mouse models of cancer to aid drug development: an industry perspective,” *Clin. Cancer Res.*, vol. 12, no. 18, pp. 5312–5328, 2006.
- [74] V. Fendrich *et al.*, “The angiotensin-I-converting enzyme inhibitor enalapril and aspirin delay progression of pancreatic intraepithelial neoplasia and cancer formation in a genetically engineered mouse model of pancreatic cancer,” *Gut*, vol. 59, no. 5, pp. 630–637, May 2010.
- [75] I. F. Gilson Sena *et al.*, “Identification of early biomarkers in saliva in genetically engineered mouse model C(3)1-TAg of breast cancer,” *Sci. Rep.*, vol. 12, no. 1, p. 11544, 2022.
- [76] O. J. Becher and E. C. Holland, “Genetically Engineered Models Have Advantages over Xenografts for Preclinical Studies,” *Cancer Res.*, vol. 66, no. 7, pp. 3355–3359, Apr. 2006.
- [77] J. Sapi, L. Kovacs, D. A. Drexler, P. Kocsis, D. Gajari, and Z. Sapi, “Tumor Volume Estimation and Quasi-Continuous Administration for Most Effective Bevacizumab Therapy,” *PLoS One*, vol. 10, no. 11, p. e0142190, Nov. 2015.
- [78] M. M. Tomayko and C. P. Reynolds, “Determination of subcutaneous tumor size in athymic (nude) mice.,” *Cancer Chemother. Pharmacol.*, vol. 24, no. 3, pp. 148–54, Jan. 1989.
- [79] D. M. Euhus, C. Hudd, M. C. LaRegina, and F. E. Johnson, “Tumor measurement in the nude mouse.,” *J. Surg. Oncol.*, vol. 31, no. 4, pp. 229–34, Apr. 1986.
- [80] V. Kersemans, B. Cornelissen, P. D. Allen, J. S. Beech, and S. C. Smart,

- “Subcutaneous tumor volume measurement in the awake, manually restrained mouse using MRI,” *J. Magn. Reson. Imaging*, vol. 37, no. 6, pp. 1499–504, Jun. 2013.
- [81] M. M. Jensen, J. T. Jørgensen, T. Binderup, and A. Kjær, “Tumor volume in subcutaneous mouse xenografts measured by microCT is more accurate and reproducible than determined by <sup>18</sup>F-FDG-microPET or external caliper,” *BMC Med. Imaging*, vol. 8, no. 1, p. 16, 2008.
- [82] S. R. Stock, *MicroComputed Tomography: Methodology and Applications*. CRC Press, 2008.
- [83] D. W. Holdsworth and M. M. Thornton, “Micro-CT in small animal and specimen imaging,” *Trends Biotechnol.*, vol. 20, no. 8, pp. S34–S39, Aug. 2002.
- [84] A. C. O’Farrell, S. D. Shnyder, G. Marston, P. L. Coletta, and J. H. Gill, “Non-invasive molecular imaging for preclinical cancer therapeutic development,” *Br. J. Pharmacol.*, vol. 169, no. 4, pp. 719–735, 2013.
- [85] L. Abou-Elkacem, F. Gremse, S. Barth, R. M. Hoffman, F. Kiessling, and W. Lederle, “Comparison of  $\mu$ CT, MRI and optical reflectance imaging for assessing the growth of GFP/RFP-expressing tumors.,” *Anticancer Res.*, vol. 31, no. 9, pp. 2907–13, Sep. 2011.
- [86] R. Boellaard *et al.*, “FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0.,” *Eur. J. Nucl. Med. Mol. Imaging*, vol. 37, no. 1, pp. 181–200, Jan. 2010.
- [87] S. Rodríguez-Enríquez, A. Marín-Hernández, J. C. Gallardo-Pérez, and R. Moreno-Sánchez, “Kinetics of transport and phosphorylation of glucose in cancer cells.,” *J. Cell. Physiol.*, vol. 221, no. 3, pp. 552–9, Dec. 2009.
- [88] C. Schiepers *et al.*, “Contribution of PET in the diagnosis of recurrent

- colorectal cancer: comparison with conventional imaging," *Eur. J. Surg. Oncol.*, vol. 21, no. 5, pp. 517–522, Oct. 1995.
- [89] P. D. Shreve, Y. Anzai, and R. L. Wahl, "Pitfalls in oncologic diagnosis with FDG PET imaging: physiologic and benign variants.," *Radiographics*, vol. 19, no. 1, pp. 61–77; quiz 150–1, Jan. 1999.
- [90] M. Montelius, M. Ljungberg, M. Horn, and E. Forssell-Aronsson, "Tumour size measurement in a mouse model using high resolution MRI.," *BMC Med. Imaging*, vol. 12, no. 1, p. 12, Jan. 2012.
- [91] A. Berger, "Magnetic resonance imaging.," *BMJ*, vol. 324, no. 7328, p. 35, Jan. 2002.
- [92] R. W. Brown, Y.-C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan, *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley, 2014.
- [93] J. Dazai, N. A. Bock, B. J. Nieman, L. M. Davidson, R. M. Henkelman, and X. J. Chen, "Multiple mouse biological loading and monitoring system for MRI.," *Magn. Reson. Med.*, vol. 52, no. 4, pp. 709–15, Oct. 2004.
- [94] G. D. Ayers *et al.*, "Volume of Preclinical Xenograft Tumors Is More Accurately Assessed by Ultrasound Imaging Than Manual Caliper Measurements," *J. Ultrasound Med.*, vol. 29, no. 6, pp. 891–901, 2010.
- [95] R. Pflanzner *et al.*, "Advanced 3D-Sonographic Imaging as a Precise Technique to Evaluate Tumor Volume," *Transl. Oncol.*, vol. 7, no. 6, pp. 681–686, 2014.
- [96] K. Lloyd Williams, F. J. Lloyd Williams, and R. S. Handley, "INFRA-RED THERMOMETRY IN THE DIAGNOSIS OF BREAST DISEASE," *Lancet*, vol. 278, no. 7217, pp. 1378–1381, Dec. 1961.
- [97] M. P. Cunningham, "The Breast Cancer Detection Demonstration Project

- 25 years later,” *CA. Cancer J. Clin.*, vol. 47, no. 3, pp. 131–133, May 1997.
- [98] E.-K. Ng, “A review of thermography as promising non-invasive detection modality for breast tumor,” *Int. J. Therm. Sci.*, vol. 48, no. 5, pp. 849–859, 2009.
- [99] J. F. Head, F. Wang, C. A. Lipari, and R. L. Elliott, “The important role of infrared imaging in breast cancer,” *IEEE Eng. Med. Biol. Mag.*, vol. 19, no. 3, pp. 52–57, 2000.
- [100] C. Song *et al.*, “Thermographic assessment of tumor growth in mouse xenografts,” *Int. J. cancer*, vol. 121, no. 5, pp. 1055–1058, 2007.
- [101] W. Xie, P. McCahon, K. Jakobsen, and C. Parish, “Evaluation of the ability of digital infrared imaging to detect vascular changes in experimental animal tumours,” *Int. J. cancer*, vol. 108, no. 5, pp. 790–794, 2004.
- [102] N. Hussain *et al.*, “The use of thermographic imaging to evaluate therapeutic response in human tumour xenograft models,” *Sci. Rep.*, vol. 6, p. 31136, Aug. 2016.
- [103] J. M. Chaves-González, M. A. Vega-Rodríguez, J. A. Gómez-Pulido, and J. M. Sánchez-Pérez, “Detecting skin in face recognition systems: A colour spaces study,” *Digit. Signal Process.*, vol. 20, no. 3, pp. 806–823, May 2010.
- [104] J. C. Whitaker and J. C., *Electronic displays: technology, design, and applications*. McGraw-Hill, 1994.
- [105] S. Arslan, E. Ozyurek, and C. Gunduz-Demir, “A color and shape based algorithm for segmentation of white blood cells in peripheral blood and bone marrow images,” *Cytom. Part A*, vol. 85, no. 6, pp. 480–490, Jun. 2014.
- [106] J. C. Russ, *The Image processing handbook*. .
- [107] D. Ziou and S. Tabbone, “Edge Detection Techniques - An Overview.”

- [108] R. Jain, R. Kasturi, and B. G. Schunck, *Machine vision*. McGraw-Hill, 1995.
- [109] G. Ravivarma, K. Gavaskar, D. Malathi, K. G. Asha, B. Ashok, and S. Aarthi, "Implementation of Sobel operator based image edge detection on FPGA," *Mater. Today Proc.*, vol. 45, pp. 2401–2407, Jan. 2021.
- [110] Erwin and T. Yuningsih, "Detection of Blood Vessels in Optic Disc with Maximum Principal Curvature and Wolf Thresholding Algorithms for Vessel Segmentation and Prewitt Edge Detection and Circular Hough Transform for Optic Disc Detection," *Iran. J. Sci. Technol. Trans. Electr. Eng.*, vol. 45, no. 2, pp. 435–446, 2021.
- [111] X. Xie, S. Ge, M. Xie, F. Hu, and N. Jiang, "An improved industrial sub-pixel edge detection algorithm based on coarse and precise location," *J. Ambient Intell. Humaniz. Comput.*, vol. 11, no. 5, pp. 2061–2070, 2020.
- [112] I. Lorencin, N. Anđelić, J. Španjol, and Z. Car, "Using multi-layer perceptron with Laplacian edge detector for bladder cancer diagnosis," *Artif. Intell. Med.*, vol. 102, p. 101746, 2020.
- [113] M. Kalbasi and H. Nikmehr, "Noise-Robust, Reconfigurable Canny Edge Detection and its Hardware Realization," *IEEE Access*, vol. 8, pp. 39934–39945, 2020.
- [114] J. Canny, "A Computational Approach to Edge Detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
- [115] G. Taubin, "Curve and surface smoothing without shrinkage," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 852–857.
- [116] J. Zhang, Y. Yang, and Q. Dai, "A novel 2D-to-3D scheme by visual attention and occlusion analysis," in *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2011*, 2011,

- pp. 1–4.
- [117] S. Battiato, S. Curti, M. La Cascia, M. Tortora, and E. Scordato, “Depth map generation by image classification,” in *Electronic Imaging 2004*, 2004, pp. 95–104.
- [118] D. A. Forsyth, “Shape from texture and integrability,” in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, 2001, vol. 2, pp. 447–452.
- [119] Q. Wei, “Covertig 2D to 3D: A survey,” *Int. Conf.*, vol. 7, p. 14, 2005.
- [120] V. S. Ramachandran, “Perception of shape from shading,” 1988.
- [121] Y. Iwahori, H. Sugie, and N. Ishii, “Reconstructing shape from shading images under point light source illumination,” in *[1990] Proceedings. 10th International Conference on Pattern Recognition*, vol. i, pp. 83–87.
- [122] B. K. . Horn and M. J. Brooks, “The variational approach to shape from shading,” *Comput. Vision, Graph. Image Process.*, vol. 33, no. 2, pp. 174–208, Feb. 1986.
- [123] E. J. Ciaccio, C. A. Tennyson, G. Bhagat, S. K. Lewis, and P. H. R. Green, “Use of shape-from-shading to estimate three-dimensional architecture in the small intestinal lumen of celiac and control patients.,” *Comput. Methods Programs Biomed.*, vol. 111, no. 3, pp. 676–84, Sep. 2013.
- [124] P. H. R. Green, B. Lebwohl, and R. Greywoode, “Celiac disease.,” *J. Allergy Clin. Immunol.*, vol. 135, no. 5, pp. 1099–106; quiz 1107, May 2015.
- [125] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, 2001, vol. 1, pp. I-511-I-518.
- [126] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for

- Large-Scale Image Recognition,” 2014.
- [127] P. Workman *et al.*, “Guidelines for the welfare and use of animals in cancer research,” *Br. J. Cancer*, vol. 102, no. 11, pp. 1555–77, May 2010.
- [128] J. Lammers, “Maya 4.5 fundamentals,” p. 574, 2003.
- [129] W. F. Pfeffer, *The divergence theorem and sets of finite perimeter*. CRC Press, 2012.
- [130] A. Sharif, R. H. Azizpour, J. Sullivan, and S. Carlsson, “CNN Features off-the-shelf: an Astounding Baseline for Recognition.”
- [131] P. M. M. Pereira *et al.*, “Skin lesion classification enhancement using border-line features – The melanoma vs nevus problem,” *Biomed. Signal Process. Control*, vol. 57, p. 101765, Mar. 2020.
- [132] T. H. Johansen *et al.*, “Recent advances in hyperspectral imaging for melanoma detection,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 12, no. 1. Wiley-Blackwell, p. e1465, 01-Jan-2020.
- [133] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat, and J. Jatakia, “Human Skin Detection Using RGB, HSV and YCbCr Color Models,” vol. 137, pp. 324–332, Aug. 2017.
- [134] N. F. Lattoofi *et al.*, “Melanoma Skin Cancer Detection Based on ABCD Rule,” *1st Int. Sci. Conf. Comput. Appl. Sci. CAS 2019*, pp. 154–157, Dec. 2019.
- [135] S. Sengupta, N. Mittal, and M. Modi, “Improved skin lesion edge detection method using Ant Colony Optimization,” *Ski. Res. Technol.*, vol. 25, no. 6, pp. 846–856, Nov. 2019.
- [136] A. Mohan and S. Poobal, “Crack detection using image processing: A critical review and analysis,” *Alexandria Eng. J.*, vol. 57, no. 2, pp. 787–798, Jun. 2018.

- [137] B. Liu and T. Yang, "Image analysis for detection of bugholes on concrete surface," *Constr. Build. Mater.*, vol. 137, pp. 432–440, Apr. 2017.
- [138] L. C. Ngugi, M. Abelwahab, and M. Abo-Zahhad, "Recent advances in image processing techniques for automated leaf pest and disease recognition – A review," *Inf. Process. Agric.*, vol. 8, no. 1, pp. 27–51, Mar. 2021.
- [139] J. E. C. Cruz, E. H. Shiguemori, and L. N. F. Guimarães, "A comparison of Haar-like, LBP and HOG approaches to concrete and asphalt runway detection in high resolution imagery," *J. Comp. Int. Sci.*, vol. 6, no. 3, pp. 3–20, 2015.
- [140] S. O. Adeshina, H. Ibrahim, S. S. Teoh, and S. C. Hoo, "Custom Face Classification Model for Classroom Using Haar-Like and LBP Features with Their Performance Comparisons," *Electron. 2021, Vol. 10, Page 102*, vol. 10, no. 2, p. 102, Jan. 2021.
- [141] S. Sanjana, S. Sanjana, V. R. Shriya, G. Vaishnavi, and K. Ashwini, "A review on various methodologies used for vehicle classification, helmet detection and number plate recognition," *Evol. Intell. 2020 142*, vol. 14, no. 2, pp. 979–987, Sep. 2020.
- [142] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, 2006.
- [143] A. Arunmozhi and J. Park, "Comparison of HOG, LBP and Haar-Like Features for On-Road Vehicle Detection," *IEEE Int. Conf. Electro Inf. Technol.*, vol. 2018-May, pp. 362–367, Oct. 2018.
- [144] Y. Pang, Y. Yuan, X. Li, and J. Pan, "Efficient HOG human detection," *Signal Processing*, vol. 91, no. 4, pp. 773–781, Apr. 2011.



- [145] C. Li, Z. Qi, N. Jia, and J. Wu, "Human face detection algorithm via Haar cascade classifier combined with three additional classifiers," in *ICEMI 2017 - Proceedings of IEEE 13th International Conference on Electronic Measurement and Instruments*, 2017, vol. 2018-January, pp. 483–487.
- [146] M. Q. Khan *et al.*, "Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer," *IEEE Access*, vol. 7, pp. 90132–90144, 2019.
- [147] P. Kickingreder *et al.*, "Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study," *Lancet Oncol.*, vol. 20, no. 5, pp. 728–740, May 2019.
- [148] J. Delgado-SanMartin *et al.*, "An innovative non-invasive technique for subcutaneous tumour measurements," *PLoS One*, vol. 14, no. 10, p. e0216690, Oct. 2019.