Parsa Paiman

ADVANCING VULVAR TUMOUR DETECTION: A SYNERGY OF HYPERSPECTRAL IMAGING AND ARTIFICIAL INTELLIGENCE

Advancing Vulvar Tumour Detection: A Synergy of Hyperspectral Imaging and Artificial Intelligence

By

Parsa Paiman

In partial fulfilment of the requirements for the degree of

Master of Science
in Biomedical Engineering

at the Delft University of Technology,
to be defended publicly on Tuesday, August 17, 2023, at 13:45.

Student number:	5185068
Supervisor:	Dr. ir. N. J. P. van de Berg
Supervisor:	Prof. dr, J. Dankelman
Thesis committee:	Dr. ir. N. J. P. van de Berg
	Prof. dr, J. Dankelman
	Dr.ir. M.C. Goorden
	Dr. H.C.L. van Doorn

TU Delft

Erasmus MC

Preface

It is with great joy and a sense of accomplishment that I write this foreword to conclude my master's thesis, marking the completion of my studies in Biomedical Engineering. First and foremost, I would like to express my sincere gratitude to my supervisor, Nick van de Berg. His valuable guidance and support throughout this period have played a crucial role in the successful completion of my research project. I would also like to extend my appreciation to Laurie van de Weerd. Her invaluable help and guidance have facilitated my quick adaptation and integration into this project. Additionally, I am deeply grateful to Jenny Dankelman for her dedicated supervision throughout this journey. I would also like to thank Lena van Doorn for her invaluable insights and career advice. Furthermore, I express my thanks to Marlies Goorden for her willingness to be a part of my thesis committee. Lastly, I would like to thank my family, friends, and loved ones, with a special mention to Shabnam, for their unwavering support, encouragement, and understanding throughout this intensive period of my studies. Their love and support have given me the strength to persevere and achieve my goals. My master's thesis represents the culmination of my academic journey, and I am grateful for the opportunities and support I have received.

With pride and a sense of fulfilment, I present my master's thesis.

Parsa Paiman 2023

Abstract

Vulvar cancer has significant implications for women's health and quality of life, emphasizing the need for prompt diagnosis and treatment. Surgical treatment aims to remove both visible and non-visible malignant cells, incorporating a safety margin beyond the tumour perimeter. However, accurately assessing cancer spread and determining the optimal safety margin size remains challenging. Hyperspectral imaging allows non-invasive identification of vulvar tumour boundaries, capturing both spatial and spectral details of the region of interest. However, tissue type classification in hyperspectral imaging is complex due to the abundance of wavebands and limited training samples. Artificial intelligence (AI) methods show promise for accurate and automated tumour classification in hyperspectral imaging. In this study, four AI models (Support Vector Mechanism, Neural Network, 1D and 3D Convolutional Neural Networks) were trained to classify tumour tissue, skin tissue, and mucosa tissue. Additionally, the study aimed to explore the use of specific explanatory parameters (Tissue Water Index, Near Infrared Perfusion Index, and Tissue Haemoglobin Index) to reduce the spectral dimension, potentially improving acquisition time and interpretability without compromising classification accuracy. The study included 25 patients (mean age 71), with hyperspectral cubes captured from each patient. All cubes contained healthy skin tissue, while 22 cubes included tumour tissue and 10 cubes contained mucosa tissue. All models exhibited comparable performance for tumour detection, with F1 scores ranging between 0.91 and 0.93 and AUC-ROC scores ranging between 0.91 and 0.94. Moreover, the Neural Network trained only on the explanatory parameters achieved excellent results for tumour detection with an F1 score of 0.93 and AUC-ROC of 0.91. Overall, combining hyperspectral imaging with all models shows great potential for in vivo tumour detection, while leveraging physiological explanatory parameters with a Neural Network model can enhance acquisition time and data interpretability without compromising the classification accuracy.

Contents

Introduction

Vulvar cancer is a relatively rare form of cancer that affects the external genitalia of women. It arises from the skin and mucous membranes of the vulva, which includes labia majora and labia minora, clitoris, vaginal opening, and perineum. Vulvar cancer can have a significant impact on a woman's health and quality of life and requires prompt diagnosis and treatment. Vulvar cancer accounts for approximately 5% of all gynaecological cancers [1]. While it is considered a rare malignancy, its incidence has been gradually increasing in recent years. Two main causes that can lead to progressive vulvar cancer are vulvar intraepithelial neoplasia (VIN) and vulvar nonneoplastic epithelial disorders (VNED). The first group encompasses cases associated with an infection of the human papillomavirus (HPV), resulting in the development of VIN. Approximately 80% of women with untreated VIN are projected to progress to invasive vulvar cancer. Notably, this particular subtype of vulvar cancer predominantly affects a demographic of younger patients, typically within the age range of 35 to 65 years [1]. The second category encompasses VNED, which contributes to the development of cellular atypia, ultimately leading up to the progression of cancer. Notably, this form of vulvar cancer predominantly affects elderly patients within the age range of 55 to 85 years [1]. Mortality rates vary depending on the stage (I-IV) at diagnosis and individual characteristics [2]. Early detection and appropriate treatment significantly improve the prognosis [3]. The diagnosis of vulvar cancer involves a comprehensive evaluation that includes a medical history review, physical examination, and diagnostic tests. Treatment options for vulvar cancer depend on several factors, including the stage and location of the tumour, as well as the overall health and preferences of the patient. Common treatment modalities include surgery, radiation therapy and chemotherapy. Surgery is the primary treatment for vulvar cancer. It involves the removal of the tumour and a margin of healthy tissue surrounding it. The extra margin of healthy tissue is resected to be sure that non-visible cancer cells are also removed and thereby reducing the chance of recurrence. The extent of surgery depends on the stage and size of the tumour and may involve a partial vulvectomy (removal of a portion of the vulva) or a vulvectomy alongside inguinal lymph node dissection. Radiation therapy uses high-energy X-rays or other forms of radiation to kill cancer cells or prevent their growth. It may be used before surgery to shrink the tumour, after surgery to destroy remaining cancer cells, or as the primary treatment for patients who cannot undergo surgery. Chemotherapy involves the use of powerful drugs to kill cancer cells. It may be administered before or after surgery or as a palliative treatment for advanced or recurrent vulvar cancer [1].

Problem definition

The primary objective of surgical excision in the treatment of vulvar tumours, i.e., vulvectomy, is the comprehensive removal of all observable and non-observable malignant cells. Surgeons, as a precautionary measure, incorporate a safety margin beyond the visible perimeter of the tumour of up to 0.8 cm, ensuring the removal of this additional tissue to diminish the likelihood of recurring malignant cells that may be imperceptible [4]. Despite surgeons possessing extensive experience and having encountered various types and occurrences of vulvar tumours, accurately determining the extent of microscopic cancer cell spread and determining the optimal size of the safety margin remains a challenging task. In the domain of tumour detection, real-time visualization of tumour tissue during surgery is essential for accurate and precise surgical interventions. To address this need, researchers are developing innovative in vivo tumour detection techniques that allow surgeons to identify tumours in real-time, enhancing surgical outcomes and patient care. One prominent approach under research is fluorescence-guided surgery, which utilizes fluorescent dyes that would ideally selectively accumulate in tumour tissues. During surgery, a specialized camera system detects and visualizes the fluorescence emitted by these dyes, providing surgeons with immediate feedback on tumour location and extent. This technique could potentially aid in distinguishing tumour tissue from healthy tissue, facilitating more accurate tumour removal and reducing the risk of leaving residual tumour cells behind [5]. Another promising technique is photoacoustic imaging, which combines laser-induced light absorption and ultrasound detection. By delivering laser pulses to tissues, photoacoustic imaging generates ultrasonic waves that are subsequently captured and processed to create detailed images. In case specific tumour markers could be developed, this technique can potentially target them and enable real-time visualization of tumour tissue, enhancing surgical precision and improving the chances of complete tumour resection [6]. Raman spectroscopy represents yet another potentially valuable tool in tumour detection during surgery. By shining laser light onto tissues and analysing the scattered light, Raman spectroscopy provides molecular information about the tissues. By detecting unique molecular signatures associated with tumours, surgeons could differentiate tumour tissue from healthy tissue, guiding them during surgery [7]. Another emerging technique, known as Hyperspectral Imaging, offers a non-invasive and practical approach to potentially delineate in vivo tumour boundaries. Hyperspectral imaging (HSI) refers to the technique of capturing and analysing spectral information from an object across a wide range of electromagnetic wavelengths. Unlike traditional imaging techniques that capture only three spectral bands; red, green, and blue, HSI records data in hundreds or even thousands of narrow contiguous spectral bands. This rich spectral information can reveal valuable insights about the composition,

properties, and overall characteristics of the imaged tissue [8]. While these techniques demonstrate significant potential in real-time tumour detection during surgery, it is essential to acknowledge that further research and development are necessary to optimize their clinical application. Future advancements in these in vivo techniques hold the promise of improving surgical precision, minimizing surgical risks, and ultimately enhancing patient outcomes in the battle against cancer.

The Department of Oncological Gynaecology at the Erasmus Medical Centre (EMC) is currently engaged in an ongoing research focused on the detection of vulvar tumours through hyperspectral imaging and utilizing machine learning techniques. This master's thesis builds upon this research endeavour by providing an expanded examination that encompasses the description and comparative analysis of various machine learning and deep learning methods employed for the identification of in vivo vulvar tumours.

Research aim

This master's thesis centres around the following research question and sub-questions:

1. Which Artificial Intelligence model achieves the highest performance in terms of evaluation metrics for automated in vivo detection of vulvar tumours in combination with hyperspectral imaging?

The sub-questions that further clarify the main research question are as follows:

1.1. What types of artificial intelligence models and architectures are used in the literature for the detection of in vivo tumour tissue using hyperspectral imaging?

1.2. How does the scarcity of available data impact the performance of artificial intelligence models in terms of model accuracy?

1.3. Can specific explanatory parameters possessing physiological significance be leveraged to enhance the data interpretability and classification speed of the employed models without compromising their classification performance?

Hyperspectral Imaging

Fundamentals of Hyperspectral Imaging

Hyperspectral imaging is an advanced imaging technique that combines spectroscopy and digital imaging to capture a wide range of wavelengths. It provides detailed spectral information for each pixel in an image, enabling the identification and characterization of materials based on their unique spectral signatures.

In hyperspectral imaging, light plays a crucial role in gathering detailed and comprehensive information about the object being illuminated. Notably, a hyperspectral imaging system possesses the capability to capture light across a wide range of wavelengths. This characteristic empowers the extraction of supplementary information extending beyond the range of wavelengths visible to humans. The visible spectrum corresponds to a narrow band of frequencies with wavelengths ranging from 400 to 700 nm, which is perceptible to the human eye and various organisms [9], see Figure 1.

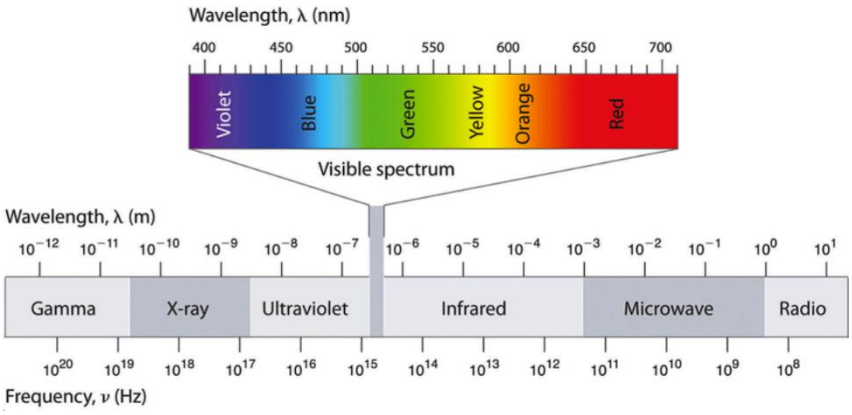

Figure 1: Illustration of the electromagnetic spectrum [8].

As electromagnetic waves cross interfaces between different media, alterations occur in their velocities while their frequencies remain unaltered. Upon entering biological tissue, the electromagnetic wave undergoes numerous scattering and absorption phenomena as it travels through the tissue [10], see Figure 2. Biological tissues are heterogeneous in composition with spatial variations in optical properties. Scattering occurs where there is a spatial variation in the refractive index. The shape and size of the cells vary among different tissue types, with dimensions of a few microns and larger. The penetration depth of light into biological tissues depends on how strongly the tissue absorbs light. When a molecule absorbs energy, it can undergo transitions between two specific energy levels. These transitions are associated with distinct wavelengths. These specific wavelengths can act as a unique spectral fingerprint of the molecule, which can be utilized for diagnostic purposes [10]. For instance, absorption spectra can indicate the concentration and oxygen

saturation of haemoglobin. These measurements could potentially reveal two important signs of cancer: angiogenesis (formation of new blood vessels) and hypermetabolism (increased metabolic activity) [11]. Incident electromagnetic waves can be directly reflected on the surface of the tissue or be scattered due to random spatial variations in tissue density and then be remitted to the tissue surface. Electromagnetic waves become randomized in direction due to scattering, and this is known as diffuse reflectance, which provides information about scattering and absorbing components deep within the tissue. The measured reflectance signal represents the light that has undergone a variety of scatterings and absorptions within the tissue and is, therefore, an average measure of the properties over a certain thin layer of tissue. Knowledge of the origin of the scattering and absorption signals would facilitate accurate modelling and interpretation of the reflectance data. Alterations in tissue morphology can affect the scattering signals.

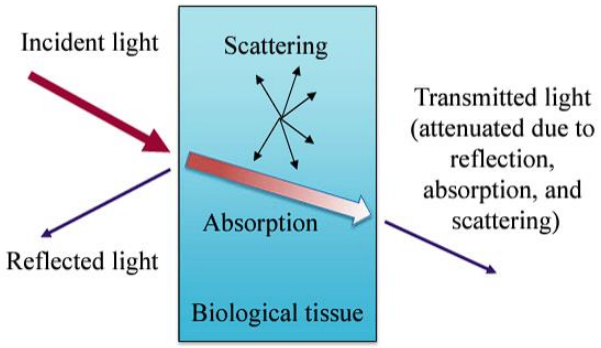

Figure 2: Different types of interaction of light with tissue [9].

Spectral Range and resolution

The spectral range denotes the particular wavelength intervals encompassed by the hyperspectral imaging system. The selection of the spectral range depends upon the intended application and may span ultraviolet, visible, near-infrared, or infrared wavelengths. In the medical domain, hyperspectral imaging systems operating within the visible and very near-infrared spectrum, such as 380-800 nm or 400-1000 nm, are extensively employed [12]. The ability of a hyperspectral imaging system to distinguish features in the electromagnetic spectrum is known as spectral resolution. It quantifies the system's capability to differentiate between neighbouring monochromatic spectral bands. Spectral resolution depends on the wavelength dispersion properties of the spectrograph and the sizes of the entrance and exit apertures.

Spatial Resolution

spatial resolution refers to the level of detail captured by the imaging system. It quantifies the ability of the system to resolve small-scale spatial features and discriminate between neighbouring objects

or areas. Spatial resolution is influenced by the size of the individual pixels or spatial elements in the captured image. A higher spatial resolution corresponds to smaller pixel sizes, enabling the system to capture finer details and provide a more precise representation of the object's spatial structure. This facilitates the identification and characterization of smaller objects, subtle variations, and intricate spatial patterns within the scene. The determination of spatial resolution involves factors such as the sensor's physical characteristics, optics, and the imaging system's design parameters.

Hyperspectral Image Data

Hyperspectral image data comprises of several coinciding images that represent intensity values at different wavelength bands. These images consist of vector pixels, which contain both two-dimensional spatial information and spectral information. This synthesis of spatial and spectral data forms a three-dimensional hyperspectral cube, which provides physical and chemical information about the imaged surface. Creating this three-dimensional dataset is a complex process and requires extensive computing power. An example of a hyperspectral cube extracted from a vulvectomy procedure is illustrated in Figure 3. Each vector in a hyperspectral image contains the spectrum, acting as a unique fingerprint that characterizes the composition of that particular site. The hypercube can be examined in the spatial domain as images at different wavelengths or in the spectral domain as spectral vectors at different sites.

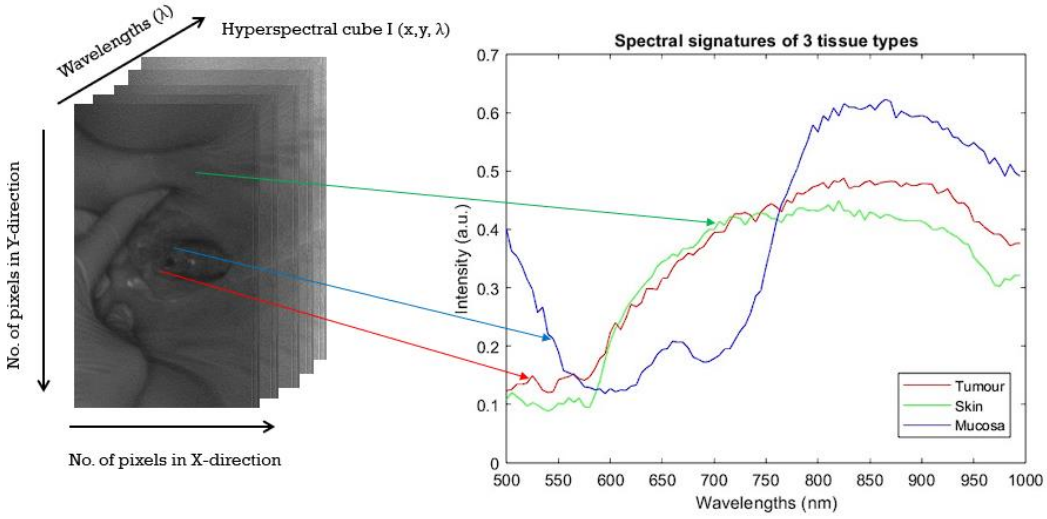

Figure 3: Representation of a hyperspectral cube from a vulva with cancer, alongside with spectral signatures of tumour, skin and mucosa wavelengths.

Acquisition Modes of Hyperspectral Images

Three commonly used methods exist for generating spectral images: area scanning, point scanning, and line scanning, as depicted in Figure 4. These methods involve capturing a subset of the hyperspectral data cube, either one or two-dimensionally, and then temporally scanning the remaining dimensions to obtain the complete hyperspectral data cube. The area scanning design maintains a fixed image field of view and acquires images sequentially, one wavelength at a time. This approach is conceptually referred to as the wavelength-scanning method or band sequential method. The point-scanning method, also known as whiskbroom, involves measuring the spectrum of a single point and then moving the sample to acquire another spectrum. The third method is line scanning, also known as push broom, which entails capturing spectral measurements from a line of the sample simultaneously recorded by an array detector [13, 14]. Although these high-dimensional data offer access to rich information and present new possibilities for uncovering detailed hidden insights, they also pose a challenge for data processing, particularly in real-time applications.

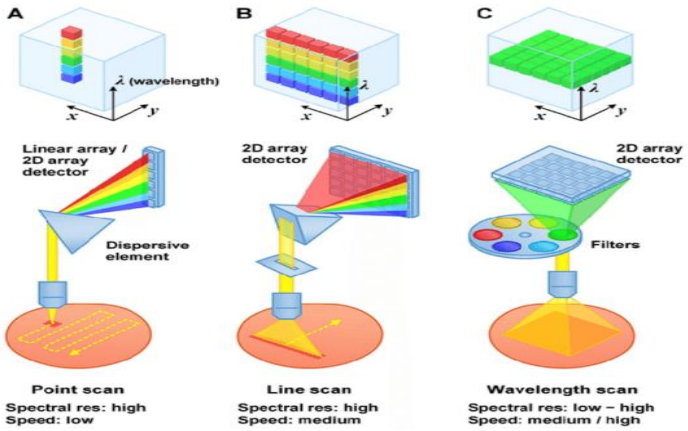

Figure 4: Schematic diagram of a hyperspectral imaging system and methods of spectral image acquisitions with (a) point scan, (b) line scan, and (c) wavelength scan [12,14].

Artificial intelligence and Hyperspectral imaging

Hyperspectral imaging is characterized by its ability to capture rich spatial and spectral information in high-resolution, high-dimensional images composed of hundreds of adjacent bands across the electromagnetic spectrum. However, the task of tissue type classification in HSI is inherently complex. The large number of wavebands present in HSI data poses challenges for classification models, particularly in achieving higher accuracy when dealing with limited training samples. Traditional

methods rely on expert knowledge and manual adjustment of hyperparameters to design and extract relevant features. Machine learning techniques such as support vector machines and artificial neural networks have already been applied to HSI tumour classifications [15-17]. Moreover, deep learning approaches such as convolutional neural networks have emerged as efficient methods for automatically extracting highly robust and discriminative features. These deep learning methods have demonstrated their ability to achieve more accurate classification results compared to other traditional learning methods [15, 17, 18].

Artificial intelligence applied in tumour detection

Machine learning

Support vector machine

Support Vector Machine (SVM) is one of the most popular machine-learning models for tumour-related HSI image classification tasks and has demonstrated excellent performance in various studies [16, 17, 19-21]. One advantage of SVMs is their ability to handle high-dimensional data, which is often the case with HSI datasets. SVMs can effectively handle this high dimensionality and capture complex relationships between the spectral information and the corresponding classes. Additionally, SVMs are known for their excellent generalization capabilities, even with limited training data [22].

SVM constructs a model within a high-dimensional feature space to achieve the highest possible accuracy in its predictions. To train the model, an optimization algorithm based on learning theory is employed, incorporating a learning bias to enhance its performance [23]. The primary focus is on efficiently finding a suitable hyperplane that can effectively separate different classes within the high-dimensional feature space, see Figure 5. The quality of this hyperplane is evaluated based on its ability to generalize well to unlabelled test data. A key measure used for evaluating the hyperplane's quality is the concept of margin. The margin refers to the minimum distance between the separating hyperplane and the closest data points from each class. By maximizing the margin, the optimal hyperplane is obtained, which results in the maximum separation between the different classes. This separation allows for better classification and improves the ability of the SVM model to accurately predict the class labels of unlabelled test data [22].

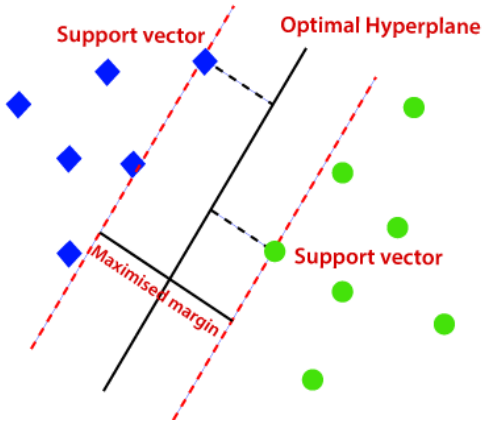

Figure 5: The hyperplane, represented by the solid black line, serves as the separation boundary, while the dotted red line indicates the separation margin. The coloured squares and circles represent the different classes present in the classification problem [23].

Artificial Neural network

Hyperspectral images consist of numerous spectral bands, providing detailed spectral information for each pixel. Artificial neural networks (ANNs) excel in processing such complex data and learning the underlying patterns that distinguish different classes or characteristics within the hyperspectral cubes [21, 25].

ANNs are computational models inspired by the structure and functioning of biological neural networks in the human brain. ANNs consist of interconnected nodes organized into layers [26]. The basic architecture of an ANN includes an input layer, one or more hidden layers, and an output layer. Each neuron in the input layer represents a feature or variable of the input data. For example, in the hyperspectral cube, each neuron in the input layer represents a spectrum. The hidden layers are not directly accessible from the input or output. Each hidden layer is composed of multiple neurons that process and transform the information received from the previous layer. The output layer produces the final result or prediction of the neural network (NN). Each neuron in the NN is associated with a weight and a threshold. The weight determines the strength or significance of the connection between neurons. It reflects the impact of the input signal on the output of a neuron. Neurons with larger weights have a more substantial influence on the final output compared to those with smaller weights. To calculate the output of a neuron, each input is multiplied by its corresponding weight, and the results are summed. This weighted sum is then passed through an activation function, which introduces non-linearities, enabling the NN to learn complex patterns, and determines the final output value, see Figure 6 [26, 27].

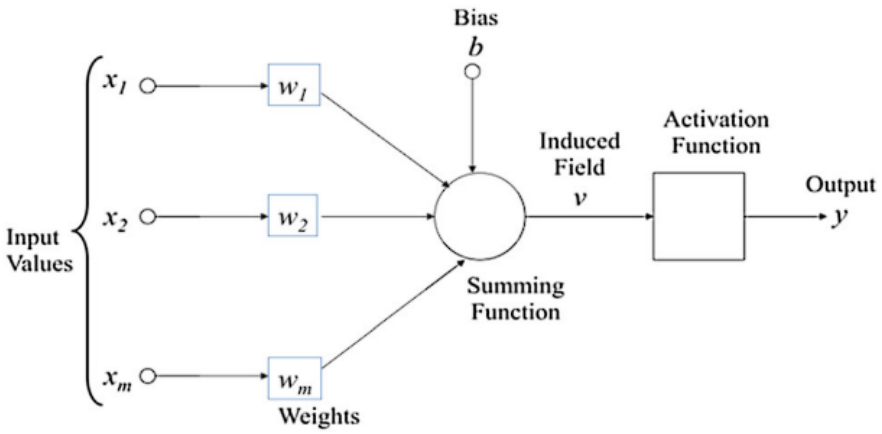

Figure 6: A schematic overview of an artificial neural network and its various layers. The input values are multiplied by corresponding weights, and the resulting products are summed. This aggregated value is then passed through an activation function, which produces an output corresponding to the class with the highest probability [27].

Deep learning

Convolution neural network

The third model is a Convolutional Neural Network (CNN), which has proven to be highly effective in analysing various types of image data, including hyperspectral images [15, 18, 20, 29]. By utilizing the multiple channels present in hyperspectral cubes, CNNs can effectively exploit the spatial and spectral relationships in the data, resulting in improved classification performance [30-32].

Convolutional Neural Networks (CNNs) are a specific class of AI systems that leverage multi-layered neural networks to effectively detect, recognize, and classify. CNNs are particularly useful because they can learn directly from the input data without the need for manual feature extraction. A typical CNN consists of four layers: convolution, activation, pooling, and a fully connected layer, see Figure 7 [30].

The convolutional layer is a crucial component of CNN's overall structure. It applies a set of filters to the input data, where each filter has its dimensions and weights. By adjusting their weights during training, these filters slide over the input data to perform element-wise multiplications and summations to extract important features from the input data. This allows the CNN to detect edges, textures, or other meaningful patterns and features [30].

The activation function layer, which follows the convolutional layer in the CNN architecture, serves an essential role in the mapping of inputs to outputs. Activation functions are fundamental components in all types of neural networks. By incorporating these non-linear activation layers, CNNs gain the

ability to capture and learn intricate patterns within the data. This enables the network to model complex relationships and enhance its capacity for accurate classification and recognition tasks [33].

The pooling layer, also referred to as the down-sampling layer, plays a vital role in reducing the size of the feature maps while retaining essential information. This layer utilizes pooling operations, such as max pooling or average pooling, to simplify the input data. By down sampling, this layer effectively simplifies the subsequent layers, making it particularly valuable for handling high-dimensional data sources.

The fully connected layer serves as a connection between every neuron in the layer and those in the preceding layer. Its primary function is to comprehend intricate patterns and relationships among features within the input data. Each neuron within the fully connected layer determines its output by considering the inputs it receives from all neurons in the previous layer. This extensive connectivity allows the neural network to capture complex relationships and acquire more sophisticated representations of the data, thereby enhancing its ability to comprehend and classify intricate patterns.

In addition to the standard CNN architecture, regularization layers can be incorporated into the network. Batch regularization introduces randomness and noise during training, effectively controlling network complexity and mitigating the risk of overfitting. It achieves this by including a regularization term in the loss function, which penalizes large weights. This leads to the network learning more intricate details and prevents individual weights from overpowering the learning process. Moreover, batch regularization contributes to the stability of the training process by attenuating the impact of noisy or outlier examples within a batch. Since weight updates are based on the average error computed over a batch, extreme errors from individual examples are dampened, resulting in a smoother and more consistent training trajectory.

The multiple layers in a CNN collaborate to perform the network's operations. During the training phase, a critical procedure called backpropagation is employed to optimize the network's performance. Backpropagation calculates the error value by comparing the output of the previous layer with the desired output and then adjusts the weights of each neuron within that layer accordingly.

This iterative process aims to minimize the error or loss function. By repeatedly updating the weights based on the calculated error and minimizing the cost function, the network gradually learns to make better predictions and improve its overall performance [31, 32].

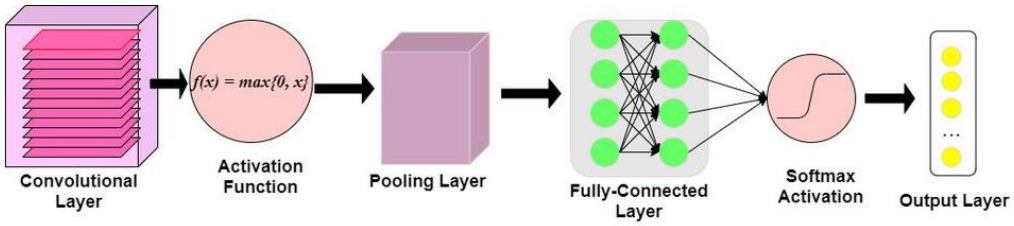

Figure 7: Schematic overview of a CNN. The input data initially passes through a convolution layer, followed by an activation layer. Subsequently, the data is down sampled using the pooling layer. The fully connected layer analyses complex patterns and relationships among the input features. Finally, the SoftMax activation calculates the probabilities for the highest class [32].

Methods and Materials

Data acquisition

Hyperspectral cubes of in vivo vulvar tumours and RGB images were acquired with the Tivita 2.0 (Diaspective Vision GmbH) via the push-broom method with an image resolution of 640*480 pixels. The Field of view (FOV) of the camera is approximately 8 x 6 cm, which gives a spatial resolution of approximately 0.125 mm/pixel. The hyperspectral cubes and RGB images are acquired with two separate sensors positioned next to each other. Each hyperspectral cube consists of wavelength bands ranging from 500 nm to 1000 nm with a spectral resolution of 5 nm, resulting in 100 different wavelengths. Higher wavelengths can achieve a penetration depth of up to 6 mm, depending on the wavelength. For the measurement process, the camera is placed at a distance of approximately 50 cm, and one measurement takes approximately 6.4 seconds.

Data preprocessing

Raw hyperspectral cubes directly from the Tivita 2.0 were first processed in Python into a readable format for MATLAB (2023a). Thereafter a calibration was performed on the raw data for normalization of the data and removal of light glares. In addition, the data was filtered with a median filter for noise reduction.

To label the data for tumour tissue, healthy skin tissue, and healthy mucosa, an experienced surgeon annotated these regions in the RGB images. Additionally, the resected tumour tissues were confirmed as tumour tissue by a pathologist, and the annotations were further refined based on this confirmation. Given that the annotations were drawn on the RGB images, their alignment with the hyperspectral cubes was not entirely accurate. Linear geometric image registration employing

rotation, translation and scaling was therefore performed on the RGB images to fully align the annotations with the hyperspectral cubes, see Figure 8.

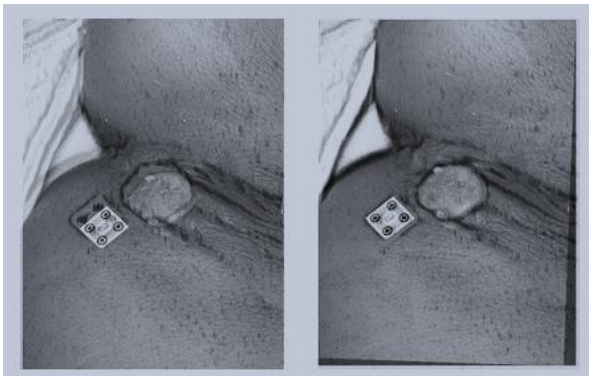

Figure 8: Overlay of the RGB image on which the annotation was made by the surgeon and the HS Cube. Left: Overlay of the RGB and hyperspectral images before image registration. Right: Overlay of the images after image registration.

To train the AI methodologies, the annotated regions were extracted from the hyperspectral cubes and divided into sub-patches measuring 1x1x100 pixels for the machine learning approaches, while for the deep learning approach, which also incorporates spatial information, sub-patches of size 5x5x100 pixels were employed. This cropping process allowed for the isolation of smaller sections of the annotated regions, enabling effective utilization of the respective machine learning and deep learning methods, see Figure 9 for a visual representation. In addition, subsets of the hyperspectral cube were extracted and transformed into explanatory parameters possessing physiological significance (FPPS). Namely, near-infrared perfusion index (NIR), tissue water index (TWI) and Tissue haemoglobin index (THI). The NIR Perfusion Index is computed by taking the ratio of the mean values of spectra within the ranges of 655–735 nm and 825–925 nm, see Equation 1. It captures the characteristics of the NIR range, known for its deep light penetration. Similarly, the THI evaluates the distribution of haemoglobin in the tissue's microcirculatory system. The calculation procedure for both the THI index and TWI index is similar to the NIR index and involves considering the dominant absorption ranges, 530–590 nm and 785–825 nm for THI, and 880–900 nm and 955–980 nm for TWI [35].

$$FPPS = \frac{mean(HS\ cube)_{[spectral\ range\ 1]}}{mean(HS\ cube)_{[spectral\ range\ 2]}} \quad 1$$

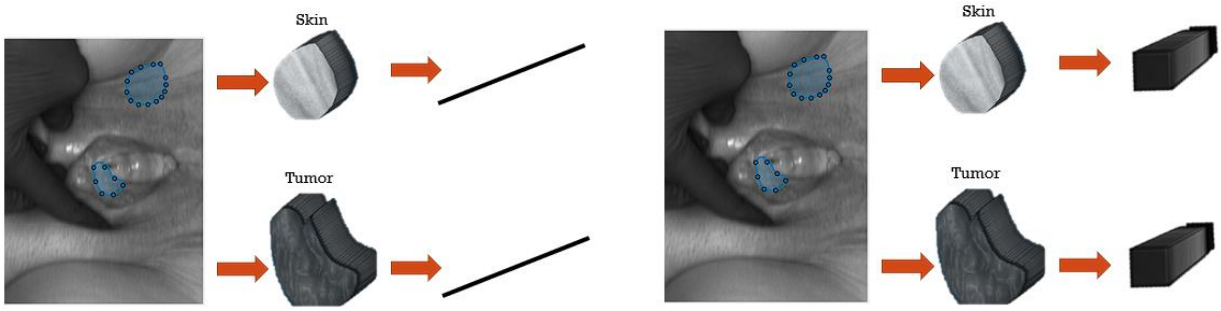

*Figure 9: Data preprocessing pipeline for the deep learning methods (left) and machine learning methods (right). Annotated regions from HS Cubes are extracted and further divided into 1*1*100 patches (left) or into 5*5*100 patches (right).*

Artificial intelligence methods

Support vector machine

The SVM classifier was implemented using the classification learner application in MATLAB. To effectively address intricate and non-linear patterns present in the data, a Gaussian kernel was utilized, incorporating an automated process to determine the appropriate kernel scale. This determination of scale is achieved through subsampling, where a subset of the data is selected, enabling the selection of a kernel size that aligns with the characteristics of this specific data subset. Unlike linear SVM, the Gaussian Kernel SVM incorporates a kernel function that enables the creation of more flexible decision boundaries [36]. This characteristic makes it particularly advantageous for datasets with intricate patterns and overlapping classes. In the implementation, the hyperparameters were set as follows: The box constraint level, which controls the trade-off between maximizing the margin and minimizing the training error, was set to 1. The classification method used was one-vs-one, which constructs binary classifiers for each pair of classes.

Artificial neural network

The ANN was also implemented using the classification learner application. A trilayer neural network architecture was employed, comprising three fully connected layers. The first layer consisted of 20 neurons, followed by two subsequent layers with 35 neurons each. To introduce non-linearity and enhance the network's power, the ReLU activation function was utilized. Regularization was applied to mitigate overfitting. Specifically, a regularization strength of 1/N was employed, where N represents the number of data points in the dataset.

Minibatch size refers to the number of training examples used in one iteration of stochastic gradient descent, while an epoch represents a complete pass through the entire training dataset during the learning process.

Convolution neural network

The CNN was constructed by adopting a pre-existing 3D CNN architecture that had been previously employed in a study on Hyperspectral Imaging (HSI) [37]. The design of the CNN used in this study also incorporates ReLU and batch normalization (BN) layers following each convolution layer, aiming to enhance both model complexity and training stability. Additionally, alterations were made to the filter sizes to optimize performance with the given dataset. Moreover, various training parameters, such as minibatch size, the number of training examples used in one iteration, and epochs, the number of times a complete pass through the entire training dataset is performed, were empirically evaluated and selected based on their ability to yield the most favourable outcomes. The CNN trained on a sub-volume of a 5x5 spatial window, accompanied by a spectral dimension of 100 wavelengths. The network comprised 20 hidden layers consisting of convolution, ReLU, and BN layers, ultimately leading to the last hidden layer composed of a fully connected layer, see Figure 10. The fully connected layer has an output of 3 neurons, representing the number of classes, connected to a softmax layer for probability computation. Furthermore, the 1D CNN network underwent training using the same architecture as the 3D CNN without including spatial convolutions. The CNN training utilized the Adam optimization algorithm with a learning rate of 0.001. Adam stands for "Adaptive Moment Estimation" and combines the benefits of two other optimization techniques, namely AdaGrad and RMSprop. It dynamically adjusts the learning rate based on the gradient's first and second moments, allowing for effective optimization of the model's parameters. A minibatch size of 64 was chosen for the training process, and the training spanned five epochs. To tackle the challenge of an imbalanced dataset, where certain classes were much more dominant than others, a weighted cross-entropy loss function was utilized [37]. The weights used in this function were assigned inversely proportional to the number of training samples per class. This approach aimed to prevent the classifier from solely prioritizing the performance of the most prevalent class, which in this case was the skin class.

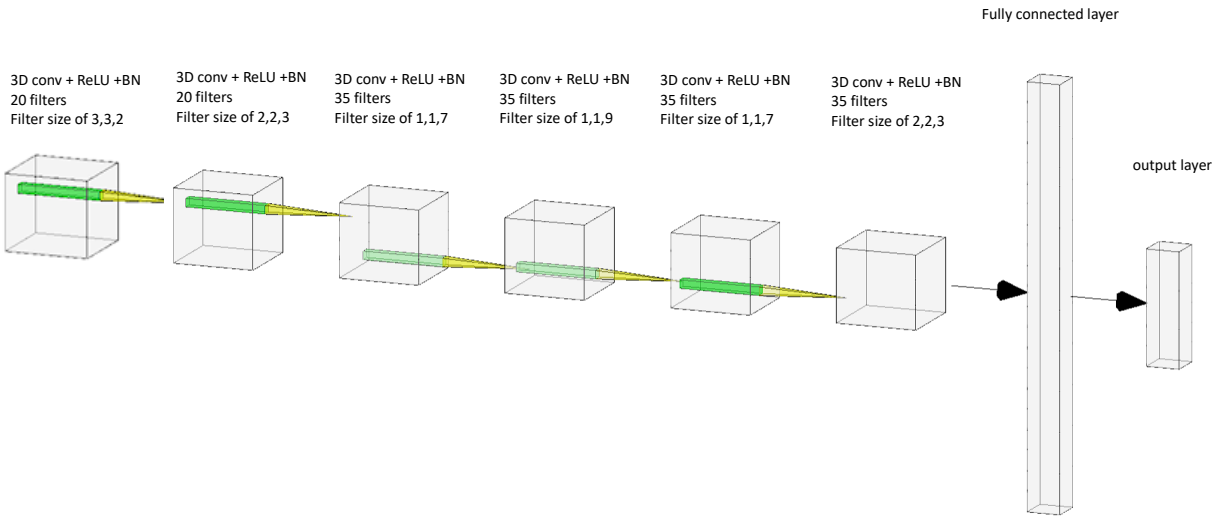

Figure 10: The convolutional neural network architecture employed in this study utilizes an HSI sub-volume as input, centred on a specific pixel. The sub-volume has spatial dimensions of 5 × 5 and consists of 100 wavelength dimensions. Down-sampling convolutional operations are applied to process the input data, transforming it into a final 1D feature space vector. Subsequently, a fully connected layer is utilized to generate the tissue prediction scores.

Model Evaluation and Learning Curve Analysis

To evaluate the model's ability to classify data accurately and make predictions on new, unlabelled data, the dataset was split into two separate sets: a training set used for learning and a testing set used for assessing the model's performance. The partitioning involved allocating 80% of the data for training purposes, while the remaining 20% was reserved for evaluating the model's performance. Given the limited amount of data available and the variability within each class, a k-cross validation method was employed. The dataset was split into five subsets, and in each iteration, patients were randomly assigned to either the training or testing set, ensuring different patient combinations for each split. This approach helps to thoroughly assess the model's classification effectiveness and its ability to generalize to new, unlabelled data. Eventually, the mean of all the trained models was taken to reduce the impact of the limited data and the variation within each class. This approach provides a reliable method to assess the model's performance and its ability to perform well on unseen data.

To establish a learning curve and examine the impact of data quantity on model performance, a rigorous experimental setup was employed. Each model underwent ten training iterations, employing a systematic approach to incrementally increase the amount of available data. The process began by utilizing data from two patients and gradually augmenting the dataset by incorporating data from an additional two or three patients in each subsequent iteration. The process was repeated five times to average the effect of any circumstances where the models performed well from the start due to the training and testing datasets having similar tumour types from the start.

Performance Metrics

When evaluating AI methods, it is imperative to utilize a range of measures to comprehensively assess their performance and effectiveness. In this study, key evaluation metrics, such as accuracy, sensitivity, specificity, F1 score, as well as receiver operating characteristic (ROC) curves and the corresponding area under the curve (AUC) metrics, were employed. The utilization of these measures aims to provide a comprehensive understanding of the significance and implications of the evaluated AI methods. Accuracy quantifies the correct classification rate by dividing the sum of true positives (TP) and true negatives (TN) by the total number of samples, see Equation 2. Sensitivity measures the accurate identification of positive instances by dividing TP by the sum of TP and false negatives (FN), see Equation 3. Specificity evaluates the correct identification of negative instances by dividing TN by the sum of TN and false positives (FP), see Equation 4. The F1 score, a balanced metric combining accuracy and sensitivity, see Equation 5, provides a consolidated measure of overall model performance in classifying different tissue types. Furthermore, ROC curves and the corresponding AUC in AI models were utilized in the evaluation and comparison of the different models. ROC curves provide a graphical representation of the balance between the true positive rate and the false positive rate across diverse classification thresholds. They offer a visual means to assess the model's discriminatory capacity. The AUC metric, derived from the ROC curve, provides a concise numerical measure of the overall performance and discriminative ability of the model. A higher AUC value signifies a better capability to accurately classify instances, irrespective of the selected threshold.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Senistivity = \frac{TP}{TP + FN} \quad (3)$$

$$Specificity = \frac{TN}{TN + FP} \quad (4)$$

$$F1 = 2 * \frac{Accuracy * Sensitivty}{Accuracy + Sensitivity} \quad (5)$$

Results

The study comprised 25 patients with squamous cell carcinoma, either on the skin tissue or mucosa tissue, ranging from stages 1 to 3C, see Table 1. In vivo, hyperspectral cubes of vulvar tumours were captured from these patients. All 25 hyperspectral cubes included regions of healthy skin tissue, while 22 cubes had areas with tumour tissue, and ten cubes contained mucosa tissue. Figure 11 displays the average spectra of the three tissue types across the entire dataset of patients, accompanied by their respective standard deviations. The training of the machine learning models involved 269897 trainable spectra, of which 89873 belong to the tumour tissue, 154255 belong to the skin tissue, and 25769 belong to the mucosa tissue. Each spectrum represented a single pixel alongside a spectral dimension consisting of 100 wavelengths. The deep learning models trained on 10796 spectra, of which 3595 existed from tumour tissue, 6170 from skin tissue and 1031 from mucosa tissue. Each spectrum represented a 5*5 pixel alongside a spectral dimension consisting of 100 wavelengths. The ratio among the various classes exhibited an approximate proportion of 6 to 3.5 for the skin class in comparison to the tumour class and 6 to 1 for the skin class in comparison to the mucosa class. Detailed visual representations showing the distinct tissue types and their differences within each group can be seen in Figure 12.

Table 1: Patient information and clinical picture.

Total Patients	25
Tissue type annotation	Tumour: 22 Mucosa: 10 Skin: 25
Mean Age	71
Tumour Type	Squamous Cell Carcinoma
Location	Skin: 18% Mucosa: 82%
Stadium	1: 10% 1B: 57% 3: 5% 3A: 10% 3B: 5% 3C: 14%

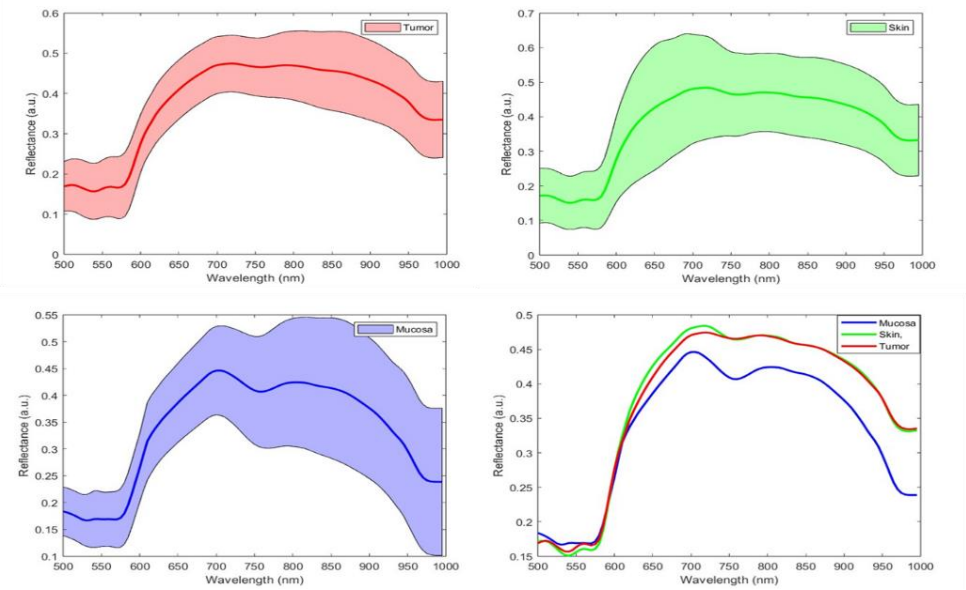

Figure 11: Mean spectra of all three tissue types with their standard deviation plotted as a shaded region around the mean spectra.

The study results are presented in Tables 2-5, providing the mean and standard deviation of the metrics across all validation sets. Notably, the NN model and 3D CNN exhibit the highest accuracy in both tumour and skin classification tasks, with the 3D CNN model also achieving the highest accuracy for mucosa classification. Regarding specificity, the 1D CNN model demonstrates the highest score for tumour classification, while the NN model attains the highest specificity scores for skin and mucosa classification. In terms of sensitivity, both the NN and SVM models achieve the highest scores for tumour classification, while the 3D CNN model scores the highest for both mucosa and skin classification. Additionally, the NN model outperforms other models in terms of F1 scores for tumour classification, while the 3D CNN model exhibits the highest F1 scores for skin and mucosa classification. Figure 13 shows a visual representation of the classifications generated by the various models. It is worth noting that all performance metrics consistently exhibit relatively high standard deviations, indicating some variability in the results. The models trained solely on fractions with physiological significance, such as THI, TWI, and NIR, demonstrate inconsistent performance across the metrics. While the NN + FPPS model performs comparably to models trained on the full spectra, the SVM + FPPS model consistently exhibits lower scores compared to other trained models.

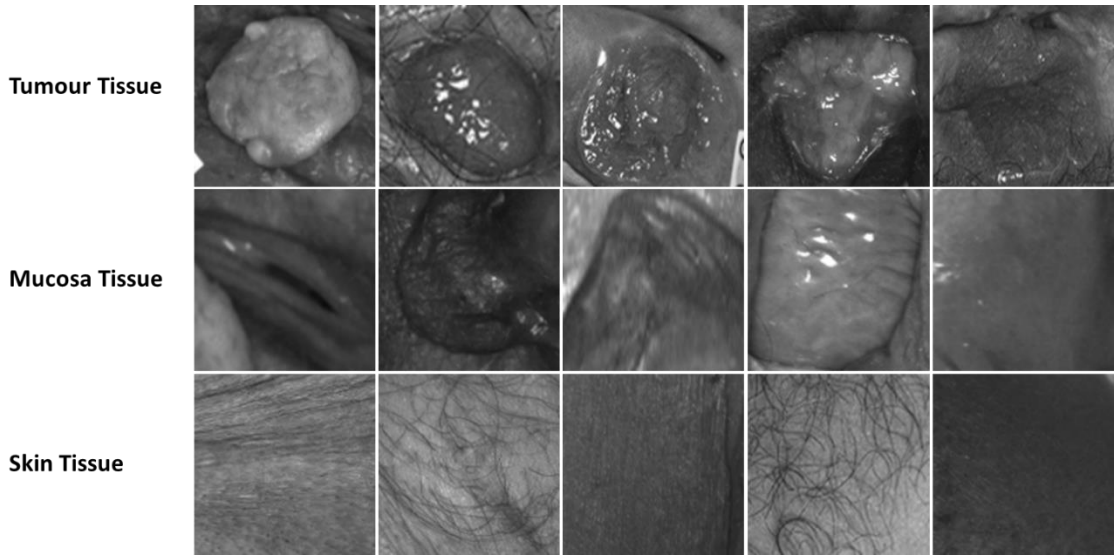

Figure 12: Examples of the 3 tissue types and their diversity within each group. The first row displays examples of tumour tissue. The second row displays examples of mucosa tissue. The third row displays examples of skin tissue.

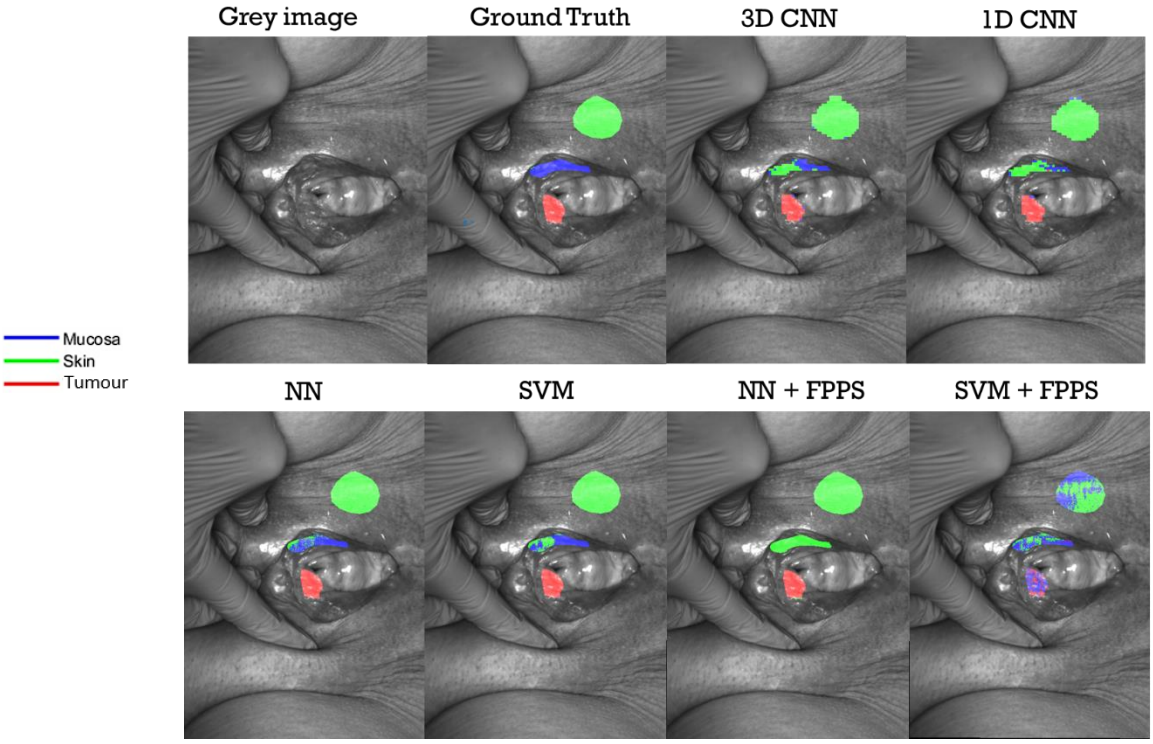

Figure 13: Visual portrayal of the classification accuracy of the trained models alongside the RGB and Ground Truth images. Green represents tissue classified as healthy skin, blue represents tissue classified as mucosa and red represents tissue classified as tumour.

Table 2: Performance of the Six classification models in terms of Accuracy.

Accuracy	Tumour	Skin	Mucosa	Overall
3D CNN	0.92 ± 0.07	0.89 ± 0.13	0.78 ± 0.06	0.86
1D CNN	0.91 ± 0.07	0.87 ± 0.12	0.76 ± 0.03	0.85
NN	0.92 ± 0.06	0.89 ± 0.12	0.80 ± 0.10	0.87
SVM	0.91 ± 0.08	0.86 ± 0.14	0.76 ± 0.07	0.84
NN + FPPS	0.91 ± 0.09	0.87 ± 0.14	0.66 ± 0.08	0.81
SVM + FPPS	0.67 ± 0.11	0.67 ± 0.17	0.54 ± 0.04	0.63

Table 3: Performance of the Six classification models in terms of specificity.

Specificity	Tumour	Skin	Mucosa	Overall
3D CNN	0.91 ± 0.14	0.89 ± 0.15	0.97 ± 0.02	0.90
1D CNN	0.94 ± 0.09	0.89 ± 0.15	0.97 ± 0.02	0.93
NN	0.89 ± 0.11	0.92 ± 0.12	0.99 ± 0.01	0.93
SVM	0.83 ± 0.19	0.91 ± 0.12	0.96 ± 0.04	0.90
NN + FPPS	0.89 ± 0.11	0.84 ± 0.20	0.97 ± 0.04	0.90
SVM + FPPS	0.80 ± 0.18	0.57 ± 0.14	0.85 ± 0.01	74

Table 4: Performance of the Six classification models in terms of Sensitivity.

Sensitivity	Tumour	Skin	Mucosa	Overall
3D CNN	0.90 ± 0.10	0.89 ± 0.15	0.55 ± 0.13	0.78
1D CNN	0.89 ± 0.09	0.86 ± 0.17	0.47 ± 0.07	0.74
NN	0.97 ± 0.05	0.85 ± 0.20	0.55 ± 0.25	0.79
SVM	0.97 ± 0.03	0.81 ± 0.25	0.50 ± 0.17	0.76
NN + FPPS	0.93 ± 0.06	0.90 ± 0.16	0.23 ± 0.16	0.69
SVM + FPPS	0.54 ± 0.04	0.76 ± 0.22	0.22 ± 0.04	0.51

Table 5: Performance of the Six classification models in terms of F1.

F1	Tumour	Skin	Mucosa	Overall
3D CNN	0.91 ± 0.08	0.88 ± 0.16	0.65 ± 0.11	0.81
1D CNN	0.91 ± 0.08	0.86 ± 0.15	0.56 ± 0.06	0.78
NN	0.93 ± 0.06	0.86 ± 0.16	0.62 ± 0.25	0.80
SVM	0.92 ± 0.07	0.82 ± 0.21	0.59 ± 0.19	0.78
NN + FPPS	0.91 ± 0.09	0.87 ± 0.14	0.39 ± 0.15	0.72
SVM + FPPS	0.55 ± 0.12	0.68 ± 0.18	0.21 ± 0.05	0.48

Model performance was further evaluated using ROC curves and AUC metrics. ROC curves and corresponding AUC values were calculated for each model. The results of the ROC curves are presented in Figure 14, and the AUC scores are summarized in Table 6. Notably, the 1D-CNN model exhibited the highest ROC-AUC score for the tumour class. Conversely, the 3D-CNN model showcased the highest ROC-AUC scores for skin and mucosa classification. Furthermore, it is noteworthy that the mucosa class consistently displayed the lowest ROC-AUC score across all models. On the other hand, the skin and tumour classes exhibited relatively comparable trends in terms of ROC-AUC scores. Figure 15 illustrates the average outcomes of the experiment examining the impact of dataset expansion on model accuracy. The error bars, depicted as standard deviations, provide a measure of variability. The graph reveals that after reaching a dataset size of approximately 10, the rate of accuracy improvement begins to diminish. The accuracy values subsequently fluctuate between 0.7 and 0.9, depending on the tissue type. Notably, this trend is observed earlier for the 3D CNN model. Additionally, the standard deviations of the models are relatively high, particularly at smaller dataset sizes and at the maximum available dataset size.

Table 6: Performance of the Six classification models in terms of AUC.

AUC	Tumour	Skin	Mucosa	Overall
3D CNN	0.91	0.95	0.80	0.89
1D CNN	0.94	0.91	0.72	0.86
NN	0.91	0.91	0.78	0.87
SVM	0.91	0.90	0.80	0.87
NN + FPPS	0.91	0.86	0.62	0.80
SVM + FPPS	0.83	0.78	0.59	0.73

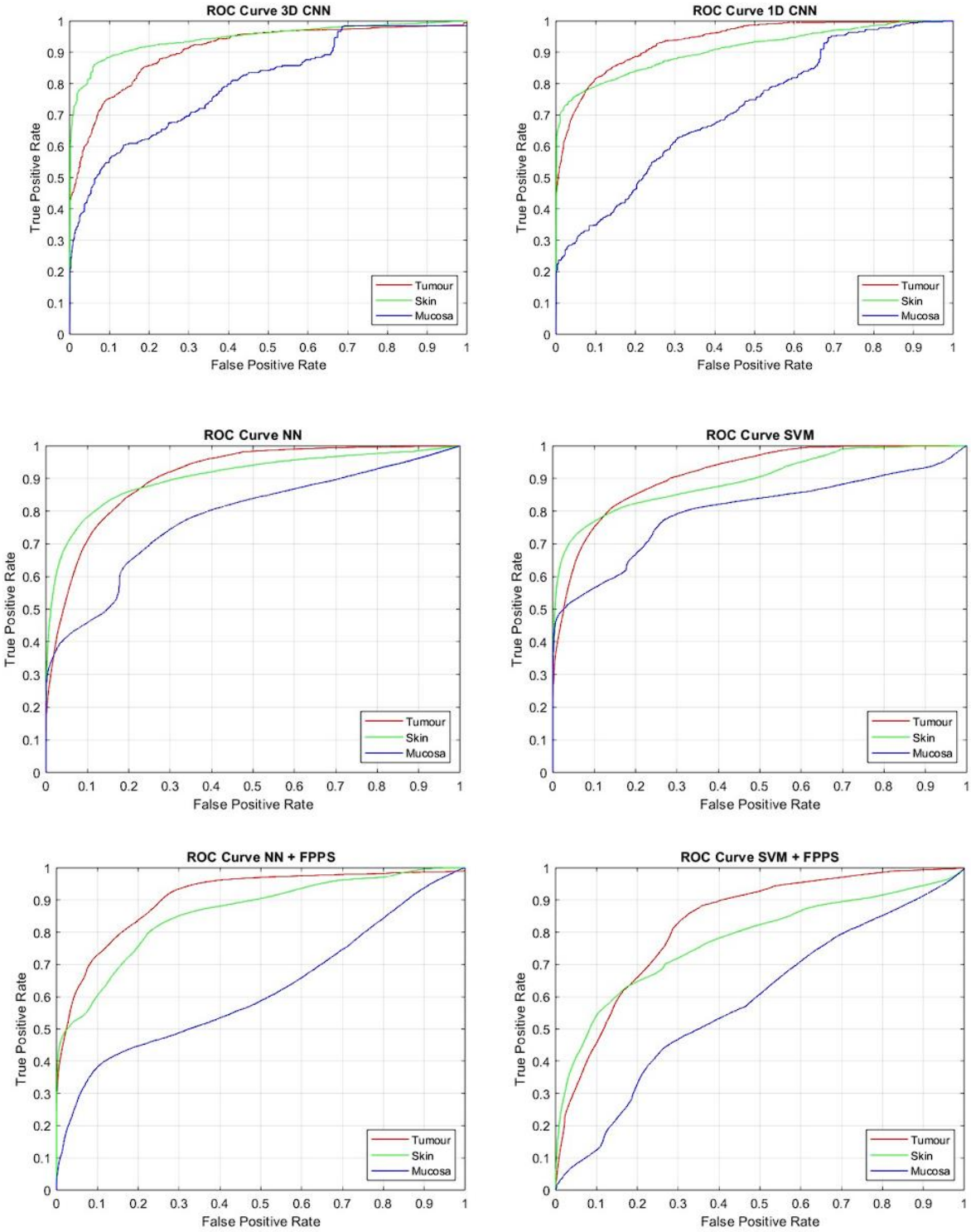

Figure 14: Receiver operator characteristic curves: Receiver operator characteristic curves for the six models. Each plot, represents the mean ROC curve for a specific class.

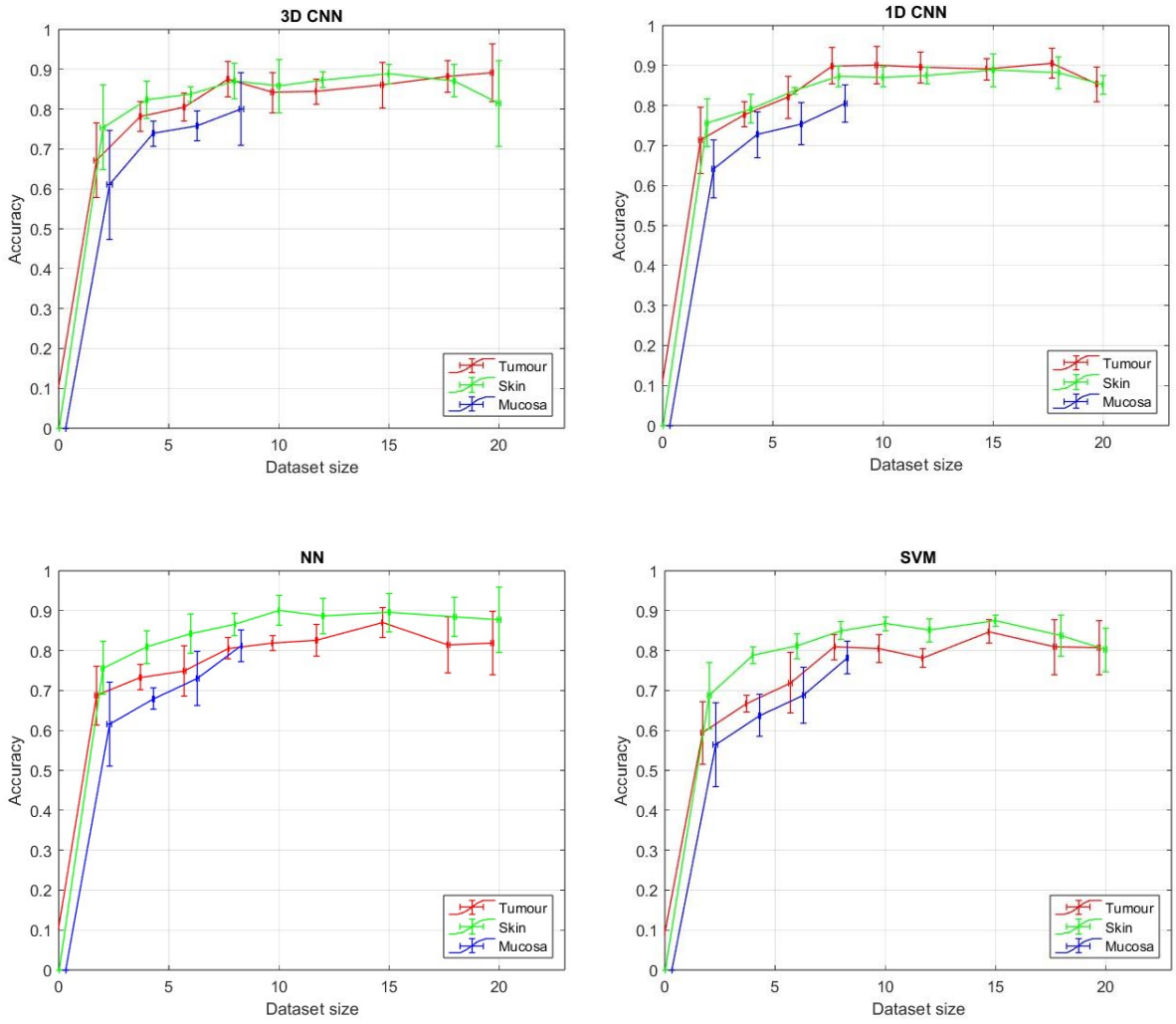

Figure 15: Graphical plots on the effect of Dataset size on the accuracy of the models for the different tissue types alongside the standard deviation represented as error bars.

Discussion

In this study, a significant achievement was made in classifying three different tissue classes using four AI methods, despite dealing with relatively small and imbalanced sample sizes in an in vivo setting. The four AI methods—SVM, NN, 3D CNN, and 1DCNN—all demonstrated comparable performance in classifying tumour tissue, healthy mucosa tissue, and skin tissue. The F1 scores for tumour detection ranged from 0.91 for the 3D CNN to 0.93 for NN , indicating the model performance. Additionally, the four models also showed similar results in tumour detection for the ROC-AUC metric, with scores ranging between 0.94 for the 1D CNN and 0.91 for the other models. However, it is essential to emphasize that all four models struggled relatively more in classifying mucosa tissue. This challenge can be attributed to the limited availability of data specific to the

mucosa class. Therefore, expanding the dataset for mucosa tissue could potentially enhance the model's classification performance.

In this study, the detection of vulva tumours through hyperspectral imaging and AI was investigated to gain insights into the possibilities of the synergy of these techniques. Although this study stands as a novel exploration, it is essential to acknowledge that there are no direct comparative studies available to compare the results with. Nevertheless, the absence of similar studies also underscores the significance of this research, as it fills a notable gap in the existing literature. Despite the lack of direct comparisons, this study can be contextualized within the broader body of research on tumour detection through hyperspectral imaging and AI. In the domain of in vivo HSI-based cancer detection with AI, other studies have been conducted [15, 17, 38-40]. These pioneering investigations have demonstrated promising outcomes, showcasing the ability to accurately discriminate cancerous lesions through the integration of HSI and AI algorithms. Such literature context underscores the relevance and importance of this research, guiding future investigations and applications in this domain.

Regarding the models trained on only fractions with physiological significance, it is worth noting that the NN model performed similarly to the models that used the entire spectral range, except for mucosa classification. On the other hand, the SVM model demonstrated inferior performance across all three classification classes compared to its counterpart, which utilized the complete spectral range. In terms of classification and training speed, no noticeable differences were observed between using only the fractions and utilizing the entire spectral dataset. However, the findings suggest that the NN model yields comparable outcomes to those achieved by utilizing the entire spectrum for tumour tissue and skin tissue, thereby indicating the feasibility of employing solely the spectra required for calculating physiologically meaningful fractions within the HSI camera instead of the entire 500-1000 nm range. In this study, only three explainable parameters were utilized: THI, NIR, and TWI. However, it's worth considering that these three parameters might not be the most optimal for effectively classifying and distinguishing between the given tissue types. Other explainable parameters could potentially be more suitable, such as tissue oxygenation (StO_2) or tissue lipid index (TLI), or even different combinations of these explainable parameters. Exploring alternative parameters could lead to improved accuracy and performance in tissue type classification. This approach holds the potential to substantially reduce data acquisition time, thereby rendering it more practical for real-time feedback and facilitating its widespread utilization in image-guided surgical procedures.

When comparing the 3D CNN model and 1D models, it is evident that the 3D CNN consistently performs as well as or better than the 1D models, especially for the mucosa class. This suggests that including spatial information has a positive impact on the model's ability to make accurate classification decisions. However, the use of spatial information in the form of 5*5*100 patches for the 3D CNN classification has a noteworthy impact on the model's ability to detect small regions containing tumours, even though the overall model performance may not exhibit significant changes. This approach reduces the model's spatial resolution to 3.125mm/pixel from the original 0.125mm/pixel. Consequently, the model's ability to identify subtle tumour regions can be compromised, potentially leading to less precise detection of small or early-stage tumours. The resolution described in this section is a theoretical spatial resolution. However, in practical applications, several factors can worsen this resolution, including environmental conditions, data processing techniques, sampling frequency, and computational constraints. As a result, the actual spatial resolution achieved in practice may be lower, making the use of spatial dimensions in models even less desirable.

As shown in Figure 14, the results of the learning curve analysis reveal an interesting trend. All models demonstrate a slower rate of accuracy improvement after approximately 10 data points. This suggests that these models have reached a point where they understand the underlying patterns of their capacity and experience a plateau in performance. One plausible explanation for this observation is that the models used in this study might not possess sufficient complexity to capture more intricate patterns beyond a certain dataset size. It's possible that these models lack the ability to extract highly detailed features and exploit subtle relationships within the data. Consequently, adding more data may not lead to a significant improvement in model performance with the current designs. Additionally, it's crucial to take into account the specific characteristics of the dataset and the underlying patterns. Factors such as data homogeneity, redundancy, or the presence of a dominant pattern that the models have already learned could influence the saturation point where further data fails to improve performance.

The ground truth for tissue classification was established in two steps, involving annotations from an experienced surgeon and further refinement through pathological examination, as explained in the methods section. However, it's important to understand that the pathological examination was only performed on the resected areas to confirm them as tumour tissues and potentially as skin or mucosa on the edges. For healthy mucosa tissue and healthy skin tissue, the annotation was solely based on the surgeon's annotations. However, identifying healthy mucosa tissue posed a challenge due to the difficulty in visually distinguishing it from mucosa affected by cancer cells or other pre-malignancies. Consequently, many mucosa regions visible in the RGB images were not annotated,

resulting in a scarcity of mucosa tissue data. Moreover, the annotated mucosa regions consisted of mucosa from various parts of the vulva, as well as mucosa from prolapse regions. This diverse inclusion of mucosa subtypes could potentially have had an adverse impact on the model's ability to detect mucosa accurately.

The ultimate goal of this study is to develop an HSI-based automatic in vivo operative tumour detection system for clinical use. However, the AI models trained in this study are only capable of classifying small selected regions of interest containing the three predetermined tissue classes with overall F1 scores ranging between 0.78 and 0.81. During surgical procedures, a more comprehensive range of tissue types may be encountered, which extends beyond the specific classes considered in this study. As a result, the performance and accuracy of the trained AI models might not be optimal when tasked with analysing a complete HS cube. Figure 16 displays a heatmap indicating the probability of a tumour in a HS cube. The image shows that the model classified the expected tumour region as a high probability tumour region, as visible in the ground truth image. However, it also misclassified a part of the prolapse as a high-probability tumour region. Therefore, it becomes crucial to expand the training data by incorporating additional classes for commonly encountered tissue types, encompassing both healthy and pre-malignant conditions.

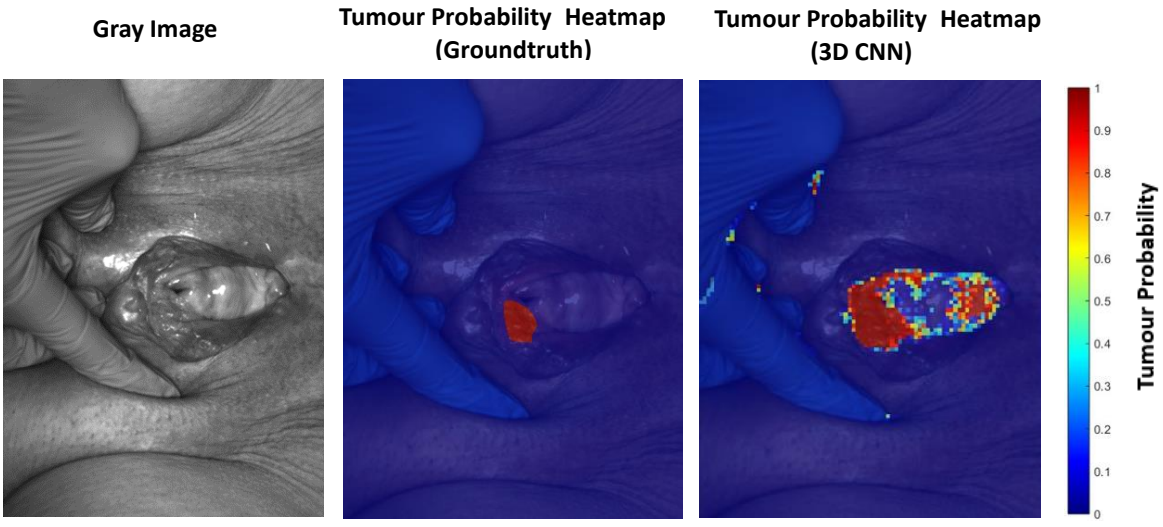

Figure 16: The tumour prediction heatmap shows values ranging from 1 (dark red) for high certainty in tumour classification to 0 (dark blue) for high certainty in non-tumour classification. Left: RGB image. Middle: Tumour prediction heatmap based on the ground truth. Right: Tumour prediction heatmap based on 3D CNN prediction.

Conclusion

In conclusion, this thesis successfully demonstrates the viability of in vivo detection of vulvar tumours through the integration of hyperspectral imaging and artificial intelligence techniques. Notably, all models demonstrated similar results for tumour detection, with F1 scores ranging between 0.91 and 0.93 and ROC-AUC scores ranging from 0.94 to 0.91. Furthermore, the NN model, which employed only physiologically significant fractions, exhibited particularly promising accuracy with an impressive F1 score of 0.93 and an AUC-ROC score of 0.91 for tumour detection. Additionally, the utilization of the NN model presents an opportunity to reduce the inherent spectral range of hyperspectral imaging systems without compromising detection accuracy or spatial resolution. This has the potential to reduce image acquisition time, data compression and potentially enable real-time image-guided surgery. However, it is crucial to acknowledge that further comprehensive investigations, incorporating more extensive and more diverse datasets, are imperative to validate and extrapolate these findings for broader clinical applicability.

Bibliography

[1] I. Alkatout *et al.*, "Vulvar cancer: epidemiology, clinical presentation, and management options," (in eng), *Int J Womens Health,* vol. 7, pp. 305-13, 2015.

[2] A. B. Olawaiye *et al.*, "FIGO staging for carcinoma of the vulva: 2021 revision," vol. 155, no. 1, pp. 43-47, 2021.

[3] T. P. Canavan and D. Cohen, "Vulvar cancer," (in eng), *Am Fam Physician,* vol. 66, no. 7, pp. 1269-74, Oct 1 2002.

[4] G. Baiocchi and R. M. Rocha, "Vulvar cancer surgery," vol. 26, no. 1, pp. 9-17, 2014.

[5] T. Nagaya, Y. A. Nakamura, P. L. Choyke, and H. Kobayashi, "Fluorescence-Guided Surgery," (in English), Review vol. 7, 2017-December-22 2017.

[6] M. A. Lediju Bell, "Photoacoustic imaging for surgical guidance: Principles, applications, and outlook," (in eng), *J Appl Phys,* vol. 128, no. 6, p. 060904, Aug 14 2020.

[7] H. C. van Doorn *et al.*, "Raman spectroscopy for guidance of vulvar cancer surgery: a pilot study," (in eng), *Biomed Opt Express,* vol. 12, no. 5, pp. 3008-3020, May 1 2021.

[8] E. J. Baltussen *et al.*, "Hyperspectral imaging for tissue classification, a way toward smart laparoscopic colorectal surgery," vol. 24 %J Journal of Biomedical Optics, no. 1, p. 016002, 2019.

[9] A. Ziemann, "A manifold learning approach to target detection in high-resolution hyperspectral imagery," 2015.

[10] G. Keiser, "Light-Tissue Interactions," in *Biophotonics: Concepts to Applications*, G. Keiser, Ed. Singapore: Springer Nature Singapore, 2022, pp. 169-221.

[11] P. Vaupel, A. Mayer, and M. Höckel, "Impact of hemoglobin levels on tumor oxygenation: the higher, the better?," (in eng), *Strahlenther Onkol,* vol. 182, no. 2, pp. 63-71, Feb 2006.

[12] L. Guolan and F. Baowei, "Medical hyperspectral imaging: a review," *Journal of Biomedical Optics,* vol. 19, no. 1, p. 010901, 1/1 2014.

[13] G. ElMasry and D.-W. Sun, "CHAPTER 1 - Principles of Hyperspectral Imaging Technology," in *Hyperspectral Imaging for Food Quality Analysis and Control*, D.-W. Sun, Ed. San Diego: Academic Press, 2010, pp. 3-43.

[14] Y. Wang, N. Reder, S. Kang, A. Glaser, and J. Liu, "Multiplexed Optical Imaging of Tumor-Directed Nanoparticles: A Review of Imaging Systems and Approaches," *Nanotheranostics,* vol. 1, pp. 369-388, 08/19 2017.

[15] F. Manni *et al.*, "Hyperspectral Imaging for Glioblastoma Surgery: Improving Tumor Identification Using a Deep Spectral-Spatial Approach," *Sensors*, vol. 20, no. 23. doi: 10.3390/s20236955

[16] O. Okwuashi and C. E. Ndehedehe, "Deep support vector machine for hyperspectral image classification," *Pattern Recognition,* vol. 103, p. 107298, 2020/07/01/ 2020.

[17] G. Urbanos *et al.*, "Supervised Machine Learning Methods and Hyperspectral Imaging Techniques Jointly Applied for Brain Cancer Classification," *Sensors*, vol. 21, no. 11. doi: 10.3390/s21113827

[18] Q. Zhang, Y. Wang, S. Qiu, J. Chen, L. Sun, and Q. Li, "3D-PulCNN: Pulmonary cancer classification from hyperspectral images using convolution combination unit based CNN," *Journal of Biophotonics,* https://doi.org/10.1002/jbio.202100142 vol. 14, no. 12, p. e202100142, 2021/12/01 2021.

[19] L. Ruiz *et al.*, "Multiclass Brain Tumor Classification Using Hyperspectral Imaging and Supervised Machine Learning," in *2020 XXXV Conference on Design of Circuits and Integrated Systems (DCIS)*, 2020, pp. 1-6.

[20] H. Martin *et al.*, "Hyperspectral imaging for head and neck cancer detection: specular glare and variance of the tumor margin in surgical specimens," *Journal of Medical Imaging,* vol. 6, no. 3, p. 035004, 9/1 2019.

[21] M. Nathan, A. S. Kabatznik, and A. Mahmood, "Hyperspectral imaging for cancer detection and classification," in *2018 3rd Biennial South African Biomedical Engineering Conference (SAIBMEC)*, 2018, pp. 1-4.
[22] M. A. Chandra and S. S. Bedi, "Survey on SVM and their application in imageclassification," *International Journal of Information Technology,* vol. 13, no. 5, pp. 1-11, 2021/10/01 2021.
[23] W. Xiangying and Z. Yixin, "Statistical learning theory and state of the art in SVM," in *The Second IEEE International Conference on Cognitive Informatics, 2003. Proceedings.*, 2003, pp. 55-59.
[24] A. Vidhya, "Suport Vector Mechanism (SVM): A complete guide for beginners," vol. 2023, no. 6-7, 2023.
[25] B. Jansen-Winkeln *et al.*, "Feedforward Artificial Neural Network-Based Colorectal Cancer Detection Using Hyperspectral Imaging: A Step towards Automatic Optical Biopsy," (in eng), *Cancers (Basel),* vol. 13, no. 5, Feb 25 2021.
[26] A. Krogh, "What are artificial neural networks?," *Nature Biotechnology,* vol. 26, no. 2, pp. 195-197, 2008/02/01 2008.
[27] A. K. Jain, M. Jianchang, and K. M. Mohiuddin, "Artificial neural networks: a tutorial," *Computer,* vol. 29, no. 3, pp. 31-44, 1996.
[28] D. Grigoryan, "Activation functions in neural networks," 6-7 2022.
[29] Q. Hao *et al.*, "Fusing Multiple Deep Models for In Vivo Human Brain Hyperspectral Image Classification to Identify Glioblastoma Tumor," *IEEE Transactions on Instrumentation and Measurement,* vol. 70, pp. 1-14, 2021.
[30] M. M. Taye, "Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions," vol. 11, no. 3, p. 52, 2023.
[31] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*, 2017, pp. 1-6.
[32] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," *ArXiv e-prints,* 11/01 2015.
[33] M. M. Taye, "Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions," *Computation*, vol. 11, no. 3. doi: 10.3390/computation11030052
[34] H. Latt, "CNN Deep Learning With Transfer Learning. Part 1," 6-7 2021.
[35] A. Holmer, J. Marotz, P. Wahl, M. Dau, and P. W. Kämmerer, "Hyperspectral imaging in perfusion and wound diagnostics - methods and algorithms for the determination of tissue parameters," (in eng), *Biomed Tech (Berl),* vol. 63, no. 5, pp. 547-556, Oct 25 2018.
[36] M. Rana, "BREAST CANCER DIAGNOSIS AND RECURRENCE PREDICTION USING MACHINE LEARNING TECHNIQUES," *International Journal of Research in Engineering and Technology,* vol. 04, pp. 372-376, 04/25 2015.
[37] M. Barberio *et al.*, "Deep Learning Analysis of In Vivo Hyperspectral Images for Automated Intraoperative Nerve Detection," (in eng), *Diagnostics (Basel),* vol. 11, no. 8, Aug 21 2021.
[38] Z. Han, A. Zhang, X. Wang, Z. Sun, M. D. Wang, and T. Xie, "In vivo use of hyperspectral imaging to develop a noncontact endoscopic diagnosis support system for malignant colorectal tumors," vol. 21 %J Journal of Biomedical Optics, no. 1, p. 016001, 2016.
[39] C.-L. Tsai *et al.*, "Hyperspectral Imaging Combined with Artificial Intelligence in the Early Detection of Esophageal Cancer," vol. 13, no. 18, p. 4593, 2021.
[40] D. Eggert *et al.*, "In vivo detection of head and neck tumors by hyperspectral imaging combined with deep learning methods," *Journal of Biophotonics,* vol. 15, 12/01 2021.