Universidade Estadual de Londrina

JESSICA FERNANDES LOPES

DUAL STAGE IMAGE ANALYSIS:

A COMPUTER VISION TECHNIQUE APPLIED FOR ANALYSING COMPLEX IMAGE PATTERNS IN RAW HAM VEINING DEFECT

Londrina
2019

JESSICA FERNANDES LOPES

DUAL STAGE IMAGE ANALYSIS:

A COMPUTER VISION TECHNIQUE APPLIED FOR ANALYSING COMPLEX IMAGE PATTERNS IN RAW HAM VEINING DEFECT

Dissertação apresentada ao Programa de Mestrado em Ciência da Computação da Universidade Estadual de Londrina para obtenção do título de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Sylvio Barbon Junior

Londrina
2019

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

L864 Lopes, Jessica Fernandes.
Dual Stage Image Analysis: A Computer Vision technique applied for analysing complex image patterns in raw ham veining defect / Jessica Fernandes Lopes. - Londrina, 2019.
83 f.

Orientador: Sylvio Barbon Junior.
Dissertação (Mestrado em Ciência da Computação) - Universidade Estadual de Londrina, Centro de Ciências Exatas, Programa de Pós-Graduação em Ciência da Computação, 2019.
Inclui bibliografia.

1. Aprendizado de Máquina - Tese. 2. Processamento de Imagens - Tese. 3. Reconhecimento de Padrões - Tese. I. Barbon Junior, Sylvio. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Ciência da Computação. IV. Título.

CDU 519

JESSICA FERNANDES LOPES

DUAL STAGE IMAGE ANALYSIS:
A COMPUTER VISION TECHNIQUE APPLIED FOR ANALYSING COMPLEX IMAGE PATTERNS IN RAW HAM VEINING DEFECT

Dissertação apresentada ao Programa de Mestrado em Ciência da Computação da Universidade Estadual de Londrina para obtenção do título de Mestre em Ciência da Computação.

BANCA EXAMINADORA

_____

Orientador: Prof. Dr. Sylvio Barbon Junior
Universidade Estadual de Londrina - UEL

_____

Prof. Dr. Alan Salvany Felinto Universidade Estadual de Londrina - UEL

_____

Prof. Dr. Fabrício Martins Lopes Universidade Tecnológica Federal do Paraná - UTFPR

Londrina, 26 de agosto de 2019.

ACKNOWLEDGEMENTS

Primeiramente, agradeço a Deus, por me manter e prover todas as condições e circunstâncias para realização deste trabalho.

Assim como, agradeço aos meus pais, Teresinha e Roberto, por sempre me guiarem e incentivarem. Não me faltam motivos para agradecê-los.

Agradeço aos demais que contribuíram e influenciaram nos caminhos trilhados.

Gostaria de agradecer a CAPES pelo suporte financeiro durante o decorrer da pesquisa realizada.

Agradeço ao meu orientador, professor Dr. Sylvio, por todo apoio e orientação no decorrer do mestrado. Assim como, agradeço a oportunidade e por compartilhar o conhecimento que possui.

Por fim, agradeço aos colegas de laboratório e estudo, que sempre me auxiliaram no decorrer do mestrado.

"Dreams determine what you want. Actions determine what you gain." - Aldo Novak

LOPES, J. F.. **Dual Stage Image Analysis:** a computer vision technique applied for analysing complex image patterns in raw ham veining defect. 2019. 83 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Estadual de Londrina, Londrina, 2019.

RESUMO

Lidar com reconhecimento de padrões requer a construção de métodos robustos e flexíveis de modo a propiciar alta performance na identificação de diferentes amostras para os mais variados problemas. O sistema de visão computacional é utilizado na análise de vários processos em que, associa-se algoritmos, para predizer um conjunto de imagens, simulando a visão humana. Na indústria alimentícia, o controle de qualidade requer técnicas precisas que proporcionem maior confiabilidade no processo. Considerando o processo industrial de presunto crú, a presença de veias no pernil suíno está diretamente relacionada com a qualidade do produto final, consequentemente, em seu valor de mercado. Alguns defeitos relacionados a presença de veia são detectados através de inspeção visual. No entanto, uma avaliação automática da qualidade de presunto crú através da análise de imagens permeia desafios para tradicionais técnicas de visão computacional devido ao padrão complexo relacionado a cada nível de defeito. Sendo assim, estratégias para distinção entre diferentes classes são necessárias visto que a extração de características torna-se ineficiente para se obter um bom desempenho. Dentro deste contexto, esta dissertação objetiva propor o *Dual Stage Image Analysis* (DSIA), uma técnica baseada em visão computacional para propiciar melhor performance na identificação de classes baseadas na quantificação de características determinantes presentes nas amostras. Por meio do princípio de "dividir-para-conquistar", o DSIA propõe a divisão do problema em subconjuntos para resolução de subtarefas de modo a propiciar uma melhora significativa no entendimento entre as classes a serem preditas. Para avaliar a performance da abordagem proposta, foram realizados comparativos entre o método Tradicional de sistema de visão computacional com o DSIA. Após observar diversas análises, nota-se que a técnica proposta é capaz de aumentar, significativamente, a capacidade preditiva, obtendo acurácia **25%** maior quando comparada ao método tradicional.

Palavras-chave: Aprendizado de máquina.reconhecimento de padrões. Sistema de visão computacional.

LOPES, J. F.. **Dual Stage Image Analysis:** a computer vision technique applied for analysing complex image patterns in raw ham veining defect. 2019. 83 p. Master's Thesis (Master in Computer Science) – Universidade Estadual de Londrina, Londrina, 2019.

ABSTRACT

Dealing with pattern recognition requires the construction of robust and flexible methods to provide high performance in identifying different sample classes. The computer vision system is used in the analysis of several processes in which, it is associated algorithms, to predict information set. In the food industry, quality control requires precise techniques that provide accurate process. Considering the industrial process of Dry-cured ham, vein in the pork's thigh carcass are directly related to the quality of the final product, consequently to its market value. Some veining defects are easily detected by visual inspection. However, the automatic evaluation of raw ham quality by image analysis poses some challenges to the traditional computer vision technique grounded on the complex image pattern related to each defect level. Thus, strategies for distinguishing between different classes are often necessary since feature extraction becomes inefficient to achieve good performance. In this context, this dissertation aims to propose Dual Stage Image Analysis (DSIA), a technique based on computer vision to provide better performance in the identification of classes based on the quantification of determinant characteristics present in the samples. Through the divide-to-conquer principle, DSIA proposes to split the problem into subsets to solve sub-tasks to provide a significant improvement in understanding between classes to be predicted. To demonstrate the performance of the proposed approach, comparisons were made between the traditional method of computer vision system with the DSIA. After observing several analyses, it is noticed that the proposed technique is able to increase the predictive capacity significantly, the obtained accuracy is higher at **25%** in comparison to the traditional method.

Keywords: Computer vision system. Machine learning. Pattern recognition.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AI	Artificial Intelligence
CVS	Computer Vision System
DFT	Discrete Fourier Transform
DSIA	Dual Stage Image Analysis
FFT	Fast Fourier Transform
FN	False Negative values of Confusion Matrix
FP	False Positive values of Confusion Matrix
GLCM	Grey Level Co-occurrence Matrix
HLS	Hue Saturation Lightness channels
HSI	Hue Saturation Intensity channels
HSV	Hue Saturation Value channels
LBP	Local Binary Pattern
ML	Machine Learning
PCA	Principal Component Analysis
PDF	Probability Density Function
PQI	Parma Quality Institute
RBF	Radial Basis Function
RF	Random Forest
RGB	Red Green Blue channels
ROI	Region of Interest
SRM	Structural Risk Minimisation
SVM	Support Vector Machine
TN	True Negative values of Confusion Matrix
TP	True Positive values of Confusion Matrix

CONTENTS

1 INTRODUCTION

The industrial process requires precisely quality control for final products. Although the basis of quality assessment is often based on visual inspection, which is considered a subjective method being based on human attributes perception, such as appearance, smell, texture and flavour. Thus, the production environment has increased the demand for enhanced quality monitoring, which has propitiated research for automatic inspection systems based on image analysis [9].

Quality evaluation can be performed by a Computer Vision System (CVS). It is based on hardware and software for digital image processing (digital camera, inexpensive and broadly available) for providing an automatic classification [10, 11]. Machine vision presents several advantages, as rapidity, low-cost and accuracy and can be applied to a diverse industrial process. Therefore, this type of approach is suitable for many tasks, being considered a non-invasive method, which aims high accurate results [12, 13, 14]. Image classification has received increasing attention related to computer vision targets [15, 16]. Several image tasks are complex since the object of interest cannot be described based on small and meaningful features set. In this way, strategies and methodologies based on image analysis and CVS are developed to achieve higher performance, and generalised approaches capable of handling varied environmental conditions [17, 10].

The agri-food industry requires an effective evaluation of the production process to establish the quality of the final product [13]. Thereby, CVS has been a meaningful application in order to improve food quality assessment and control. Considering the applicability of quality control process, the analysis of meat quality measurements in the industrial process involves multivariate characteristics, such as breed, genotype, feeding, pre-slaughter handling, stunning, slaughter method, chilling and storage conditions. The analysis of pork properties has relevance for the suitability of the meat for further processing. Thus, the most popular technique assign grades or quality labels in pork meat is the visual inspection performed by specialists [18, 19, 10].

Pork research has lagged behind efficient methodologies for computer vision-based evaluation of quality. Quality classification would ensure that various components were used in the most suitable product. For reducing product variation and remain competitive in a quality-sensitive market, the industrial process has to reach an objective quality measurement for pork is essential [20]. More specifically, in dry-cured ham production is required control of specific quality properties which interferes in production steps. One of the most important properties is the globosity index, as there are limits suggested by the literature for being respected, which should not exceed [21, 22, 23]. Moreover, research studies have shown the presence of veins in pork's thigh carcass influences in the adipose

cover and globosity index, both representing essential quality measurements in dry-cured ham industry [22]. Therefore, it is essential to highlight the veins presence, as well as red bark appearance, are defects that characterise a low-quality sample [21, 23].

Concerning meat products, the analysis of appearance is considered a complex task which involves animal genetics, ante- and postmortem conditions, fundamental muscle chemistry, and factors related to meat processing to final production. Therefore, strategies and methodologies with digital imaging techniques are applied to improve the available measures [15]. O'sullivan et al. [24] compared instrumental colour measurements and features extracted of the colour channel as RGB and CIE L*a*b* to evaluate the pork meat quality. The comparison was proposed to analyse the effectiveness of the sensory visual quality prediction, resulting from computer vision method in a high correlation and performance. Shiranita et al. [25] implemented grading systems based on image processing to analysis the meat quality. The conception evaluated through the marbling score showed effectiveness as observing experimental results.

Nowadays, CVS are valuable tools for many tasks. There are several applications of Machine Learning (ML) able to build automatic classifiers. In other words, ML creates a classifier that can be used to predict outcomes from new instances [26]. Thus, a large number of techniques have been developed based on Artificial Intelligence to resolve the problems using CVS [20, 15, 27, 28, 29, 30, 16, 10]. Considering variability of samples as the environmental conditions such as different visual characteristics, strategies are essential to handle of handcrafted solutions for each scenario [10]. Frequently, the image analysis by ML involves robust techniques due to the complexity in distinguishing the different classes. Dealing with pattern recognition is a challenge considering the structural description of image samples supported by image features [17].

Concerning visual defects, such as veins presence and their structure, it provides an increase difficulty for sample classification by a CVS [31]. The veining defect in raw ham is usually localised under the skin and it can affect a large surface of the carcass. Most skin lesions cause the highest percentage of trimming, as veining defects, resulting in a degrading of the product quality [32]. Considering the standards of acceptable dry-cured ham for the market, low-quality products are those with a high presence of veining defect, which can be identified with different levels of vein quantity. Dealing with image representation of veining defects involves the use of complex features for their description. These complex image patterns demand an effective CVS, including the ability to identify the relevant properties to discriminate each class of veining defects.

To provide a suitable classification of veining defect level, in this work the Dual Stage Image Analysis (DSIA) fashioned over the "divide and conquer" strategy is proposed. The proposed approach aims to split the Region of Interest (ROI) into sub-regions toward extracting micro-descriptors, which is adopted in order to support a robust im-

age pattern recognition in complex pattern scenarios. The First Stage is carried by ML recognition over each sub-region. The Second Stage provides the final decision based on a voting strategy. The method was evaluated embedded in a CVS with a set of image features based on colour, intensity, histogram analysis, border and texture in comparison to traditional CVS [20, 27, 28, 29, 30, 10]. To evaluate the First Stage, the experiments were analysed applying two different machine learning algorithms: Random Forest (RF) and Support Vector Machine (SVM). These methods were applied to discriminate three different veining levels among raw ham samples. The obtained results have shown the DSIA achieved superior performance considering traditional CVS.

1.1 Hypothesis

Automatic evaluation by image analysis has to deal with challenges to Computer Vision Systems as complex image pattern. For this reason, strategies are required to improve predictive performance [15, 16, 10]. Concerning raw ham quality evaluation, each defect level is grounded on complex image pattern tasks, which adds difficulty to represent those characteristics. To do so, decomposing the image classification task into sub-patterns and to develop rules by which these sub-patterns are interrelated can propitiate to a high predictive performance. Nonetheless, the hypothesis is the following:

"It is possible to improve the CVS classification performance without overburden feature extraction by splitting images into sub-regions to learn about different parts of a scenario."

1.2 Objective and Contributions

In this work a technique called Dual Stage Image Analysis is proposed, which is based on computer vision system to improve predictive performance for scenarios with complex characteristics.

The main contributions of this work can be summarised as:

1. Proposing the DSIA which improves predictive performance in comparison to Traditional CVS based on "divide and conquer" strategy;
2. Comparing two machine learning algorithms as kernel of First Stage;
3. Applying the DSIA in a raw ham dataset classification considering predictive performance.

1.3 Outline

The remainder of this work is divided as follow: Chapter 2 presents a theoretical foundation about CVS and traditional ML, including the general idea of Support Vector Machine and Random Forest algorithms. Chapter 3 presents, as related work, image classification tasks with CVS. Emphasising, the classification problems related to food industry which involves quality evaluation. In Chapter 4, the DSIA technique is presented in details. Materials and methods are discussed in Chapter 5. Results obtained are presented in Chapter 6, as the discussion about the performances evaluated. Chapter 7 presents the Conclusion. Lastly, the references which were used to support this work were listed.

2 COMPUTER VISION SYSTEM

Human perception is considered a challenge for computer tasks. Computer vision can be considered a data transformation from digital images for making a decision [33]. Thus, to reach the goals from image analysis, it is necessary to construct models with object descriptions for pattern recognition, which provides essential features from scenario based on images [34].

Dealing with systems which replicate the capacity of the human visual system is a challenge for the CVS. Through the image analysis, seeking out to obtain descriptions which contain information to distinguish the ROI without human intervention. Thus, computer vision assists in solving complex problems based on human cognition and decision-making ability [35]. CVS is defined as the science which characterises image interest information. CVS consists of four main steps: Acquisition, Image Pre-processing, Feature Extraction and Classification [1], as showed in Figure 1.

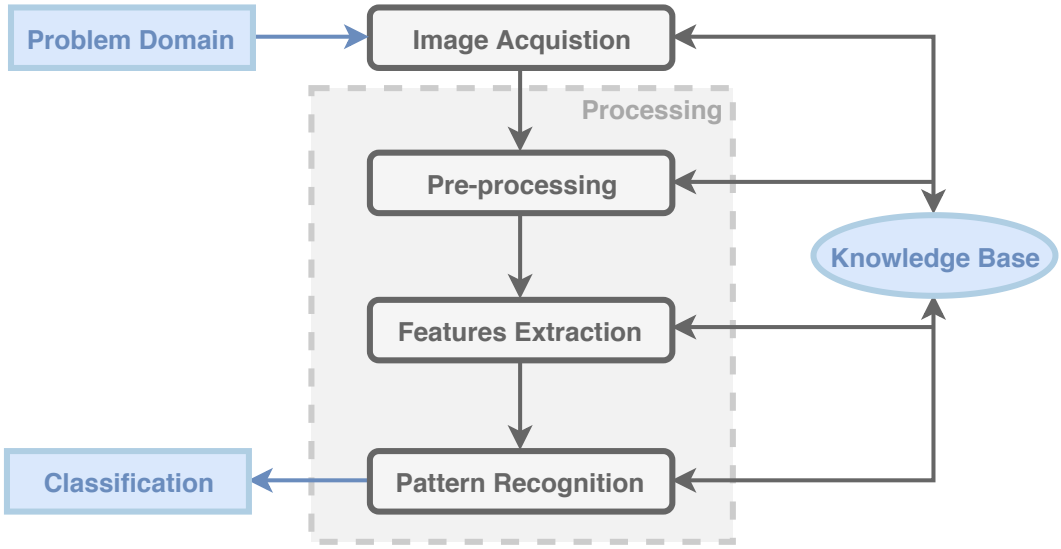


Figure 1 – General overview of Computer Vision System. Image adapted from [1]

The digital image is composed of finite elements, each with specific value and position. So, it provides a matrix of numeric values which can be analysed and processed [36, 37]. The computer vision aims to develop tools able to determine scenarios and properties observing datasets. The knowledge about the problem domain is coded in a Digital Image Processing system through a composed base which leds to the feature analysis [2]. In this way, the typical steps of CVS can be briefly described as:

1. *Acquisition* - to perform the images acquisition, it is necessary a physical device sensitive to the energy radiated by the object and a digitising device to convert the

output of the physical devices into a digital format. Thus, it captures the image through the device and converts it into a suitable representation for subsequent digital processing [38, 3].

2. *Pre-processing* - aims to improve image quality which includes methods to facilitate the identification of an object [3]. Segmentation step are performed to extract and to identify an area of interest. Usually, it is based on discontinuities detection (edges) or image similarity (regions). It seeks to isolate image regions belonging to objects for after feature extraction and descriptive parameters calculation [3, 37].

3. *Features Extraction* - from the segmented images, it is sought to obtain relevant data or attributes of the highlighted regions or objects. It aims at extracting characteristics which can be used in the discrimination between objects classes [39, 2].

4. *Pattern Recognition/Classification* - consists of distinguishing the image classes according to its similarity as belonging to a previously established group or class. The decision is made from analysis of the features extracted from images. The recognition process assigns an identifier or label to the objects basing on its descriptors. The process of interpretation consists in assigning a meaning to the recognised objects set [40].

Abstraction levels on CVS are usually divided into three categories: low level (acquisition and pre-processing); intermediate level (segmentation, feature extraction and characteristics normalisation) and high level (classification and pattern recognition). The purpose is to get tools able to understand a particular scenario or context characteristics. For all the steps, the goal is to create a system that models or represents a corresponding input image in the knowledge base through the interpretation, decision or description of the classifiers [35]. Considering CVS targets, through digital image processing, it aims to analyse images to obtain results based on human perception. Thus, computer vision is a technique based on image feature extracted for pattern recognition.

2.1 Image Acquisition System

The simplified physical model of the image acquisition is based on a camera device which is situated between the image plane and the three-dimensional scene environment. Any light rays emitted or reflected determine the surface through the object before being projected in the digital image plane [41]. The signals representation (images) can be made by symbols in the objects identification process, being described as a continuous function. Mathematically, the Real function $\mathfrak{R}$ has a value for each provided x having a single

value of *f(x)* obtained [42]. Formally, the relation between image dimension is defined by Equation 2.1.

$$\begin{aligned} f &: \Re \rightarrow \Re \\ x &\rightarrow f(x) \end{aligned} \tag{2.1}$$

Computer image processing consists in making discrete points along x towards storing the $f(x)$ corresponding value. The domain function discontinuation on the x axis is termed as sampling, which is composed of equal intervals. Sampling is uniformly spaced since each sample is drawn at equal intervals for coordinate values. The digital image discontinuity representation needs to be converted into matrix values. The sampling spatial coordinates form the image Cartesian Plane projection whose components (x, y) are represented by a finite numbers points (pixels). The image projection makes two-dimensional the real graphic of model [2]. A grey and two-dimensional image has only one band ($k = 1$) in which the *spels* are called pixels (picture elements) [37].

(a) Original image (b) Low resolution

Figure 2 – Split the pixels into a coloured image. Image adapted from [2]

Pixels are defined as the distribution and quantity variation being directly proportional to the spatial obtained information amount. Each pixel is associated with its quantified $f(x,y)$ value symbolised by small squares in the image, as exposed Figure 2. A discontinuous image is presented as a pixels array (M x N), where M is the length (rows number) and N is the width (columns number), respectively. The pixel values $p(x, y)$ positive integers indicate the intensity colour at each image position [x, y]. Thus, the image resolution refers the interval for space region, the smaller sampling corresponds to a minor resolution then larger interval for the same space refers a better one [37].

2.2 Image Pre-processing

The image pre-processing step is responsible for highlighting and restoring operations. The techniques application is fundamental for subsequent images analysis [37]. Several global operations can be performed in the pre-processing step, considering the spatial domain, which aims to improve image quality [43].

Using digital images requires the environment must be prepared for minimising insufficient lighting influences, dealing with dynamic scale limitations or adjust in an improper setting of the lens aperture during image acquisition. Incorrect procedures can result in poor image quality and low contrast [37]. Image pre-processing enhancements are used to optimise for specific feature measurement methods, for instance, sharpening and colour balancing. In this way, this step is dependent on the application and descriptor using the images analysis [43].

Considering steps of image pre-processing, the segmentation defines regions with internal similarity identifying the ROI. The image segmentation process refers to the step which joins transformation techniques are applied for reducing noises and extracting the ROI from the image. ROI identification allows the pattern recognition for later interpretation [44]. The interest elements are separated from the complete image scene for posterior information extraction supports the classification, as shown in Section 2.2.1.

2.2.1 Segmentation

Respecting the most important step of image processing, the segmentation consists of gathering image processing operations, like noise reduction and object of interest for pattern recognition by scenarios interpretation. The segmentation process is considered the first step of pre-processing images, which can be based on various properties of intensity, as discontinuity or similarity pixels value.

Segmentation subdivides an image into regions or objects that compose it by being guided by the problem to be solved. This task is considered difficult; accuracy and success are demanded at the end of the procedure. Through the pixel discontinuity, properties with the rest of the image elements partitioning are extracted. As well, homogeneous regions with similar properties are identified. Among the traditional techniques used in segmentation, there is the region growing technique by pixel grouping, threshold and split and merge process, which is based on similarity per pixel neighbourhood [45].

The segmentation purpose is join image parts that probably belongs to the same structure and split into parts or constituent objects. The segmentation step needs to be robust, as it substantially favours the final solution, if the segmentation is poorly performed this almost always determines in the final processing failures. The segmentation complexity is determined by the object of interest [4]. One of the techniques for image

segmentation is through edge detection. Edge is defined as the boundary or boundary between two regions with relatively distinct grey level properties. Thus, in image segmentation task in scenes which present sufficiently homogeneous grey levels, the transition between two regions and consequent identification can be determined simplifying by the grey levels discontinuity.

The region growing technique is based on indicating one or more pixels that will be used to determine the intensity value in comparison to other pixels. If multiple points are selected, the average is usually calculated to be used. The values pixel intensity in each neighbourhood is compared to the centre [4]. Figure 3 shows an example of the progression of the region growing technique. The object of interest to be considered is the blank spaces, as shown in Figure 3. The four pixels sides are analysed: left, right, top and bottom. The letter C refers to the central pixel and the letter N the neighbours. The sequential regions expose marked the neighbourhoods analysed and selected, as observed in Figure 3(a). If the difference is within a predefined threshold, the pixel is added to the evaluated region. Otherwise, this is discarded. The process continues as long as there are pixels to be parsed. The sequence of six images and their respective pixel position are observed, as shown in figures 3(b)-(d). The figures corresponding to 3(e) and 3(f) are sequences which shown no homogeneity.

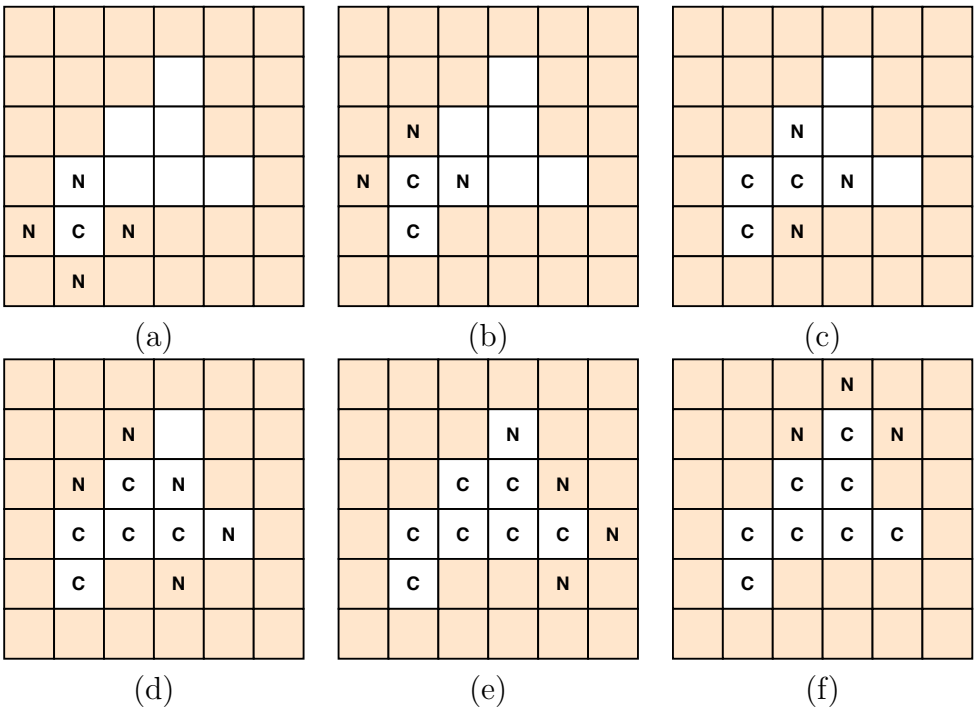
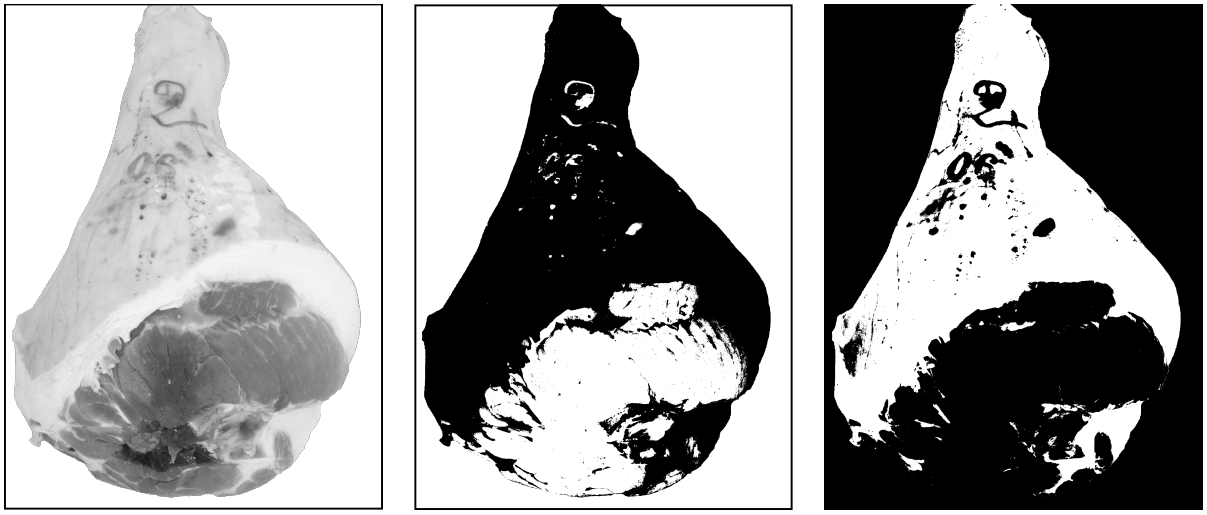


Figure 3 – Region growing graphical representation. Image created by the author.

The thresholding technique is performed for image segmentation using one or more grey level threshold values. The values are usually obtained through the histogram when dealing with greyscale images. After image histogram analysis, a threshold value is chosen and the pixels which have an intensity value smaller than the threshold are assigned white

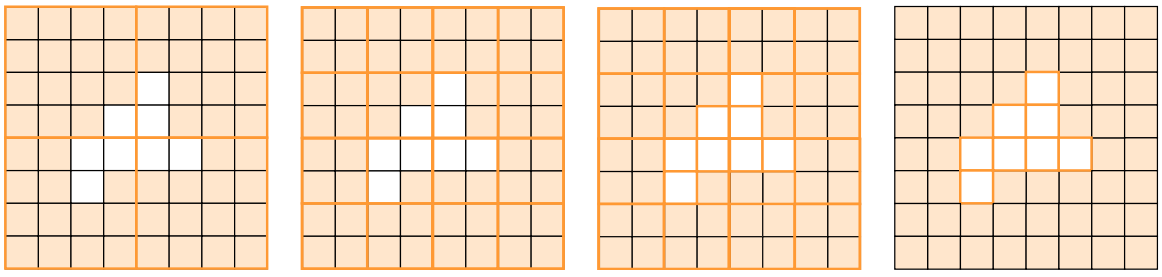
value and the remaining pixels the black value (binary image). This method allows simple separation between objects and the image background through varying intensity values [4]. Figure 4 shows two examples of image thresholding. In Figure 4(a) the threshold of 150 and in Figure 4(b) were delimited above threshold 184. The difference between the delimited regions can be observed. The technique challenge is to find the threshold value.



(a) Original grey image (b) Less 150 (c) Above 184

Figure 4 – Thresholding segmentation with two different threshold values. Image adapted from [2]

Another technique for image segmentation is split and merge. This is based on the image homogeneity. The split and merge considers as a homogeneous region all pixels that have similar or equal intensity to an average intensity value. Otherwise, the region is not homogeneous, it is partitioned into four smaller regions. Each region is analysed at either iteration, in case non-homogeneity stills the partition process occur again [46]. The sequence process can be observed as Figure 5. The highlighted lines refer to the regions representing the partitioning sequence and subsequent merge. Behind the splitting image process, the merge process is based on the combinations of adjacent regions whose average pixel value differs is below threshold or regions are too small.



(a) Fist split (b) Second split (c) Third split (d) Merge

Figure 5 – Split and merge sequence process. Image created by the author.

Concerning colour images, the split into object of interest and background in im-

age pre-processing can be procedure based on colour properties of the samples images. This approach consists in converting each RGB (Red, Green and Blue channels) image of the dataset into a 4900-points long signal, namely the colourgram [47, 29, 48, 49]. More in detail, colougrams are one-dimensional signals that codify the whole colour-related information of an RGB image, which represents properties of images for subsequent analysis. Basically, it propitiates the identification of the colour-related aspects of the ROI, as colourgrams allows to evaluate in a critical manner the forms of images to be segmented. It must be highlighted the colourgrams is performed without requiring image pre-processing [8].

Colourgram considers every single RGB image with size $\{r, c, 3\}$ (r is the number of pixel rows, c is the number of pixel columns and 3 refers the R, G and B channels) for unfolding its into two-dimensional matrix $\{(r \text{ x } c), 3\}$. Therefore, each of the three channels can assume integer values in the greyscale images, considering the range 0-255. Considering the steps for region detection, each colourgram is obtained by merging the frequency distribution curves of the R, G, and B channels. Additionally, the colour parameters derived from those channels are considered: the lightness (L), which is calculated from a sum of the RGB channels; the relative colours of each channel, i.e., it is calculated by ratio between each channel and lightness obtaining the relative red (RR), relative green (RG) and relative blue (RB) values; the hue (H), saturation (S) and intensity (I) obtained by converting the RGB coordinate into the HSI colour space; lastly, the scores, which is obtained by applying the PCA (Principal Component Analysis) to the raw, mean centred and auto-scaled RGB matrix [8].

Regarding previously steps, the acquired RGB images are converted into colourgrams and the signals are visually inspected in order to identify the colour parameters for leading to the optimal segmentation for the object of interest. Analysing colour parameters based on colourgram approach, it allows to remove the pixels related to background and isolates the ROI. Basically, it consists of finding a threshold value for segmenting images [8, 11].

2.3 Image Feature Extraction

Characteristics are grounded in quantify useful information based on image analysis toward pattern recognition. Object characteristics are represented through numerical values encoded into a feature vector. Feature extraction, as shown in Figure 1, propitiates interpretative information present in the image for ML algorithms. Most of the pattern recognition by images requires the local or structures isolation is necessary. Usually, these procedures are performed manually, for the automatic development methods, the analysing stage and recognition is hugely relevant [39].

In CVS, there is a vast type of descriptors which can be representing image properties, as colour, texture, edge, among others. Thus, pattern recognition encompasses quantitative pattern descriptors for finding a pattern is based on the descriptors arrangement composed of feature sets of a given class. CVS identifies objects through descriptors, playing as a search key for the content-based image retrieval [2].

In this work, 92 image-base features were applied for inducing ML algorithms. The amount of descriptors can be divide, as: Colour, Intensity analysis, Histogram analysis, Border and Texture. The concepts that cover each cited group are detailed in the next sections.

2.3.1 Colour Spaces

Grey level images are composed by a single value per pixel which is called intensity or brightness, as previously mentioned. The intensity represents the amount of light reflected or emitted by an object. Thus, pixels can represent multiple values for light at different frequencies or colours [37].

A colour space can be described as a method for expressing the image aspects. Colour is the perception resulted of the light that penetrates the retina in photo-receptor cells [37]. The colour information is encoded in relation to light intensity and colour differences, called luminance. Thus, it is very relevant attends to distinguish objects in the image allowing spatial acuity and information that can be quite compacted. Regarding colour models, the representation is based on the relationship between coloured light and perception. Light may be associated an electromagnetic wave. Therefore, the waves hit an objects and light frequencies are absorbed, likewise, some other light frequencies are reflected. It process is perceptive with the eyes and thus creating what we observe as colours. Similarly, it occurs with the camera's sensors. The reflected light towards the sensors which obtains an intensity measure by adding energy on a frequencies range [2].

The colour can be decomposed into three components: hue, saturation and intensity. Summarising, the intensity refers to the brightness sensation, the hue by the colour sensation which it is connected by the wavelength and the saturation refers the colour purity degree in relation to the white. Coloured images are stored of three components forming a colour space [37]. In colour image processing, the values are obtained by selecting different frequencies. Colour models gives meaning to this information by organising colours in a way that can be related to the colours we perceive. Thus, a colour model is defined as a three-dimensional image representation in which each colour is specified by a point in the coordinate system. There is no model that independently describes all colour aspects, so different models are used to aid the characteristics description.

There are many different colour spaces proposed in the literature, each with its own properties, as RGB, HSI, CIE colour spaces, among others [50]. Due to its simplicity,

the RGB colour space is most commonly used. It is composed by red (R), green (G) and blue (B) chromaticities. Basing on the *Tristimulus Colour Theory* additive system, the human eye perceives colour through the three visual pigments stimulation present in the retina cones. The stimuli have sensitivity peaks approaching wavelengths 700.0 $m\mu$ (red), 546.1 $m\mu$ (green) and 435.8 $m\mu$ (blue). Thus, the RGB model was defined [37].

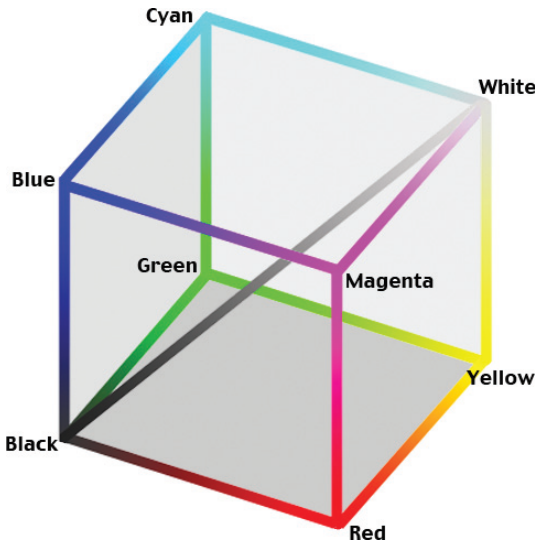


Figure 6 – RGB colour space with combining primary colours. Image from [3]

Inbuilt hardware operations, the RGB colour space can be obtained through a cube, according to Figure 6. Achieving a certain colour a pre-specified range is used defining 0 to 255. The black colour is obtained by the combination (0, 0, 0) and the white colour (255, 255, 255). For the representation of the other tonalities, channel R is represented by red (255, 0, 0), V (green) by (0, 255, 0) and B (blue) by (0, 255). Likewise, other combinations are composed of the primary colour complements.

Defining by signals, RGB colour space is represented with Cartesian cubic space it is composed since the red, green, and blue signals. The colours are independent and can be added to produce any colour within the cube [3], as previously mentioned. Combining colour channels it is possible to produce other colours tonalities as exposed in Figure 6. Varying Red and Green produces Yellow, Green adding Blue allows Cyan, and Blue with Red produces Magenta, it colours are defined by subtracting primaries colour printed. Thus, the primary colours allow the other colours composition. Colours named as secondary are the primary colours combination in equal proportions.

Orienting in human conception the HSV, HSI and HSL coordinate system are embodied on this approach. These are related to each other and are defined as Hue (H), Saturation (S), Value (V) from HSV, Hue (H), Saturation and Intensity (I) from HSI and Hue (H), Saturation (S) and Lightness (L) from HSL coordinate systems. The colour space concepts are associated with tint, shade and tone which is defined Hue by colour instance, Saturation is the amount of presented colour, and Lightness/Intensity/Value is

light quantifying as dark or light colours [3]. Figure 7 exposes the schematically thought of colour space.

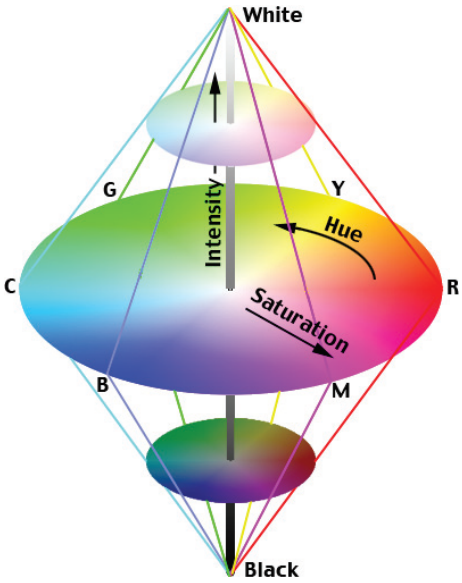


Figure 7 – HSI colour space. Image from [3]

The values are plotted in circular or hexagonal cone (cylinder). The cone axis is the greyscale progression showing the black to white. The distance from the central axis is defined from saturation and the direction is the hue. Primary RGB colours space as observed and corresponding the R (red), G (green) and B (blue). The combination composition is related as secondary as C (cyan), M (magenta) and Y (yellow). The HSI colour space has advantages for image processing and colour knowledge. The HSI components do not cause colour changes in the results by applying the channels separately. As well as the hue helps to distinguish the feature for segmentation. For example, HSI colour space disregards shading effects making processing more effective. However, the HSI components do not correspond to most hardware used from image acquisition. Thus, the RGB images conversion to HSI and back is required [37].

2.3.2 Entropy

Entropy is defined as the macroscopic organisation measurement for thermodynamic systems [51], being used to measure the uncertainty degree set. Thus, the entropy measure is directly related to pattern uncertainty, the more uncertain the particular patterns set, the greater entropy value. Information can be represented as a probabilistic process. The event occurrence i can be modelled with the p_i probability close to zero or one values. The entropy closest to zero is considered when the uncertainty degree of events is low. Similarly, the occurs probability is inversely proportional.

In image processing, the entropy corresponds to the pixel probability p_i in digital images assuming a certain intensity value i. The probability density function of the image

is analysed through the intensity histogram levels being divided into the pixels number of intensity i by the pixels number of image (N) [52]. Equation 2.2 shows the entropy measure.

$$\begin{aligned} H &= -\sum_{i=1}^{G} p_i.logp_i \\ \sum_{i=1}^{G} p_i &= 1 \\ p_i &= \frac{g_i}{N} \end{aligned} \tag{2.2}$$

By definition, the entropy H is based on the pixels amount g_i with intensity i, as G refers the grey levels quantity in the image (channel) [52]. The smallest entropy value is found (zero) in situations where all pixels are the same intensity (there is no uncertainty). Otherwise, maximum entropy is obtained when an image contains the same pixel amount for all intensities. That is, all intensities have the same occurrence probability [53]. It is important to emphasise the entropy is not related to the information spatial arrangement.

2.3.3 Histogram

The histogram is used to obtain image quality indicators regarding the contrast level. The intensity histogram exhibits individual brightness levels in an image. Figure 8 gives an example of image histogram.

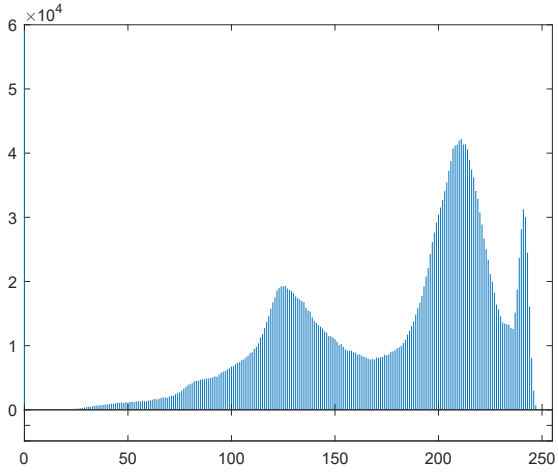
Concerning plotted histogram, it shows the brightness ranges from 0 (black) to 255 (white). Observing the Figure 8(b) most of the grey levels are used and the lowest and highest intensity levels are distinct peaks demonstrating dark portions of the image, such as the brighter points. Instead, Figure 8(d) contains overall histogram points concentrated toward black levels as dark image.

Several features such as standard deviation, skewness and kurtosis could be computed based on the first order histogram. Given a greyscale image $\tilde{\mathrm{I}}$, it can be stated which the function $h(l)$ that produces the number of occurrences of each image grey frequency. The probability distribution representation of the pixel values is computed by Equation 2.3.

$$P_y(r_k) = \frac{n_k}{n} \tag{2.3}$$

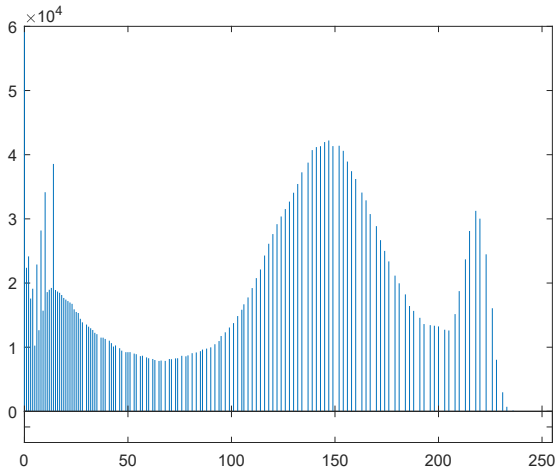
The correspondent value r_k represents the distribution of the normalised histogram being $0 \leq r_k \leq 1$, n_k is the pixels number corresponding to the k grey level, n represents the total pixels number product in the image (spatial resolution $M.N$), and $k = 0, 1, \ldots, L\text{-}1$ which L is the number of grey levels of the image.

(a) Original grey image



(b) Histogram of image

(c) Dark grey image



(d) Histogram of dark image

Figure 8 – Image and its histogram. Image adapted from [2]

The mean is computed by Equation 2.4.

$$S_{mean} = \bar{r_k} = \sum_{r_k=0}^{L-1} r_k.P_y(r_k) \tag{2.4}$$

The pixels average intensity are denoted by Standard deviation Equation 2.5.

$$S_{std} = \sigma_{r_k} = \left[\sum_{r_k=0}^{L-1} (r_k - \bar{r_k})^2.P_y(r_k)\right]^{1/2} \tag{2.5}$$

Considering statistical analyses, some fundamental measures are characterisations of location and variability set which includes skewness and kurtosis. Skewness is a measure of asymmetry assuming positive or negative values. Equation 2.6 shows computed

skewness.

$$S_{skw} = \frac{1}{\sigma^3_{r_k}} . \sum_{r_k=0}^{L-1} (r_k - \bar{r_k})^3 . P_y(r_k) \tag{2.6}$$

Kurtosis is defined as any measure of probability peak distribution. That is, high kurtosis measure tends to have outliers, on the other hand, low kurtosis tends to have lack of outliers. Similarity, kurtosis is a descriptor of probability shape distribution (Equation 2.7).

$$S_{kurt} = \frac{1}{\sigma^4_{r_k}} . \sum_{r_k=0}^{L-1} (r_k - \bar{r_k})^4 . P_y(r_k) - 3 \tag{2.7}$$

The respective Energy is exposed as Equation 2.8.

$$S_{energy} = \sum_{r_k=0}^{L-1} (P_y(r_k)^2 \tag{2.8}$$

The texture uniformity is described by Energy. Dominating grey tone transitions at image shows homogeneous when there are very few variance. Therefore, the image has fewer entries which suggests large magnitude and high energy.

2.3.4 Texture

Texture descriptors are relevant attributes for image analysis in studied of CVS. Most surfaces offer unique textures that can be discriminated in an interpretation procedure. Briefly, the texture reflects variations and optical properties in an object, just as on surfaces, different textures are produced. Therefore, texture is an important information resource for problem tasks like classification and segmentation. Essentially, there is no unique definition of texture and there are many ways to describe and extract it [2]. Descriptors of texture have many definitions depending on the specific application.

Broadly, some descriptors are related to texture visual perception derivating features measurements to describe them. Spatial intensity variation of texture conceptions can be related to structural (transform-based), statistical, and spectral approaches. Accordingly, the image frequency reflects its texture [2]. Structural approaches are the most basic texture description computed through the Fourier Transform. The texture is composed of patterns and repetitions being described with rules capable of generating and reproducing itself. Thus, the texture perception could be modelled using fractal analysis of images. In fractal methods, one measures the degree of a given image is represented by an original line repeated several times to generate textures.

Statistical methods for texture describe statistical rules relate to grey levels distribution and spatial relationship of image. The most famous statistical approach is the co-occurrence matrix [54]. The second-order features accumulated in the matrix measure the grey levels of spatial dependence on the image. A set of descriptors can be extracted from the matrix and form an image features vector. This descriptor presents a good performance in the textures analysis with low resolution, in which elements are difficult to be described [55].

The Local Binary Pattern (LBP) texture description considers local texture based on pattern and intensity. Therefore, LBP compares grey level intensities from pixels neighbours to central pixel intensity, being more precise and robust for light variability [56]. The LBP operator is a measure to extract texture characteristics of the images being a structural statistical texture analysis model. Another factor is the computational simplicity that makes it possible to analyse images in complex environments in real-time. Essentially, a value 1 (one) is attributed to the pixel is it corresponds to upper value considering a threshold based on a central pixel for a 3x3 region; otherwise, it receives 0 (zero). The basic LBP is derived by comparing the centre point with the near pixels considering the binary code. Lastly, a histogram is made of the calculated LBP codes. The histogram performs an image texture. For pixels f_p(p=0, 1,..., 7) the centre value f_c depends on thresholding, which Equation 2.9.

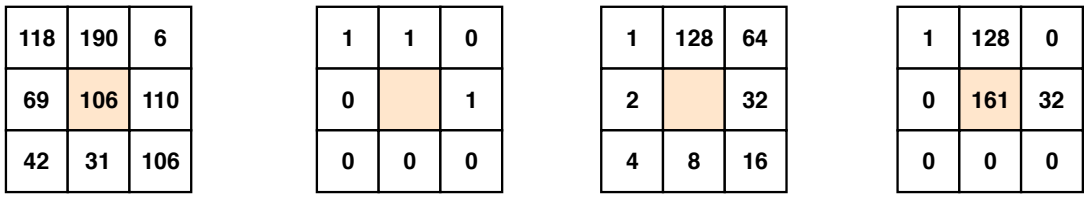
$$s(f_p - f_c) = \begin{cases} 1, & f_p > f_c \\ 0, & otherwise \end{cases} \tag{2.9}$$

Deriving the code from binary weighting applied to thresholding result the threshold points neighbouring equivalent the centre point and then unwrapping the code as a binary one [2]. Thus, the code LBP for f_c with eight neighbours f_p is exposed Equation 2.10.

$$LBP = \sum_{p=0}^{7} s(f_p - f_c).2^p \tag{2.10}$$

The spatial structure characterisation of the image texture is performed in the whole image by means of a convolution mask generating the LBP mapping. The process of extracting characteristics of the LBP operator is shows in Figure 9.

The constructing local binary pattern code exposed in Figure 9(a) refers the centre value point is exceeded three times, likewise, there are three 1 binary value in the resulting code can be observed in Figure 9(b). The unwrapped clockwise image region from the top left to point the resulting code is 10100001_2. Considering the binary code with weightings as Figure 9(c), the final value LBP=161 (Figure 9(c)). Basically, LBP



(a) 3x3 image region (b) Threshold result (c) Pixel weights (d) Code contribution

Figure 9 – LBP texture descriptor. Image adapted from [2]

code was complemented by two local measures which refer to contrast and variance. The former if these were computed from the difference between pixels encoded as 1 (one) and those as 0 (zero).

The Grey-Level Co-Occurrence Matrix (GLCM) is an usual method for texture analysis. This method is based on second-order statistics stored in the matrix and measures the spatial dependence of grey levels. The co-occurrence matrix is used to describe texture supported by the occurrence of some relationship between the image grey levels. In fact, if there is rapid variation with the distance is fine textures and slowly in thicker textures [54]. Several researchers suggested the used of grey level co-occurrence matrices, which have become one of the most well-known and widely used texture features. GLCM represents the frequency of pair of grey pixels in an image; in other words, it computes the occurrence probability of a pair of grey levels separated from a distance and angle pre-defined, which $P_{(d,\theta)}(i,\, j)$ since $(i,\, j)$ is the pair respectively. The distance d at the angle θ is the separating between the pixels. For an image $f(x,\, y)$ the GLCM is exposed in Equation 2.11 as a probability matrix, which $f(x,\, y)$ represents the tonality of a pixel at the position $(x,\, y)$, $P_{(d,\theta)}(i,\, j)$ is the value of how many times a pair of pixel has occurred and R the value of total pixel pairs.

$$\begin{gathered} f(x_1, y_1) = i \quad \text{and} \quad f(x_2, y_2) = j \\ p_{(d,\theta)}(i,j) = \frac{P_{(d,\theta)}(i,j)}{R} \end{gathered} \tag{2.11}$$

The matrix is generated with a dimension equal to the number of intensities in the image for each distance. Low values are used for fine textures and high values for thick textures. The image contrast is denoted by Equation 2.12, which the local contrast is low, the grey levels of pixels are similar. Contrast refers to the difference in visual properties that makes an object distinguishable from others and background.

$$S_{cont} = \sum_i \sum_j (i-j)^2 . p(i,j) \tag{2.12}$$

Correlations are related to the grey level linear dependence. The pixels at the specified positions relative to each other. The explanations are shown in Equation 2.13.

μ_x and μ_y correspond the mean considering the values of pixels in an image which is expressed in values of greyscale in the co-occurrence matrix as σ_x and σ_y that refer to standard deviation values.

$$
\begin{aligned}
S_{cor} &= \frac{\sum_i \sum_j (i.j).p(i,j) - \mu_x.\mu_y}{\sigma_x.\sigma_y} \\
\mu_x = \sum_i i \sum_j P_{(d,\theta)}(i,j) \quad &\text{and} \quad \mu_y = \sum_j j \sum_i P_{(d,\theta)}(i,j) \\
\sigma_x = \sum_i (i-\mu_x)^2 \sum_j P_{(d,\theta)}(i,j) \quad &\text{and} \quad \sigma_y = \sum_j (j-\mu_x)^2 \sum_i P_{(d,\theta)}(i,j)
\end{aligned}
\tag{2.13}
$$

The closeness of the distribution of elements in the GLCM with range [0 1] is measured by homogeneity, as previously mentioned (Equation 2.14).

$$
S_H = \sum_i^{n_k-1} \sum_j^{n_k-1} \frac{P_{ij}}{(1+|i-j|)} \tag{2.14}
$$

Another approach was disclosed in 1960 by J. W. Cooley and J. W. Tukey [57], the Fast Fourier Transform (FFT). FFT refers to the size change of the transform. The use of its discrete form DFT (Discrete Fourier Transform) and the FFT change the space domain (x, y) to the frequency domain (u, v). An image can be represented by a two-dimensional function. In this way, if $I(x, y)$ is an image, it can be transformed into a set of spatial frequencies $|F(u, v)|$. Thus, the Fourier spectral view is a two-dimensional sinusoidal surface [58].

2.3.5 Border

Edge detection can be performed based on constant values that differs from the background with an object. The detection methods are based on a search for discontinuity of grey levels, such as colour and texture between neighbouring pixels, being represented by the border in a region. In this way, edge detection involves image processing to find objects of interest [59]. The challenge of edge detection is related to imperfections in the image acquisition process. Noise can be one of the difficulties encountered for efficient detection. Defined as image degradation by random effect at the pixel to the pixel level, image smoothing provides edge blurring.

Most edge detection techniques seek to solve problems involving blurring or noise in the image. Detection consists basically of two steps. First step consists of evaluating the pixels when searching for edge properties and listing the current candidates. The evaluation is done through a mask that is compared to the pixels. Second step lists the preselected pixels and reduces information based on the first step. The Sobel, Prewitt, and Roberts operators are the classic methods for edge detection. These approaches use the

first derivative to identify significant changes. Through the gradient, the differentiation of images is analysed [59, 4].

The Sobel operator was developed by Irwin Sobel and Gary Deldman. It consists of an operator that calculates an approximation of the pixel intensity gradient in the image [60]. The convected 3x3 matrix is divided by horizontal and vertical variations (x and y directions). A matrix is used for horizontal variation of the image and another for the vertical calculation [37].

Figure 10 shows an example of two matrices that are convoluted. The magnitude is defined by G. It is emphasised that both matrices must be multiplied by the original image to find G. In the Sobel operator finds for horizontal (G_x) and vertical (G_y) edges. It is possible to observe the use of the value 2 in the central position of the mask, resulting in the smoothing of the image in this position.

-1	0	1
-2	0	2
-1	0	1

(a) G_x

1	2	1
0	0	0
-1	-2	-1

(b) G_y

Figure 10 – Sobel operator for edge detection. Image adapted from [4]

Another edge detection method based on first-order derivative techniques is Canny. This approach involves a list of criteria for optimal edge identification. The error rate (detection) consists of the noise signal ratio maximisation, the higher the edge detection, the lower the signal noise ratio. The second criterion is the location that represents the distance between founded pixels. The greater the nearest location the detected points will be in the true positions. The third step is the minimum response obtained by the detector consisting of only simple edges [61]. Briefly, Canny is recognised as the most complete edge detection method. The three proposed parameters alleviates noise effects, finds the graph and removes pixels that are not considered edges through the threshold.

2.4 Machine Learning and Pattern Recognition

ML is considered a subarea of AI that has emerged from computer science to provide methods for making machines able to learn. One of the most important characteristics of ML is automatic learning for complex pattern recognition toward making smart decisions based on data sets. Comprehension based on samples analysis is considered a challenge due to n possibilities of input samples [62].

Pattern recognition is based on observing data from objects for making pattern decisions. Usually, the object of interest changes according to the problem to be resolved

[37]. Classification or regression models are usually based on image characteristics which propitiate the identification of samples with respect to different classes or parameters. Thus, Machine Learning is the process which learning from instances inducting a model towards an automatic classification [26]. The modelling systems of machine learning compose several varieties of complexity, allowing adapt the appropriate conditions to each specific scenario [63].

The ML performs tasks associated with AI such as pattern recognition through prediction and construction of methods capable of improving performance [64]. Based on the inductive reasoning to perform the discovery of patterns, it generalises about a set of data with the intention of defining conclusions. The analysis is done with examples provided to the algorithm, being denominated as learning that is divided into *supervised* and *unsupervised*. Thus, supervised (predictive) category is typically used in the classification context. It is intended to map entries to output labels. For regression category the input is mapped to a continuous output. As mentioned, another category in pattern recognition task is called unsupervised (descriptive) learning which has an input vector, otherwise without the corresponding target values. The descriptive problems tasks may be to find groups with similar properties and examples composed in the data (clustering), besides determining data distribution within space, density estimation or high-dimensional space for visualisation purpose. The result of learning is the knowledge that can be used by the system to solve new problems. Using machine learning methods, there must be a distinction between the training sets (*model*) and the execution algorithm (*solutions*). In this way, ML provides the construction of models that allow the analysis of attributes contained in the image. The development of algorithms based on computer vision should be flexible and robust in order to improve performance in practical systems, being a higher level of competence and generality [65].

Pattern recognition is intended to process input variables by modifying the space of variables to facilitate the problem solution. The process called feature extraction also might be performed in order to speed up computation. Each problem task evaluated for dealing with complex pattern recognition may be computationally unfeasible. Finding useful features which preserve discriminating information is a challenge in computer vision problems. By this way, features extracted are used for input pattern recognition algorithms [66]. The classification task belongs to supervised learning and is used when dealing with instances with known classes. The main purpose of the classification is to create a model that learns and can classify objects that have not yet been seen. Thus, one searches through the space of possible hypotheses to find good representation.

The output value depends on the learning process applied in pattern recognition for classification. As mentioned, the supervised learning supposes the instances are labelled, which propitiates the generalised learning process capable of dealing with new

instances [67]. Considering it is essential for machine learning to assign properties descriptions of each class, a feature vector is composed of n-dimensional characteristics, which corresponds to numeric values that represent the object [68].

Classification is based on categorised process as supervised machine learning, which it knows the sample classes of the training set. Usually, supervised methods are carried in distinct steps considering training and test set to analysis the performance. Test step is fashioned in new samples to be classified, in this meaning, the classifier analysis each attribute class that it is unknown, and, based on training samples analysis, a designating class are inferred to the respective sample. Thus, the class are defined considering similarities from the classes of the training set [69].

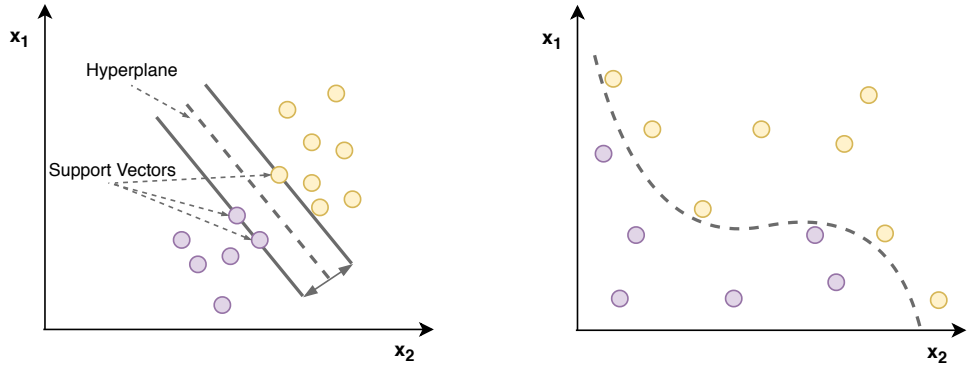
There are several classifiers available in the literature each of them has properties of making a decision. In this meaning, the proposed approach was tested with different classification algorithms to compare its performance. Dealing with the problem solution, some of the methods include non-linear machine learning algorithms, such as Support Vector Machine (SVM) [70] and Random Forest (RF) [71]. These algorithms are widely used in examples of supervised classification methods. In order to analyse the behaviour of image features groups (Colour, Intensity, Histogram, Texture and Border), the importance of each predicted component was evaluated using RF importance.

2.4.1 Support Vector Machine

SVM is a statistical learning algorithm proposed by Vapnik [70]. Belonging the kernel-based methods, SVM has a principle to find the best hyper-plane which separates the data into respective classes. Due to the good empirical performance and resources, SVM embodies the principle of Structural Risk Minimisation (SRM), which minimises a higher limit on the expected risk. By having as a characteristic the separability margin maximisation, SVM is considered with high power generalised technique, flexible and with very accurate results dealing with diverse type of problems [72, 73].

SVM stands out for having solid foundation and for achieving high performance in various applications. Considering the learning theory, it is identified precisely the factors that must be considered in order to have satisfactory results and model construction. According to Kwok [74], the SVM training involves optimising the convex quadratic function being a Mathematical Optimisation problem. Its properties involves free parameters that must be adjusted. Since there is no dependency between the size of the input space of the problem, the SVM can be efficient for problems with a large number of inputs. The SVM algorithm can be applied for pattern recognition to estimate indicator functions, estimate real value functions (regression), and feature extraction. Although fewer applicable, it has relevant fundamentals for more effective SVM. Figure 11 shows the ideal hyper-plane of separation between two classes by maximising the margin between the nearest points.

Figure 11(a) refers a linearly separable characteristics and 11(b) shows a non-linearity separable characteristics.



(a) Linear separable feature space. (b) Non-linear separable feature space.

Figure 11 – Support Vector Machine (SVM) performs classification hyperplane. Image adapted from [5]

The kernel-based method has four usually functions: Linear, Polynomial, Gaussian and Radial Basis Function (RBF). The technique is based on caring the probability distribution $P(x,y)$, which represents the vector samples of the training set x and respectively class y. In this meaning, SVM search for margin of separability among instances that situates in the edge regions considering the separability of classes. Summing up, the learning process consists in mapping $x \Rightarrow y$ to identify new samples to bring the same distribution P by training samples [70]. Thus, the amount of instances is named Support Vectors.

Regarding the maximisation of edge separability among classes, while it minimises the classification error, the SVM is related to problem square resolutions considering the hyper-plane. Concerning non-linear problems, it manipulates input spaces through kernel functions. Input samples are mapped in a new space with high dimensional, which it is possible to be linear separate [75]. According to Santos [6] example, this equates to the mapping of the input space X into a new space $Z = \{\phi(x)|x \in X\}$ of features in which ϕ_i represents the kernel function. Figure 12 demonstrates a mapping of a linearly inseparable input space transposed to a new space of larger dimension. Thus, when the characteristic space is modified, it is possible to carry out the linear separation.

Briefly, SVM is based on kernel dot product which makes it possible to stand out as:

- Efficient method with high power of generalisation, considering new samples to be analysed;
- Robustness for handling high dimensional data;
- Optimum hyper-plane for distinguishing classes with margin maximisation.

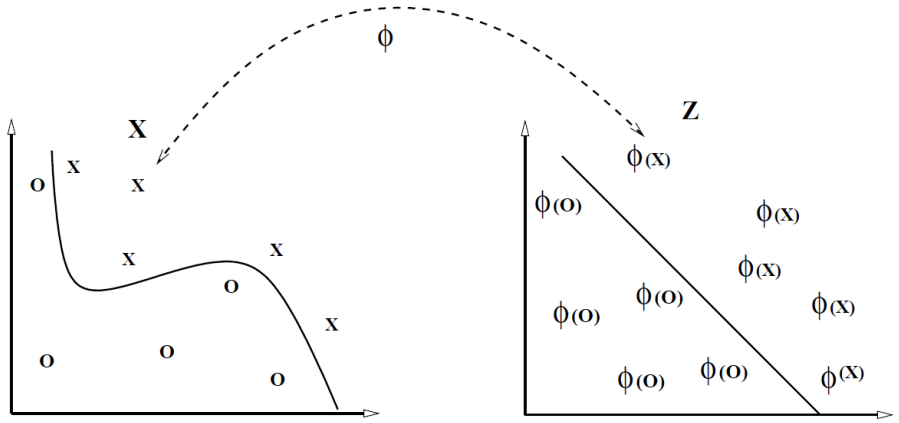


Figure 12 – Mapping the input space X to Z with the kernel function. Image from [6]

Summing up, SVM has an important advantage concerning the training phase, which is carried out by flexibility. Basing on margin maximisation and kernel dot product, the algorithm adjusts as a problem task. In this meaning, SVM is considered a method which controls the complexity irrespective of data dimension [76].

2.4.2 Random Forest

RF is considered a classifier algorithm based on decision trees, which is based on a set of these trees. This technique differs in that it has the objective of constructing several decision trees using a subset of randomly selected attributes. Through the bootstrap, it is possible to better analyse the data. Thus, the construction of the trees occurs by the selection of attributes randomly from the subsets. After creating the trees it is possible to carry out the classification of which has the best gain of knowledge for the solution of the problem. However, it is necessary to choose a subset of decision trees that has the best advantages in making decision. For each subset a vote is defined on which class the key attribute should belong. The vote has a weight in which the smaller the similarity between two trees the better. In addition, it is considered the strength that each tree owns individually, that more a tree needs its note will be better [77].

Decision trees are considered as inductive algorithms and strategy uses the decomposition of the larger problem into simpler sub-problems, recursively. The construction of decision trees refers to the choice of an attribute that will serve to divide the examples into subsets corresponding to each value of the attribute. For each subset the other examples that belong to the same class are observed. Associates the class with the node containing all the examples referring to the subset. Thus, a new attribute is chosen to partition the samples. The process is repeated until it obtains homogeneity of the sample classes from a final subset. Each node divides a smaller and less cluttered data set. The disorder can be calculated by Information Gain or Gini Coefficient. To avoid that a decision tree grows exponentially what makes it specific for a given database causing overfitting as pruning

techniques. By generating a more general hypothesis of the training set, pruning techniques are used to contain unique examples of a class, avoiding incorrect classification of examples [78].

The division rule is guided by measures that indicate how apt a given attribute is to discriminate classes. It is used to select the attribute that maximises the measure. Considering a node t where the probability of observing the class c_i is p_i, the probability of each class is given by $p_1, p_2, ..., p_k$, such that $\sum p_i = 1$. To find the information gain of a characteristic c_i, the entropy of each of the resultant data sets of a root node is calculated. The calculation of the entropy of a subset $A \subseteq T$ which T is a set of data represented in Equation 2.15.

$$H(A, c) = -\sum_i p_i.log_2 p_i \tag{2.15}$$

The weighted average of the entropy of each of the subsets resulting from the choice of the characteristic c, exposed in Equation 2.16. The operation μ refers to the resulting value of the data set sizes of the data subsets when it is applied to T and A_i, respectively.

$$H(T, c) = \frac{1}{\mu(T)}.\sum_i H(A_i, c).\mu(A_i) \tag{2.16}$$

The entropy $H(T)$ results in relation to the data sets and $H(T, c)$ is the weighted average of the entropy of the subsets generated from the characteristic c_i. Thus, the information gain of the samples is calculated and the one with the highest information gain is selected. Equation 2.17 shows the information gain formula.

$$IG(T, c) = H(T) - H(T, c) \tag{2.17}$$

Breiman et al. [79] proposed Gini function to measure the impurity degree of the nodes. Equation 2.18 defines the relation of characteristics for selection of attributes. An attribute in which the weighted average impurities of implicit descendent nodes is subtracted from $i(t)$ is examined. The result of the impurity decrease will be selected [80].

$$i(t) = 1 - \sum_i p_i^2 \tag{2.18}$$

Considering a training set T with a attributes and n examples. The bootstrap sample T_k of the set with m random attributes ($m \leq a$) on each node of the trees. A tree is randomly induced from a set of possible trees. Using m random attributes on each node. The combination of large sets generates accurate models. The ensemble learning

method, Random Forest (RF) was proposed by Breiman [77]. Combining many decision tree models, the main idea is into a forest that provides more accurate prediction. The technique applies diverse training set in each tree in the ensemble. Thus, the Random Forest composes by the Bagging approach (bootstrap sampling). The proposed method adds an additional layer of randomness to bagging and changes how the classification trees are constructed using different bootstraps sample of the data. In fact, the nodes are split observing the best split among all variables in the standard trees. In the Figure 13 can observe the Random Forest representation.

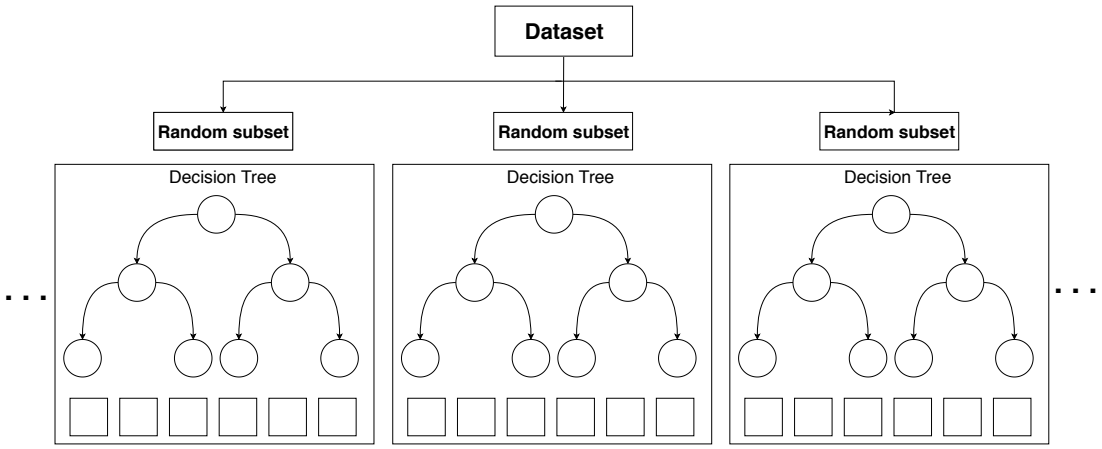


Figure 13 – *Random Forest* structure. Image adapted from [7]

For each sampling of the data set a decision tree is generated. Each decision tree contributes to the classification process of an unknown instance. The steps for RF pattern recognition are determined by the procedures forming *ntree* bootstrap samples are formed by analysing the original input dataset. For each bootstrap sample, a tree is generated in which each node chooses the best division among the predictors randomly m_{try} and predicts new instances aggregating the predictions of the *ntree* [7].

Briefly, RF is based on the divide-to-conquer feature which makes it possible to stand out as:

- Better performance compared to decision trees;
- Better accuracy when tested on different datasets;
- Avoid overfitting;
- Less sensitive to noise.

Breiman [77] justifies the use of the bagging method by improving performance are used random attributes. One can provide continuous estimates of the generalisation error of the combined trees set. Therefore, the Random Forest is considered an intuitive and robust performance strategy. The Random Forest features importance is a useful metric

which estimates the importance of a variable and inspects the impact of each characteristic of sample extract of prediction task.

2.4.2.1 An importance measure for variables

Measuring feature importance is often an interesting model for selection or interpretation. Strategies can be proposed observing the most important variables to predict patterns. To observe the provided forest of trees is a challenge, which the interpretations of variables interaction is considered critical and impracticable. To deal with, RF also provides measure of importance based on permutation test [77].

Briefly, RF is composed of several single trees built using bootstrap samples which becomes from splitting the data into random sets. To the provided prediction is based on averaged valued or majority votes of prediction from each tree [81]. To do so, the called "out of bag' (OOB) samples are the ones not used to fit the current tree, which can be applied to estimate an impartial prediction error (OOB-error) and to evaluate variable importance of X^j.

The score of importance is given by observing the misclassification rate (classification task) in the forest when the observed values of the variable are randomly permuted in the OOB samples (OOB_t) of a single tree [82]. The permutation importance measure is computed by the difference of OOB-error, denote by ${OOB^j}_{error}$, from each tree t concerning before and after the random permutation phase of a variable for prediction. In this meaning, the mean value of all the trees is the current importance of a variable [81]. Equation 2.19 defines the variable importance of X^j.

$$RF_{imp}(X^j) = \frac{1}{ntree}\sum_t (OOB^j{}_t error - OOB^t error) \tag{2.19}$$

3 VEINING DEFECT ISSUES

Dry-cured ham production is considered a time-consuming process which is affected by many factors. Raw material and process conditions mainly interfere in compounds formed by reactions postmortem, consequently, final product [15]. The production of dry-cured ham must achieve some quality control parameters, which are based on requirements of the food processing industry. In this sense, the appropriate development of its production steps is essential to ensure quality requirements. Traditional technology for dry-cured ham production is based on dehydration and salting of the raw pieces of whole muscle product. The final processing involves some attributes which depend on the proper development and quality of raw-material. In this matter, raw samples are evaluated to identify some defective characteristics which could impact the quality and acceptability of the final product [83].

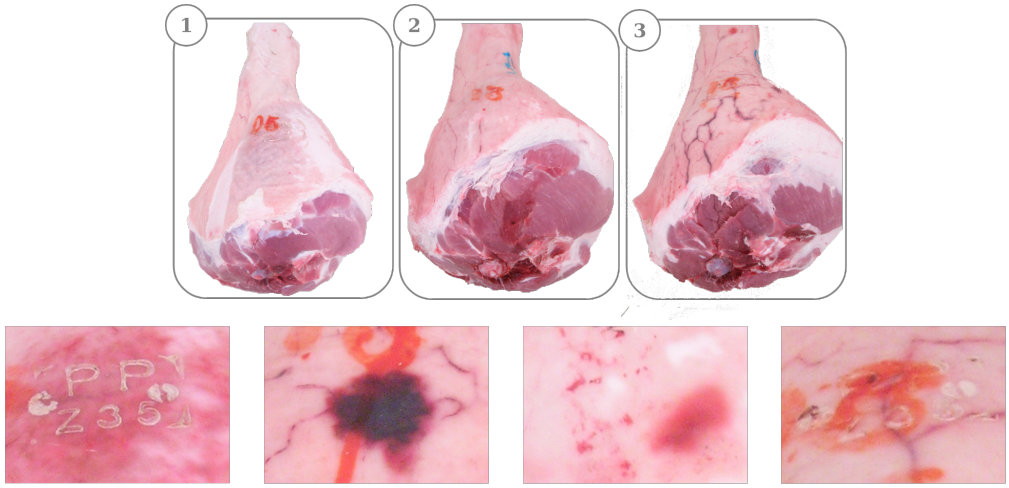
High-quality in dry-cured ham concerns appropriate visual aspects, which is primarily based on characteristics observed before curing. Skin issues and subcutaneous layers defects, e.g. the presence of visible blood vessels (veining defects), are related to low-quality products, which consequently result in ham discards. Due to the quality standards requirements, carcass lesions have a substantial economic impact in the production systems [84]. For this reason, the factors that impact the quality of dry-cured ham has observed, and it requires precise techniques for monitoring process production.

Traditional ham evaluation is based on Parma Quality Institute (PQI) [85] standards. Following the PQI statements, in the case of inappropriate physical characteristics, the sample is excluded from the dry-curing process line. In this meaning, many factors mainly determine raw ham properties. Indeed, quality parameters such as weight, the thickness of the adipose and presence of visual defects are influenced by antemortem factors [23]. Thus, according to Russo et al. [83], many raw ham defects are related to stunning methods and time-delay, as veining defects, which are exceeded in the process to refrigerate the carcass (pre-chilling time), which may occur during the dressing process.

Researches have shown an increasing presence of defects in raw hams, among which the veining defect has a higher frequency, up to reach worrying levels [86]. Moreover, trimmed hams are classified, and eventually rejected, based on visual comparison with PQI photographic standards.

The evaluation of veining defects in dry-cured ham is generally based on the quantity of veining, which is expressed in terms of increasing defectiveness levels. Considering the method exposed by Russo et al. [83], three different levels of veining defect were considered. The labelling process was carried out comparing the ham samples with a ref-

erence photographic scale in which level 1 (C_1) corresponds to hams with no defect or not-perceptible defects; level 2 (C_2) considers carcass with slight veining presence and level 3 (C_3) visible defects. Figure 14 shows a representative image for each veining level.



(a) (b) (c) (d)

Figure 14 – Images of raw ham veining defects corresponding to level 1 (C_1), level 2 (C_2) and level 3 (C_3), respectively, and examples of skin lesions: (a) fire marks with traumatic hematoma, (b) non-traumatic hematoma, (c) scratches and (d) meat stamp.

Besides veining defects, the presence of additional skin lesions and marks (figures 14(a), 14(b), 14(c) and 14(d)) pose an additional challenge for the correct identification of blood vessels, and consequently for the automatic classification. Indeed, spatial features derived from the images of ham samples may also incorporate information related to skin defects different from veining. In this context, it is necessary to develop effective strategies able to improve model performances and reduce false positives.

3.1 Related Work

The food monitoring process demands a rigorous quality control due to the consumer requirements for its products. The quality evaluation is commonly carried out by observing characteristics such as appearance, smell, texture and flavour which are often examined by human specialists. Among several techniques to support the analysis of food products, visual inspection is the usual one, employed to assess aspect-related characteristics [87]. This technique allows a straightforward evaluation based on panels for easy assessment by human inspectors. Furthermore, several defects are related to consumer perception which is quite related to the visual aspect of the food product [88]. However, the bias of human visual perception, limitations when evaluating multiple samples and environmental interference may lead to an inaccurate subjective evaluation with high variability.

Non-invasive technologies have gained popularity, which are considered useful for being applied to quality control towards the production process. To solve current drawbacks in food production, strategies and technologies have been the goal of several studies. Automatic vision inspection takes part in several processes; in this way, vision techniques become essential. Through the inclusion of CVS, the quality of components and products is verified with high accurate processes. In this meaning, the production is evaluated as well as assisting for parameters analysis for whole quality control. Moreover, efficient methods of quality control possibilities to assist in planning to a final production are required. For instance, due to the consumer requirements for food products, the monitoring process demanded to be enhanced quality control. Therefore, the food industry requires objective measurements system to determine and keep quality control in the final product.

Concerning industrial environment requirements, techniques based on image analysis has proposed for improving results in quality control. Despite the reliable process conditions, those approaches provide accurate solutions which are possible to be applied in industrial production [89]. Several CVS were proposed to automate the inspection of food products based on their aspect dealing with the limitations of visual evaluation by human operators [11, 90, 91, 10, 92, 29]. Particularly, some CVS approaches were developed for classifying sample qualities in industrial settings, including the detection of defects [30].

Exploring the potential of CVS with RF [77] and SVM in comparison to other classifiers, Xu, Riccioli and Sun [93] applied those algorithms for distinguishing between organic and conventional farm-raised salmon fillets in fresh and chill-stored conditions. The achieved result has been powerful used SVM predictor for multivariate analysis obtaining an objective and rapid categorisation of the two salmon varieties. At the same line, Barbon et al. [94] observed some attributes as water holding capacity (WHC), pH and colour descriptors to explore the optimal parameters for classification step for poultry meat samples. The methodology proposed for the authors was realised by SVM, which allowed for classification of poultry samples on quality grades 98.1% of precision.

The identification of items food process propitiated several studies related subjects as shown on this approach grouping CVS methods to improve image-based prediction. Some CVS demand sophisticated structure modelling with a complex problems solution. Thus, ML techniques were widely investigated for evaluating attributes and quality food. Introducing SVM [70] for sensory analysis, Muñoz et al. [95] proposed an automated tool for marbling classification of dry-cured ham based on segmenting intramuscular fat. SVM has shown high performance obtaining 89.00% of accuracy.

Barbon et al. [16] proposed to classify pork storage time applying ML methods in order to investigate the most appropriate method based on traditional parameters. In comparison to the other algorithms used in this approach, the RF showed the best

accuracy and precision. Ropodi, Panagou and Nychas [96] classified beef adulteration using a CVS. The classification models proved less successful compared to the case of adulteration with pork.

In this meaning, many strategies are proposed trying to solve current drawbacks in dry-cured ham production to monitor and characterise of ham during the production process. Specifically for dry-cured ham, for example, Prevolnik et al. [97] reported an attempt to classify samples according to the maturation time through of near-infrared (NIR) spectra. The overall performance was 79.7% of accuracy demonstrates being relevant for processing control. Thus, exists a wide range of applications which proves the effectiveness of CVS since classifying sample qualities as detecting meat defects in industrial settings [30]. Studies claim the animal farmers and meat industrial sector should minimise defects found through controlling the environmental factors which influence the pork meat quality [98, 99]. The evaluation of pig thighs proposed in Lo Fiego et al. [100] acquired digital images of the external surface region to find objective parameters. The samples were analysed as weight, length, circumference, thickness of fat and thigh, globosity index and colour of skin. In order to auxiliary the evaluating process, the presence of veining and red skin defects considered in the model. The obtained results were not allowed a classification of analysed thighs needing a robust technique to solve the problem. Nonetheless, the authors observed the higher incidence of globosity index value is directly related to veining and red skin defects.

Concerning the scope introduced in this work, the proposed approach has been focused on distinguishing different levels of veining defects in raw ham. This study aims at introducing a CVS approach for classifying each level achieved in DSIA process. The core of the proposed image analysis strategy is based on machine learning techniques applied via dual-stage process. This technique proposes to split the overview perspective of original samples in sub-images that composes a dual scenario to improve the prediction performance. In other words, it is based on images segmentation to enhance the information features characterising the problem by dividing the complex problem to facilitate the final classification. Thereby, the technique was proposed, as shown in the previously mentioned, to improve prediction performance in those problems for which traditional machine learning methods applied to CVS are not sufficient to correctly discriminate the classes of interest.

4 DUAL STAGE IMAGE ANALYSIS

In order to improve the classification process based on CVS, in this work was proposed a strategy called Dual Stage Image Analysis method (DSIA) to obtain a better comprehension of each sample in comparison to Traditional CVS. A vast number of characteristics might rely on performance improvement of prediction tasks. The traditional feature extraction method considers the whole image at once for extracting its features, which possible decreases important information of some image descriptors from samples. For this reason, DSIA was proposed to improve performance in problems task, which suggests splitting each image into several minor regions toward extracting local image features from sub-regions (Figure 15). Thus, the current image is viewed as its low-level visual features extracted from all sub-regions. The basic process of DSIA is turning divide the problem into sub-problems in other to simplify the prediction.

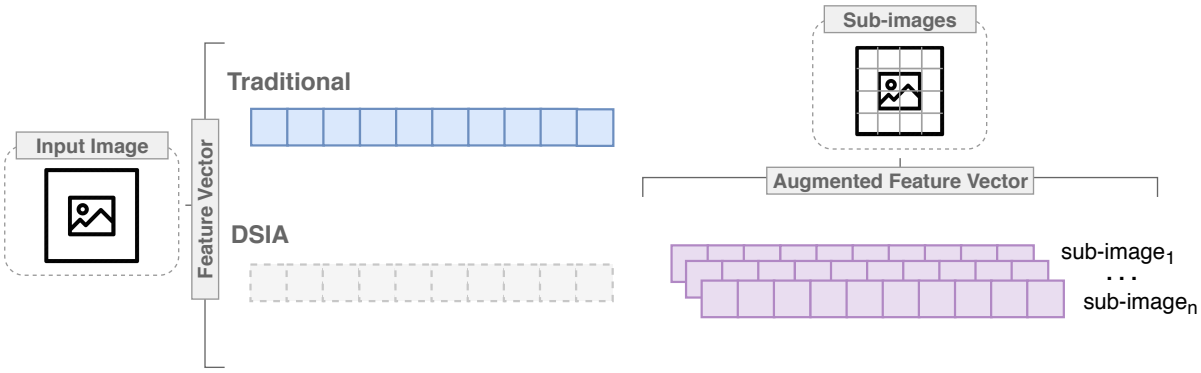


Figure 15 – General overview highlighting the differences among Traditional and DSIA approaches of image analysis (feature vector composition).

Dealing with complex samples which the understanding of the problem is difficult, it requires a robust technique to solve the classify prediction. In some problems, the feature extraction considers the region of interest in a totally, which does not provide sufficient information to distinguish the different classes, decreasing the power of ML prediction. The DSIA technique is an alternative to improve the performance of problem classification prediction.

The proposal of segmenting samples in small regions was suggested. When composing the dataset with several sub-samples, greater detail is provided for each sub-region, reducing the error caused when the image is used as a whole. The DSIA is a strategy to obtain the image classification, as observed in Figure 16. This approach creates sub-regions which extract characteristics from them toward composing a feature vector with all regions. After it is defined each minor region, a new class to be predicted through an automatic labelling scheme; in other words, it was employed the subdivision of each sam-

ple to induct a new classification, so as simplify the prediction of majority problem class, but it considers each image separately. In this way, from a new sample image, each sub-region obtained is classified, and the final decision is achieved with the sets of prediction sub-samples.

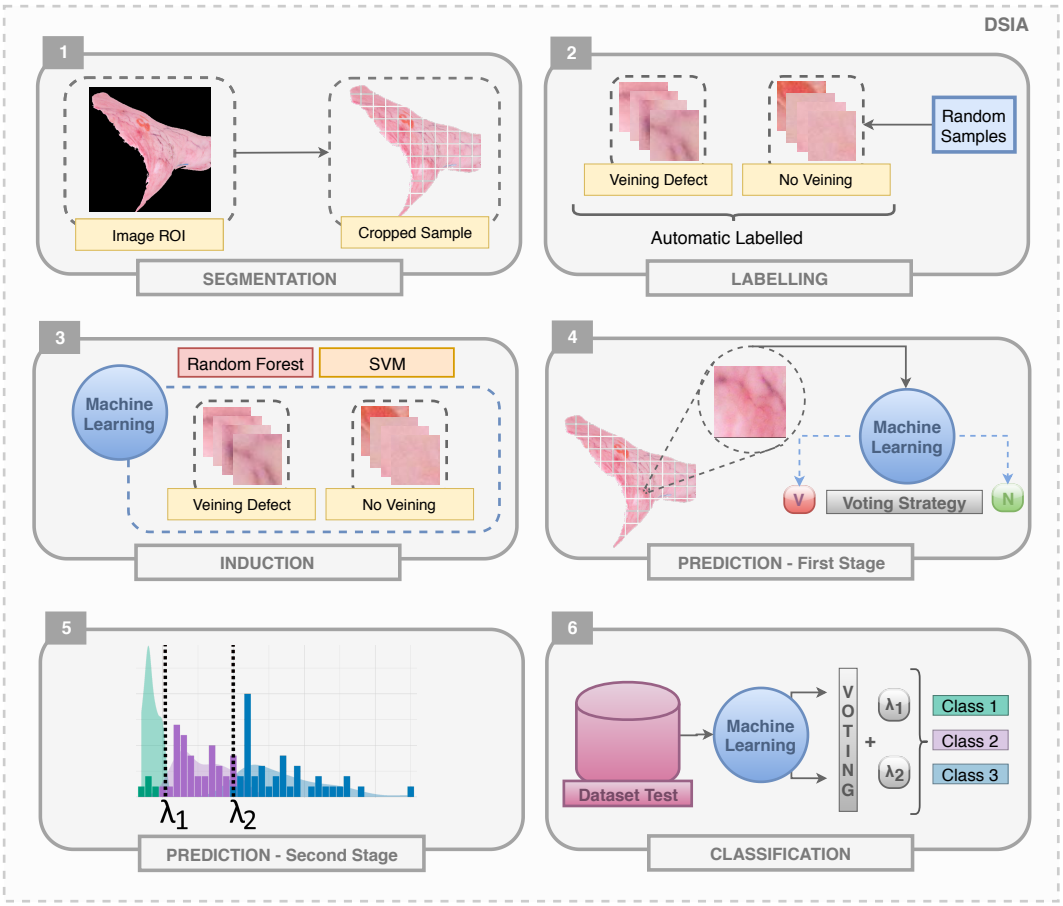


Figure 16 – DSIA proposed method.

Summarising the respective steps showed in Figure 16, the proposed approach is designated to classify samples at different levels. Basing on distinguishing each level, the analysis is measured through the relation of portion characteristics presents in ROI. Therefore, dealing with the distinction between each samples class, as previously mentioned, it is considered a challenge due to the complex samples involved with morphological characteristics to be distinct. Thus, the employed technique can be detailed in six main steps: Segmentation, Labelling, Induction, Prediction (First Stage), Prediction (Second Stage) and the final classification. It is important to highlight the proposed DSIA is a technique to improve the classification performance grounded on more informative strategy from image sample analysis which suggests the division of problem comprehension in separated stages to obtain improvements in levels distinction.

Analysing the context of the problem, the characteristic which provides information for the solution among classes is based on the presence of specific properties in each ROI part. Thereat, this work was proposed to divide the sample perception into two dif-

ferent stages: the First Stage has utilised the segmented image to predict regions; the Second Stage has observed the prediction previously mentioned toward formulating the regions according to the amount of characteristics found per sample. As part of the first step (exposed in Figure 16), in this work was proposed a strategy based on segmentation which splits each image into a smaller sub-images and extracting local image features from each sample. So, in order to find the best region size that provides informative morphological characteristics, it was investigated different sub-image resolutions. Following the sequence, the second step is essential to reach the goal, which was categorised the smaller samples into sub-problems to be predicted. As a way to maintain the reliability of the problem solution, the smaller images were sorted an automatic labelled process. This part of the dataset preparation refers to the First Stage based on original samples which each sub-region compared the respective region in the ROI.

Concerning the raw ham samples, due to the presence of textures that could confuse the identification of specific properties regions in the samples. Like decision trees, "divide and conquer" is a strategy to improve the performance for problems tasks with complex characteristics. Briefly, in order to separate the sub-set in the classes as previously mentioned, the automatic labelling scheme was used to define which sub-sample belongs to each one. Thus, the separating process checks each sub-image to detect the regions belong to the original sample. Using them as a basis, the dataset was divided in order to separate them into two sub-regions that consist of image features for the training dataset and test dataset. At the end of the iterations, the final result was computed from the sub-images prediction to measure the relation of specific properties which is composed. This procedure concerns a more reliable source of image classification by evaluating smaller portions of the image and concentrating the characteristics in each part obtaining a robust description of the sample.

The next step refers to the third (exposed in Figure 16), the extracted features of each sub-sample, and the ML is inducted. As expected, the output of prediction dataset was modelled on RF and SVM algorithms. This part of the process refers to the First Stage (step four). In order to improve the prediction performance of levels, the voting strategy is proposed. As the prediction among the sub-images is obtained, the levels are related through the amount of properties present in the region of interest. Thus, it is necessary to measure the quantity of predicted sub-samples per each subclass, obtaining two threshold values, λ_1 and λ_2. The thresholds are computed from the intervals of Probability Density Function (PDF) considering each level. Applying to distinguish sample levels, the DSIA output is based on the relation between the number of the minority and the majority sub-regions classified toward for prediction. The relation of the sub-regions of each sample is delineated the threshold is sought between the samples for the final decision (step five - Second Stage).

DSIA was performed to improve the performance obtained of the Traditional CVS method. So, subdividing the classification problem, it possibilities to extract punctually features to describe the problem and enhance the prediction. The scenario proposed in this approach requires to distinguish the levels between different classes to identify the quantity of specific properties contained in samples. Therefore, to destiny the class of each sample, after the prediction of sub-regions and having the relation of sub-classes prediction, the threshold was found between that association of sub-images to determine the final classification.

Summing up, like traditional CVS, also DSIA is a supervised modelling technique. For this reason, a calibration phase is necessary before using it in prediction. Concerning raw ham tasks, this phase requires a supervisor able to label some samples and identify veining related pixels. So, there are two main stages in DSIA to be applied: the First Stage addresses the segmented sub-regions and their predictions; the Second Stage uses a threshold over the counting of veining sub-regions to provide a final classification.

- *First Stage*: This stage focuses on sub-regions prediction. In calibration, DSIA requires an image with veining marks to automatically label the sub-regions. In other words, a segmented sub-region with veining presence is labelled as *veining*. On the other hand, if the sub-region has no veining mark, it is labelled as *non-vening*. The labelled sub-regions have their features extracted, and a machine learning model is induced to predict them as *veining* or *non-vening*. In the prediction phase, the induced model is used throughout all sub-regions from a single ROI to provide the number of *veining* regions for further classification.

- *Second Stage*: The class identification provided by the supervisor is used in the procedure called Threshold Computation. This procedure is applied to identify two threshold values, λ_1 and λ_2 through PDF considering each defect level of raw hams (C_1, C_2 and C_3). Under the assumption that the three classes are equiprobable and that the distributions are accurately estimated, λ_1 is the threshold over the area of overlap between C_1 and C_2, and λ_2 refers the separability value between C_2 and C_3. These thresholds are used as classification boundaries related to the number of *veining* sub-regions during the prediction phase.

Concerning ROI splitting in the First Stage, in the experiments different sizes of the sub-regions were evaluated to find the optimal splitting scheme: 10x10, 25x25, 50x50, 75x75 and 100x100 pixels. The image features extracted from each sub-region are reported in Section 5.3.

5 MATERIALS AND METHODS

This chapter presents the dataset applied in the experiments, the metrics employed to evaluate the different aspects of the algorithms applied to compare with DSIA and Traditional CVS, and details about the experimental setup. Classification models are designed based on the necessity of the problem solution. Such results should be informative. Thus, features are extracted in order to identify samples from different prediction levels. So, ML algorithms can induce models from image features towards classification of raw ham considering the *veining* defects presents on them.

In this way, the approach is proposed to solve the mentioned problem refers to the partitioning of the classification. The complexity of the problem to determine the level of the raw ham according to the presence of significant veins demands to create strategies to improve the accuracy results.

5.1 Data Acquisition

A total of 194 raw ham samples were collected; those of them are derived from a different heavy pig. Samples were labelled according to three different veining levels for creating the image dataset: Level 1 (C_1), which involves the clearest region considering the *veining* defects presence (17 samples), Level 2 (C_2), which is related to samples with a moderate veins (92 samples) and Level 3 (C_3), which represents the worst case for appearance of *veining* defect on the raw ham (85 samples). Each image was labelled by three expert assessors using the reference method based on the comparison with a photographic standard. The methodology applied consists in a subjective analysis based on the evaluation of the raw ham veining defect intensity of the digital image sample.

The samples were analysed in fifteen large pig farms regularly registered for slaughter in the Italian region. Initially, the results were carried out in the slaughterhouse and the respective ham factories of raw ham. The examples were collected from different farms in order to heterogeneity the data. Raw ham samples were collected in a slaughter plant during five guarantee the different days. Afterwards, each carcass was sliced and subjected to a low temperature as $0-4°C$ for 24 hours. Lastly, after thigh storage, each carcass was trimmed.

Individual raw ham sample was acquired using a Nikon Coolpix 5400 camera with a 5.8–24 mm focal length in a controlled lighting environment. Images were obtained by observing the external surface of each pork thigh. Equipment with the automatic lighting system was used in order to perform as photos of each sample and to improve the quality of the image. The lighting consisted of LED lamps (Natural daylight, 100 W) with an

angle of 45° between the camera lens and lighting source axis. Various tests and studies have been carried out to improve lighting conditions to minimise the influence of the different local of exposition. The exposure time and the camera aperture were analysed for better capture of the image details to be used in subsequent phases. To obtain a better perception of each image, the camera was located vertically over the sample. The images were acquired using the following camera settings: manual exposure with a shutter speed of 1/125 s (zoom and flash functions off) and ISO equal to 200.

As part of the process of identifying the ROI in the raw ham image, the Colourgram technique was used. The process is divided into 6 steps apart according to Figure 17. Image pre-processing was conducted considering 6 subsequent steps, with the aim of identifying and segmenting the region of interest in each image. Indeed, the isolation of the pigskin form the background and other components of the thigh was not an easy task, due to colour similarities with some parts of ham fat and muscle.

Step 1 of the pre-processing procedure was performed by means of the colourgrams approach [8] in order to obtain a general overview of the colour properties of the samples images. Basically, this approach consists in converting each RGB image of the dataset into a 4900-points long signal, namely the colourgram [47, 29, 48, 49]. Each colourgram is obtained by merging in sequence the frequency distribution curves of the R, G and B channels, and of additional colour parameters derived from RGB. These parameters include: the lightness, calculated as the sum of the RGB channels; the three relative colours, i.e. relative red, relative green and relative blue, calculated as the ratio between each channel and lightness; hue (H), saturation (S) and value (V), obtained by converting the RGB coordinated into the HSV colour space; the scores obtained by applying PCA to the raw, mean centred and autoscaled RGB matrix. Thus, a dataset of RGB images is converted into a matrix of signals, which in turn can be further analysed in order to obtain a general overview of the colour properties of the samples at the image-level. The list for description the single peaks is shown in Table 1.

Furthermore, since the colourgrams are obtained by merging in sequence the frequency distribution curves of several colour parameters, it is possible to identify colour features of interest and visualise them back at the pixel level in the original image domain. A similar approach has also been successfully proposed for the analysis of hyperspectral images [101, 102]. In the specific case under investigation, the acquired RGB images of the raw ham samples were converted into colourgrams and the signals were visually inspected in order to identify the colour parameter leading to the optimal segmentation of the ROI. Thanks to the colourgrams approach, it was therefore possible to simultaneously analyse the frequency distribution curves of several colour parameters calculated for all the images of the dataset. Among all the considered colour parameters, relative red allowed to obtain the best results to remove the pixels related to the background and ham fat.

Table 1 – List of the peaks in a colourgram and respective position. Table from [8]

No.	Colourgram Region	Description
1	1-256	Distribution curve of the Red values
2	257-512	Distribution curve of the Green values
3	513–768	Distribution curve of the Blue values
4	769–1024	Distribution curve of the Lightness values:L=R+G+B
5	1025–1280	Distribution curve of the Relative Red values:RR=R/L
6	1281–1536	Distribution curve of the Relative Green values:RG=G/L
7	1537–1792	Distribution curve of the Relative Blue values:RB=B/L
8	1793–2048	Distribution curve of the Hue values
9	2049–2304	Distribution curve of the Saturation values
10	2305–2560	Distribution curve of the Intensity values
11	2561–2816	Distribution curve of the first score vector from the PCA on the raw unfolded RGB matrix
12	2817–3072	Distribution curve of the second score vector from the PCA on the raw unfolded RGB matrix
13	3073–3328	Distribution curve of the third score vector from the PCA on the raw unfolded RGB matrix
14	3329–3584	Distribution curve of the first score vector from the PCA on the mean centred unfolded RGB matrix
15	3585–3840	Distribution curve of the second score vector from the PCA on the mean centred unfolded RGB matrix
16	3841–4096	Distribution curve of the third score vector from the PCA on the mean centred unfolded RGB matrix
17	4097–4352	Distribution curve of the first score vector from the PCA on the autoscaled unfolded RGB matrix
18	4353–4608	Distribution curve of the second score vector from the PCA on the autoscaled unfolded RGB matrix
19	4609–4864	Distribution curve of the third score vector from the PCA on the autoscaled unfolded RGB matrix
20	4865–4900	Normalised loading vectors and eigenvalues of the three PCA models: raw, mean centred, autoscaled

More in detail, a threshold value equal to 0.36 of relative red was identified through the inspection of the colourgrams: pixel with relative red values lower than the threshold are ascribable to background and ham fat, while pixels with relative red values higher than the threshold are ascribable to muscle and pigskin.

In Step 2, the segmented image (with removed background) was converted into a binary image by setting to one all the pixels with a value higher than 0.8 and setting to zero all the remainder pixels. In this way, the areas of the image ascribable to the darker parts of the muscle are eliminated (Step 3). Afterwards, the edges of the objects depicted in the binary image were identified, as shown in Figure 17 (Step 4). In order to create a mask to identify only the part related to the pigskin of the samples, the boundary of the corresponding object (Step 5) was isolated. As the final step, the mask was applied to the original image, obtaining the region of interest ROI (Step 6).

After segmenting the image in the ROI using the Colourgram algorithm, the samples were spotted in sub-regions. They were delimited in 10x10, 25x25, 50x50, 75x75 and 100x100 pixels and selected which does not contain black background part in the composition. The sub-samples were separated according to the resolution and composed five distinguish datasets. To generate the respective data of *veining* and *no veining* defects present on a small part of the sample, it has used the automatic voting technique, as before specified. The voting was set up, so each sub-region was founded with parts of *veining* defect highlighted. Only those with more than 50% relative to the number of pixels in the image column or line would be considered as part of the *veining* dataset and the others for the *no veining* dataset.

The problem was formulated to apply the ML by using an ensemble strategy to obtain the image classification. The overview of the process was composed by feature vector contains characteristics from each ROI by sub-region. Thereby, the prediction of the

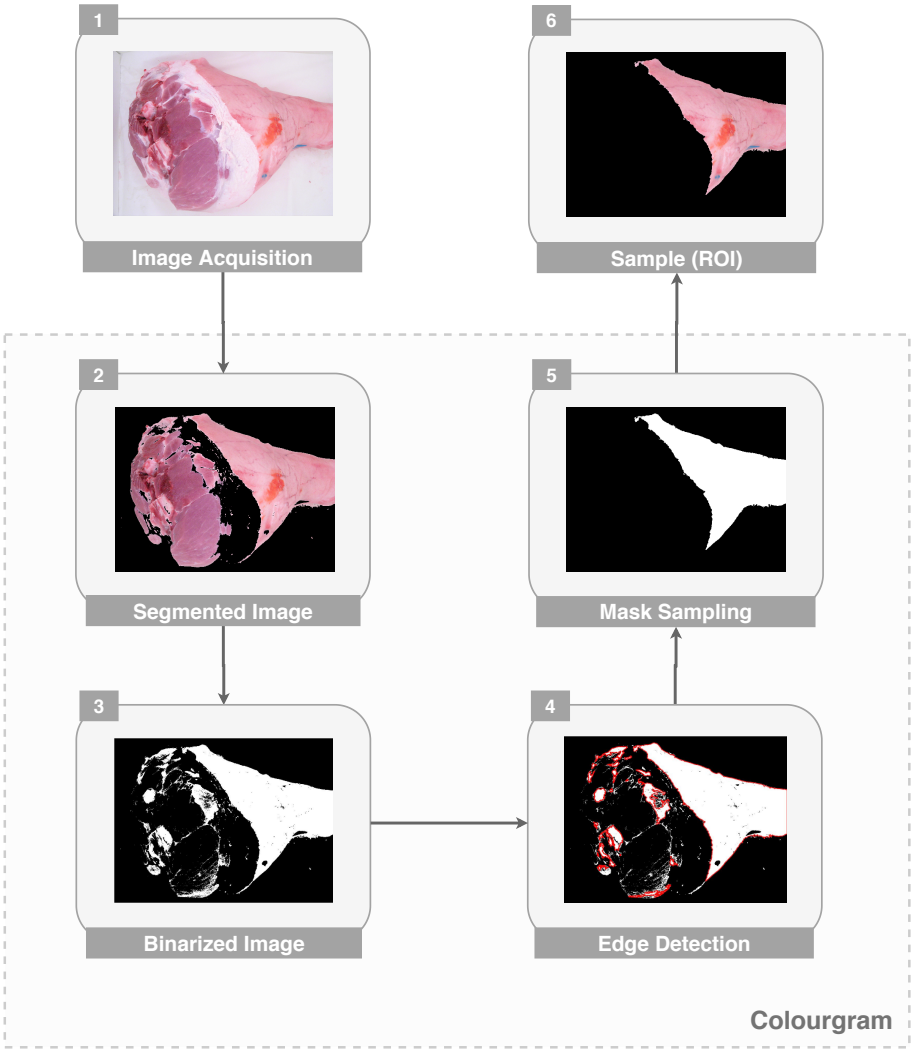


Figure 17 – Image pre-processing of identification at raw ham (ROI).

original sample is based on a threshold process about the sub-image samples classification, as previously detailed in Chapter 4.

The next step was to separate as a sample from each sub-region through a voting scheme. In other words, it was employed a segmenting samples process to induce a classifier in two distinct steps. The base classification of sub-problem was formulated observing the original sample. The voting process was constructed through an analysis of the sub-sample in the correspondent region belongs to each image. This part of the proposed approach had been defined as an automatic classification to compose the dataset to ensure efficient and trustworthy results. It is important to mention that the samples were randomly split into a training set, including 110 samples (57% of total samples), for the calibration phase and a test set containing 84 samples (43% of total samples) to evaluate the results in prediction. Both RF and SVM were compared in sub-region prediction modelling, as well as in the traditional CVS framework.

5.2 Evaluation Metrics

Supervised ML has several ways of evaluating performances. The Confusion Matrix is the essential way of measures of the quality based on correct and incorrect recognition examples for each class, which was applied to evaluate the performance of each technique applied in this work (Table 2). Its goal is to measure the similarity among classification results of CVS, considering True Positive (TP), which represents correct classifications in positive cases; True Negatives (TN), which represents correct classifications of negative cases; False Positive (FP), that represents incorrect classifications of negative cases; and False Negative (FN), which represents incorrect classifications of positive cases [103].

Table 2 – Confusion Matrix for binary classification.

	Truth		
	Class	Positive	Negative
Predicted	Positive	TP	FN
	Negative	FP	TN

Accuracy is an important verification measure, that considers the proportion of correct recognition of prediction analysis. So, it is the most used empirical measure. Considering the values of the Confusion Matrix, the Total Accuracy is formulated by the sum from the main diagonal compose to results of ML prediction. These diagonal values are the TP and TN which, observing the Equation 5.1, are divided by the sum of the whole values constituting the (n) matrix. Therefore, the Total Accuracy propitiates analysing the performance of the method used to predict the image samples. Consequently, it is possible to relate the achievement results to the relevance of features extracted and ML algorithms in classification.

The performance from analysis of built models was made by use of Total Accuracy method (Accuracy Matrix) [104] which is defined by Equation 5.1. The metric is computed through the Confusion Matrix based on the samples evaluation.

$$Total\ Accuracy = \frac{TP + TN}{n} \tag{5.1}$$

Additionally, Precision (Eq. 5.2) and Recall (Eq. 5.3) were used to provide a more realistic comparison since the dataset is unbalanced. Focus on one class, Precision and Recall are applied in case the number of examples belonging to one class is often substantially lower than the overall number of examples. Those metrics are based on FP and FN to distinguish the correct classification of labels within different classes being concentrate on one class. Precision is a function of its correctly classified examples and

its misclassified examples, whereas Recall considers of TP and examples misclassified as positive ones.

$$Precision = \frac{TP}{TP + FP} \tag{5.2}$$

$$Recall = \frac{TP}{TP + FN} \tag{5.3}$$

As previously mentioned in Section 2.4.2, the RF importance was applied in this approach. Thus, the metric estimates the variables importance through prediction error. The RF's importance observes the extracted features from samples toward analysis according to trees construction. For estimating the interpretation of image characteristics, the impact is measured by prediction [77].

5.3 Experiment Setup

The strategy called DSIA for classification is based on the "divide and conquer" theory. Due to the difficulty of predicting some types of classification problems, only feature extraction is not enough to predict a given dataset. When it becomes necessary to analyse many characteristics, it often grows impractical due to the relation of input features in ML concerning the computational costs. Thus, the DSIA was proposed as a form of rapid optimisation for high complexity problems in which the feature extraction is not sufficient to obtain high values of accuracy.

Methodologies based on computer vision is considered precise techniques which realise process non-destructive of inspection obtaining a satisfying accuracy, rapid, economic and consistent [105]. The proposed CVS in this approach was constructed to predict raw ham in three different levels: C_1, C_2 and C_3. The CVS overview can be observed in four main steps: image acquisition, pre-processing, feature extraction and classification, as exposed in Figure 18. As mentioned, the proposed DSIA is a technique to improve the classification performance. The approach treats the image feature extraction per smaller region, which enriched the dataset with complementary local information.

In the experiments were applied algorithms developed in the `R` environment to induce models for classification and to assess the performance of DSIA method to predict samples. Briefly, the algorithms description and the corresponding packages used to implement each ML algorithm are described in Table 3. Totalling of two algorithms were evaluated: SVM and RF. Due to the particularities of the problem, the algorithms were chosen for the application [106]. The hyperparameters were used corresponds to the default values of `R` packages in order to support a fair comparison among the algorithms.

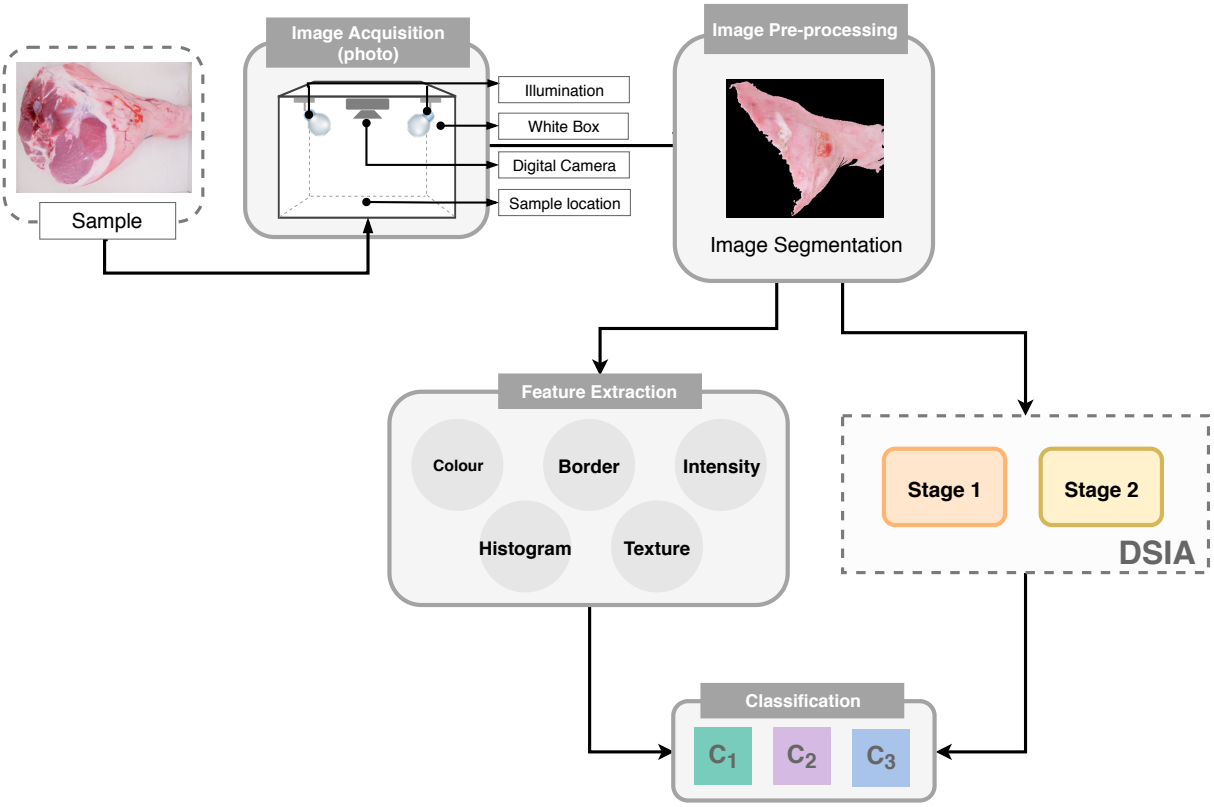


Figure 18 – Traditional Computer Vision System and Dual Stage Image Analysis technique combination for raw ham classification among levels 1, 2 and 3 (C_1, C_2 and C_3).

Table 3 – Machine learning algorithms used in the experiments and corresponding `R` packages.

Algorithm	**Description**	**R package**	**Hyperparameters**
Random Forest (RF)	Combination of decision tree models that provides more accurate prediction [77, 107].	`RandomForest`	ntree=100; mtry=7
Support Vector Machine (SVM)	A statistical learning algorithm [70], have achieved important results in food quality solutions [108].	`e1071`	kernel=polynomial; $\gamma = 0.02$, degree=3

Each type of feature in the sub-image was extracted to quantify the ROI characterisation. The RGB and HSV colour spaces were explored in this study: Red (R), Green (G), Blue (B), Hue (H), Saturation (S), Value (V) colour channels, respectively. The colour channels were analysed separately and additional features were used to auxiliary the prediction. The colour information of an image is characterised by distribution moments which could be used as image feature classification [109]. Concerning colour descriptors, statistical moments from the colour spaces were used to compose the dataset, as the mean and standard deviation. Correlations between channels were generated to improve the power of properties extracted of each sub-sample. Furthermore, other types

of feature from the histogram of the channel were computed, like standard deviation, kurtosis and skewness, totalling in 33 colour features and 21 histogram features.

The intensity information was extracted as two statistical moments for each subsample. The entropy value was calculated from the intensity channel, which is characterised as a statistical measure of randomness based on sizing texture and contrast of grey scale images [110]. Likewise, the texture is an important characteristic of space to identify objects or patterns in an image [54]. Thus, having general applicability, three texture descriptors were used: Local Binary Patterns (LBP), Grey Level Co-occurrence Matrix (GLCM) and Fast Fourier Transform (FFT), this last to uncover frequency domain characteristics. Achieving the narrow peaks of high energy in the spectrum domain by changing the perspective of image analysis is possible to identify periodic patterns in the image [36, 111]. LBP is an operator to describe local image texture features. This method turns the specific local texture in binary vector encoded comparing of a grey-scale pixel and the neighbours [56]. Likewise, GLCM involves the repetition patterns over the image. The pixel neighbourhood is observed towards the periodicity enables to detect patterns with a high level of repeatability [54].

Adding the description of each sample, the border feature was included in the group-of-features. The properties of Sobel and Canny was used to improve the power of prediction. For either type, it was observed values of white pixels present and Hu moments. The Sobel operator is based on edge detection. The differential of this technique is the process which is produced from the original image where the gradient of grey scale levels is highlighted. The edges are recognised through comparison of amplitude values to threshold level predefined [112]. The Canny edge detector is based on computing squared gradient magnitude, which involves the local threshold peak. This method derives an optimal operator from minimising the probability to detect multiple edges and the distance of the reported edge from the true edge [61].

The list including hand-crafted features used in the solution is showed in Table 4.

Table 4 – List of all image features used in the proposed approach for raw ham classification.

No.	Type	Name	Description
1	Colour	cor_RG	Correlation between Red and Green channel
2	Colour	cor_RB	Correlation between Red and Blue channel
3	Colour	cor_RH	Correlation between Red and Hue channel
4	Colour	cor_RS	Correlation between Red and Saturation channel
5	Colour	cor_RV	Correlation between Red and Value channel
6	Colour	cor_RI	Correlation between Red and Intensity channel
7	Colour	cor_GB	Correlation between Green and Blue channel
8	Colour	cor_GH	Correlation between Green and Hue channel
9	Colour	cor_GS	Correlation between Green and Saturation channel
10	Colour	cor_GV	Correlation between Green and Value channel
11	Colour	cor_GI	Correlation between Green and Intensity channel
12	Colour	cor_BH	Correlation between Blue and Hue channel
13	Colour	cor_BS	Correlation between Blue and Saturation channel
14	Colour	cor_BV	Correlation between Blue and Value channel
15	Colour	cor_BI	Correlation between Blue and Intensity channel
16	Colour	cor_HS	Correlation between Hue and Saturation channel
17	Colour	cor_HV	Correlation between Hue and Value channel
18	Colour	cor_HI	Correlation between Hue and Intensity channel
19	Colour	cor_SV	Correlation between Saturation and Value channel
20	Colour	cor_SI	Correlation between Saturation and Intensity channel
21	Colour	cor_VI	Correlation between Value and Intensity channel
22	Colour	mean_H	Mean of Hue channel
23	Colour	std_H	Standard Deviation of Hue channel
24	Colour	mean_S	Mean of Saturation channel
25	Colour	std_S	Standard Deviation of Saturation channel
26	Colour	mean_V	Mean of Value channel
27	Colour	std_V	Standard Deviation of Value channel
28	Colour	mean_R	Mean of Red channel
29	Colour	std_R	Standard Deviation of Red channel
30	Colour	mean_G	Mean of Green channel
31	Colour	std_G	Standard Deviation of Green channel
32	Colour	mean_B	Mean of Blue channel
33	Colour	std_B	Standard Deviation of Blue channel
34	Intensity	mean_I	Mean of Intensity channel
35	Intensity	std_I	Standard Deviation of Intensity channel
36	Intensity	entropy_I	Entropy of Intensity channel
37	Histogram	std_hist_H	Standard Deviation of Histogram of Hue channel
38	Histogram	kurt_hist_H	Kurtosis of Histogram of Hue channel
39	Histogram	skew_hist_H	Skewness of Histogram of Hue channel
40	Histogram	std_hist_S	Standard Deviation of Histogram of Saturation channel
41	Histogram	kurt_hist_S	Kurtosis of Histogram of Saturation channel
42	Histogram	skew_hist_S	Skewness of Histogram of Saturation channel
43	Histogram	std_hist_V	Standard Deviation of Histogram of Value channel
44	Histogram	kurt_hist_V	Kurtosis of Histogram of Value channel
45	Histogram	skew_hist_V	Skewness of Histogram of Value channel
46	Histogram	std_hist_R	Standard Deviation of Histogram of Red channel
47	Histogram	kurt_hist_R	Kurtosis of Histogram of Red channel
48	Histogram	skew_hist_R	Skewness of Histogram of Red channel
49	Histogram	std_hist_G	Standard Deviation of Histogram of Green channel
50	Histogram	kurt_hist_G	Kurtosis of Histogram of Green channel
51	Histogram	skew_hist_G	Skewness of Histogram of Green channel
52	Histogram	std_hist_B	Standard Deviation of Histogram of Blue channel
53	Histogram	kurt_hist_B	Kurtosis of Histogram of Blue channel
54	Histogram	skew_hist_B	Skewness of Histogram of Blue channel
55	Histogram	std_hist_I	Standard Deviation of Histogram of Intensity channel
56	Histogram	kurt_hist_I	Kurtosis of Histogram of Intensity channel
57	Histogram	skew_hist_I	Skewness of Histogram of Intensity channel
58	Border	nump_sobel	Number of white pixels in Sobel image
59-65	Border	hu_sobel1 - hu_sobel7	Hu Moments of Sobel image
66	Border	nump_canny	Number of white pixels in Canny image
67-73	Border	hu_canny1 - hu_canny7	Hu Moments of Canny image
74-83	Texture	lbp_0 - lbp_9	LBP Vector
84	Texture	com_entropy	Entropy of Co-occurence Matrix
85	Texture	com_inertia	Inertia of Co-occurence Matrix
86	Texture	com_energy	Energy of Co-occurence Matrix
87	Texture	com_correlation	Correlation of Co-occurence Matrix
88	Texture	com_homogeneity	Homogeneity of Co-occurence Matrix
89	Texture	FFT_energy	Energy of FFT
90	Texture	FFT_entropy	Entropy of FFT
91	Texture	FFT_inertia	Inertia of FFT
92	Texture	FFT_homogeneity	Homogeneity of FFT

6 RESULTS

The results acquired were arranged in order to compare the proposed approach performance for classification among three different levels of raw ham, as C_1, C_2 and C_3. Furthermore, the outcome achievement of DSIA was presented in comparison to the straightforward image classification (Traditional CVS) to evidence the advantages of the proposed method. Subsequently, the results were driven by a discussion related to the imaging resources, as well as RF importance, which assists in analysing the types of each raw ham in the proposed approach. Lastly, the challenges of DSIA are explored.

6.1 Comparison between Traditional CVS and DSIA performances

Table 5 summarises the obtained results for algorithms over the datasets and distinct methods. The mean performance metrics are considered for each possible combination of algorithms as Traditional CVS and DSIA. Concerning Traditional CVS, RF achieved superior results, slightly similar to SVM. Analysing the performance to distinguish raw ham levels, the results were not enough of the sample's composition in dataset being satisfactory accuracy. Nevertheless, it emphasisis that is necessary strategies for learning and to improve raw ham classification. Thus, the DSIA was proposed to distinguish levels.

Table 5 – Comparison of metrics and thresholds found with different methods and algorithms (RF and SVM) over Prediction dataset.

Method	**Sub-region Resolution**	**RF**					**SVM**				
		Accuracy	**Precision**	**Recall**	λ_1	λ_2	**Accuracy**	**Precision**	**Recall**	λ_1	λ_2
Traditional	-	63.10%	44.44%	44.55%	-	-	61.91%	43.20%	43.76%	-	-
DSIA	10x10	55.95%	53.44%	68.95%	0.1802	0.2309	61.91%	56.68%	73.18%	0.1858	0.2384
	25x25	83.33%	72.91%	82.45%	0.0465	0.0884	82.14%	71.89%	81.60%	0.0454	0.0810
	50x50	88.10%	92.71%	85.83%	0.0976	0.1827	83.33%	82.63%	88.33%	0.1034	0.1731
	75x75	79.76%	73.02%	74.16%	0.1491	0.2643	78.57%	72.77%	79.14%	0.1475	0.1301
	100x100	75.00%	66.87%	70.77%	0.1410	0.2475	84.52%	79.00%	71.60%	0.1163	0.2668

CVS coupled with DSIA presented a high performance in the majority of the tested models, except for those calculated with a sub-image resolution of 10x10 pixels. The best results were obtained with RF algorithm considering sub-images of 50x50 pixels, leading to 88.10% of accuracy, a precision value equal to 92.71% and 85.83% of recall. The worst resolution was 10x10 pixels, achieving an accuracy value of 55.95% with RF and 61.91% with SVM, and consequently low values of precision and recall. These results show that considering a too small sub-sampling area causes a loss of important characteristics, which

are fundamental to describe the problem at hand. Conversely, larger sub-sampling areas account for more complex spatial patterns, which give the possibility of accounting also for other textural features and/or defects different from veining, causing a decrease in the predictive ability of the method. Considering the traditional CVS approach, RF achieved superior results with 63.10% of accuracy, 44.44% of precision and recall value of 44.55%, slightly similar to SVM (61.91%, 43.20% and 43.76%, respectively).

From the results exposed in Table 5, it is observed different performances between the applied techniques of ML, which summarises the metrics achieved in each algorithm over the datasets with DSIA method and Traditional CVS. Regarding the performance of the algorithms over previously set, it is possible to observe superior accuracy results in the DSIA method. The DSIA was applied in different image resolutions of the segmentation process to reinforce the effectiveness of the proposed approach to problems solution in relation to the Traditional CVS method. Furthermore, it is necessary to analyse the behaviour of the proposed approach concerning different types of resolution in order to find the optimal performance.

As reported in Chapter 4, the proposed method consists of dividing complexity problems into smaller problems moreover obtaining a high of further understanding, as well as class prediction. The raw ham classification is considered a complex problem resolution being based on quantifying the presence of veins in relation to the ROI. This is considered as a defect which decreases the quality of the final product. As a pattern classification, the amount of veins per region can be distinguished between different classes according to the quality parameters. Thus, the proposed approach was carried out with analysing two types of distinct datasets. The First Stage refers to the set which samples were divided into small regions being classified as *veining* and *no veining* defects. The raw ham segmentation in minor regions was applied in five different sizes composing distinct datasets.

Summarising the *veining* and *no veining* defects over prediction set present in each raw ham sample, as shown in Figure 19. In order to obtain the classification of the samples through the DSIA method, it is necessary to analyse the relation of the sub-samples classified as *veining* in relation to the total number of sub-samples of each image. Analysing the histogram of the amount of samples classified as defective by the total of the sub-samples, the frequency ratio is obtained. Thus, it is possible to find a threshold among the different levels of raw ham. The density graphs are exposed to propitiate the knowledge of the form of data distribution over datasets. The values of λ_1 and λ_2 play an important role in the final classification. In Table 5 both values are shown for the different resolutions and models, while Figure 19 reports the corresponding PDFs. This figure shows that the best predictive performance obtained, considering the resolution of 50x50 pixels is related to the lower overlap between the PDFs of the three levels. In the

same way, RF 25x25, SVM 25x25 and RF 50x50 presented the smallest standard deviation in the number of veining sub-regions per level to compose their PDF.

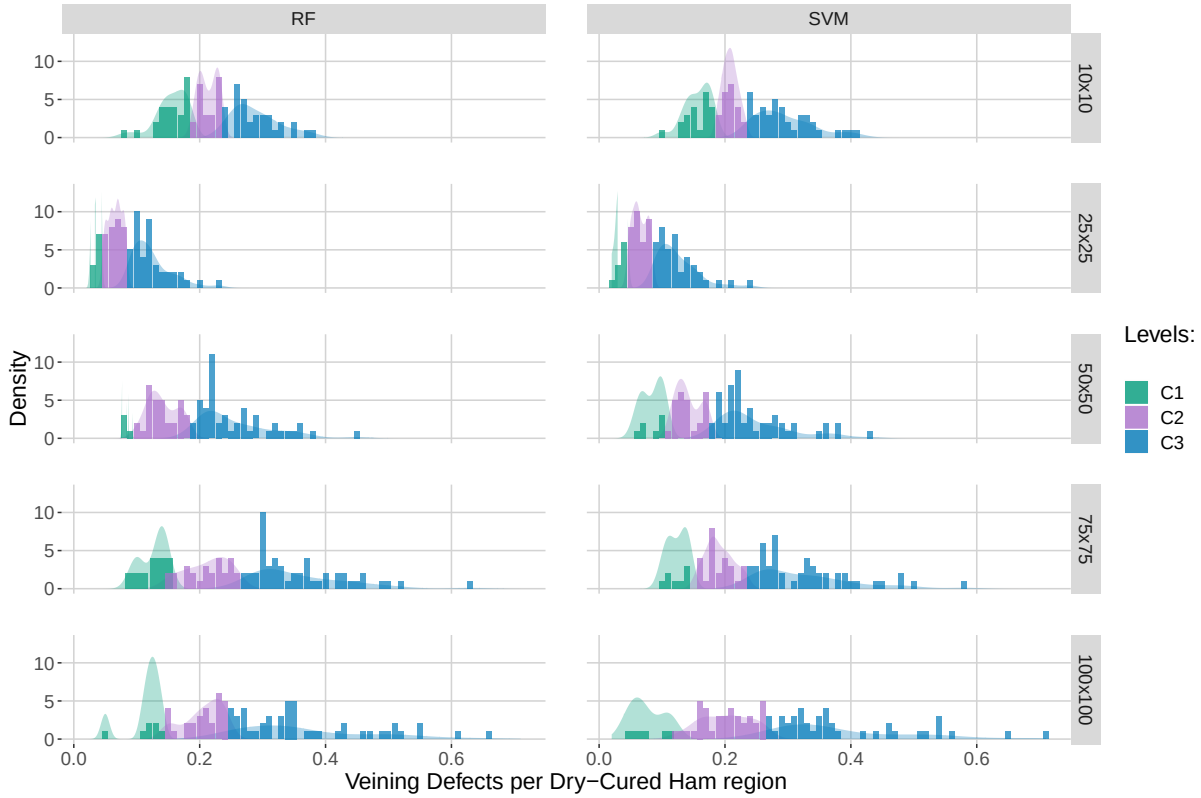
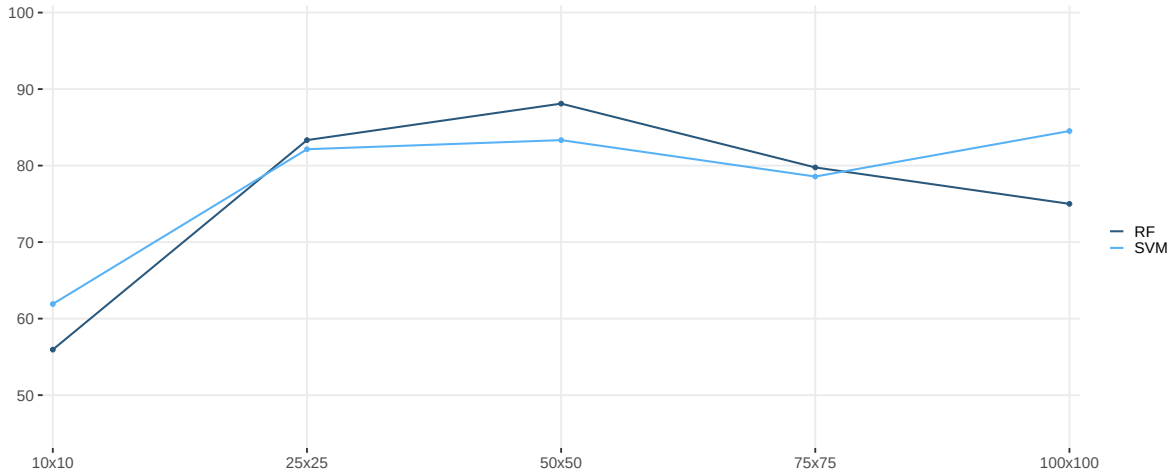


Figure 19 – Veining defects density per raw ham.

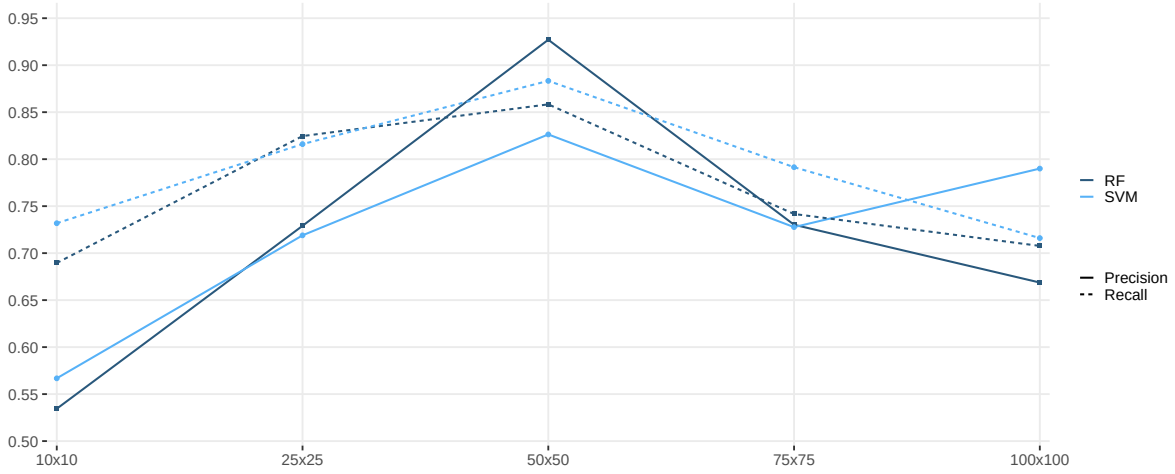
Applying to distinguish raw ham levels, the First Stage of DSIA output allowed to identify each sample class through the threshold. The Second Stage consists of finding the respective λ_1 and $\lambda 2$, which separates the relation of quantity *veining* defects per region between levels. The threshold is observed on the training set and applied in prediction set. Therefore, the obtained separation of levels can be observed in Figure 19. The relation of *veining* defects per raw ham has expressive behaviour in which it is possible to be observed in relation to the levels of samples found and their density by class is possible to distinguish the different levels.

The variability of the results in relation to the samples is related to the prediction of the sub-images related to the First Stage. In this way, it is observed that there are different behaviours in the datasets analysed. Notably there is a higher concentration between sets of resolution 10x10 and 25x25 pixels in sub-samples in comparison of the prediction sets. This distribution is a consequence of using a lower resolution, which interferes directly in the identification of the classes between *veining* and *no veining* defects. This fact has a high influence in the Second Stage, which differs levels task. The performance trend of between the RF and SVM algorithms are similar, slightly reinforces the image resolution directly impacts the final results. Testing image size variability is justified in order to find the high accuracy problem solution.

Figure 20 expose the accuracy, precision and recall over prediction datasets resolution. Analysing the results over 10x10, 25x25, 50x50, 75x75 and 100x100 pixels, it propitiates to conclude the high performance is observed in 50x50 pixels which have high precision with RF algorithm. The proposed approach has shown the smaller sub-samples loss significant characteristics, which is decisive to describe the problem task. Moreover, the biggest sub-samples concentrate more *veining* defects per region that it causes in the expressive loss of accounting of the relation of veins present by region, consequently, decreases the predictive power of the method.



(a) Accuracy



(b) Precision and Recall

Figure 20 – Accuracy, Precision and Recall plot for different datasets of raw ham.

6.2 Analysis of DSIA prediction

Summarising sub-samples prediction per set, the Table 6 shows Confusion Matrix. The comparison of DSIA method was performed using the following machine learning algorithms RF and SVM. Concerning the results, as previously mentioned, the step of

sub-samples is implied directly to find the levels C_1, C_2 and C_3. Smaller regions difficult the identification of the characteristics which inducting the lowest prediction. For all algorithms, experiments were performed executing re-sampling strategy considering the no veining label. This sampling approach attempts to retain the same portions of labels combination in all the subsets for both datasets, since the amount of sub-samples from *no veining* regions is more expressive.

Table 6 – Confusion matrix of RF and SVM classification models over sub-regions from predictions set.

Algorithm	**Dataset**		**Predicted**	
			Veining	***No Veining***
RF	10x10	*Veining*	21080	2021
		No Veining	109380	437146
	25x25	*Veining*	3129	48
		No Veining	9460	123447
	50x50	*Veining*	2564	43
		No Veining	3811	24394
	75x75	*Veining*	1690	130
		No Veining	1963	8551
	100x100	*Veining*	962	108
		No Veining	851	4250
SVM	10x10	*Veining*	21101	2000
		No Veining	116251	430275
	25x25	*Veining*	3136	41
		No Veining	9418	123489
	50x50	*Veining*	2532	85
		No Veining	3746	24449
	75x75	*Veining*	1677	143
		No Veining	1632	8882
	100x100	*Veining*	973	97
		No Veining	830	4271

Applying DSIA technique, it was possible to improve the average performance of the classification task from 63.10% in the Traditional CVS to 88.10%. It is important to highlight the lowest performance of each level was at the intermediate level (C_2). It can be associated with exposed results in Table 6. The lowest performance can be observed by finding several samples in similar regions. This means that finding a threshold between samples can influence the distinction between levels in nearby areas.

Improving of classification accuracy was obtained by applying the DSIA technique, as exposed in Figure 21. Though, the Traditional CVS has a high performance in C_2 level

prediction in comparison of other levels and techniques applied on. The result can be influenced by variability of characteristics presented in each sample, which has a mark control stamp more expressive in C_1 and C_2 because it has fewer appearance of veins at the skin. Thus, the induction setup is propitiated to induce the samples, decreasing the other levels of prediction erroneously.

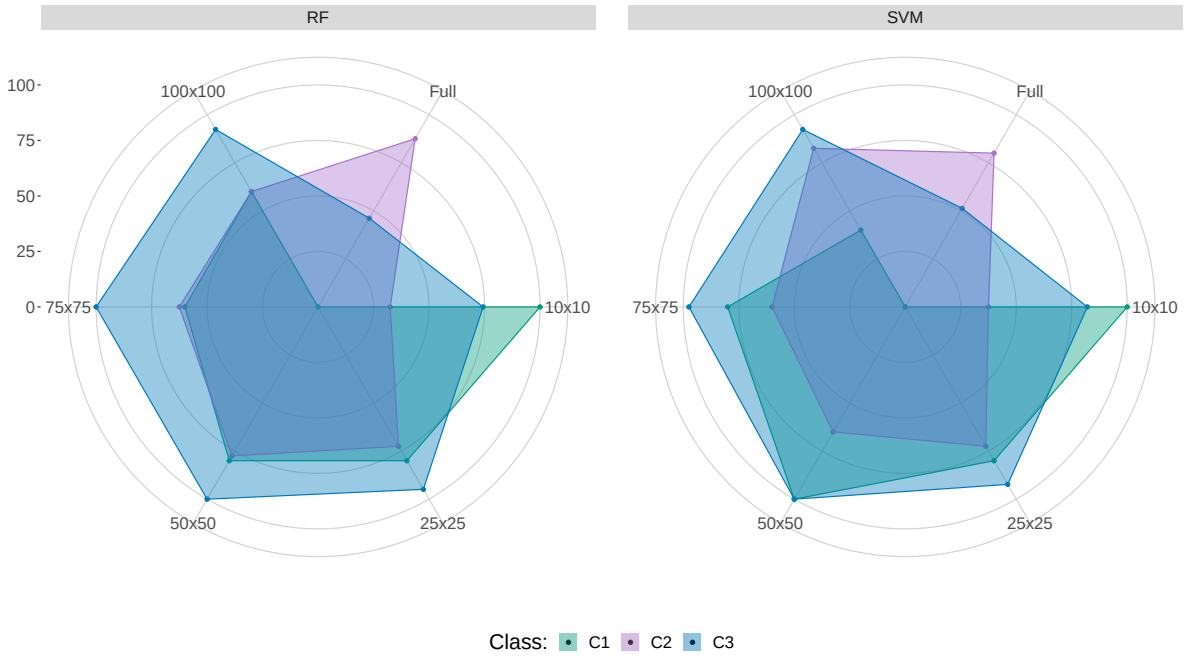


Figure 21 – Accuracy levels of applied methods on.

The influence of carcass marks was minimised with the application of the DSIA method, as can be observed in Table 7. To observe the complexity of each veining level, the table reports the precision and recall values for each sub-region and considering C_1, C_2 and C_3 separately from each other. Concerning the exposed results, it is possible to conclude the DSIA propitiates a high precision performance in comparison of Traditional CVS which reduces the insignificant importance parts of the ROI, considering the characteristics effectively to resolve the problem task. Using the DSIA method, it was possible to improve the average performance of raw ham classification. It is important to highlight that RF stood out with more high accuracy in each levels prediction. The intermediate levels in classification tasks with gradual class dispersion leads to lower performance in comparison to boundary levels. Traditional CVS obtained a biased recall result for veining level C_2, as visible in Table 7 when observing precision and recall metrics and the classification of C_1 level. DSIA with both RF and SVM obtained comparable results, with C_3 being the most easily predictable level, followed by C_1 and by the intermediate class (C_2).

Considering the results of traditional CVS, slightly better performances were observed with RF in comparison to SVM. The traditional method was not useful in distin-

guishing among the three classes. The inaccurate classification performances were mainly influenced by the fact that the CVS was not able to correctly classify the samples belonging to level C_1, since they were all assigned to class C_2. The image analysis from the whole ROI smooths the characteristics of the vein presence, degrading the image features that allow specifying the properties of the veining level. Therefore, the usage of sub-regions as proposed in DSIA was able to extract the image features to create a better classification approach. Sub-regions with low resolution (10x10 pixels) obtained the worst classification results, mainly for class C_2.

Table 7 – Precision and Recall of RF and SVM classification levels per method.

Method	**Sub-region Resolution**		**RF**		**SVM**	
			Precision	**Recall**	**Precision**	**Recall**
Traditional	-	C_1	00.00%	00.00%	00.00%	00.00%
		C_2	58.33%	87.50%	58.18%	80.00%
		C_3	75.00%	46.15%	71.43%	51.28%
DSIA	10x10	C_1	18.52%	100.00%	23.81%	100.00%
		C_2	56.52%	32.50%	68.18%	37.50%
		C_3	85.29%	74.36%	78.05%	82.05%
	25x25	C_1	40.00%	80.00%	40.00%	80.00%
		C_2	90.63%	72.50%	87.88%	72.50%
		C_3	88.10%	94.87%	87.80%	92.31%
	50x50	C_1	100.00%	80.00%	71.43%	100.00%
		C_2	96.88%	77.50%	100.00%	65.00%
		C_3	81.25%	100.00%	76.47%	100.00%
	75x75	C_1	50.00%	60.00%	50.00%	80.00%
		C_2	92.59%	62.50%	92.31%	60.00%
		C_3	76.47%	100.00%	76.00%	97.44%
	100x100	C_1	42.86%	60.00%	66.67%	40.00%
		C_2	82.76%	60.00%	84.62%	82.50%
		C_3	75.00%	92.31%	85.71%	92.31%

6.3 Misclassified samples

To identify outliers or badly labelled samples, it was created a heatmap which shows the classification results obtained from all of raw ham samples. Figure 22 exposes the results which it is possible to observe some misclassified samples comparing all performed techniques with RF and SVM algorithms in raw ham levels. The different shades of blue represent the different veining levels: a lighter blue colour indicates samples predicted as belonging to level C_1, a darker blue colour indicates samples predicted as belonging to level C_3, while the intermediate blue colour is associated to samples predicted as level C_2. Analysing the incorrect classified samples of three levels on the heatmap, it is possible to indicate the C_2 level as more complex to be predicted in comparison with the C_1 and C_3

levels, since it resulted to be the most misclassified class. For the classification algorithms, it is possible to observe similar behaviours among the samples.

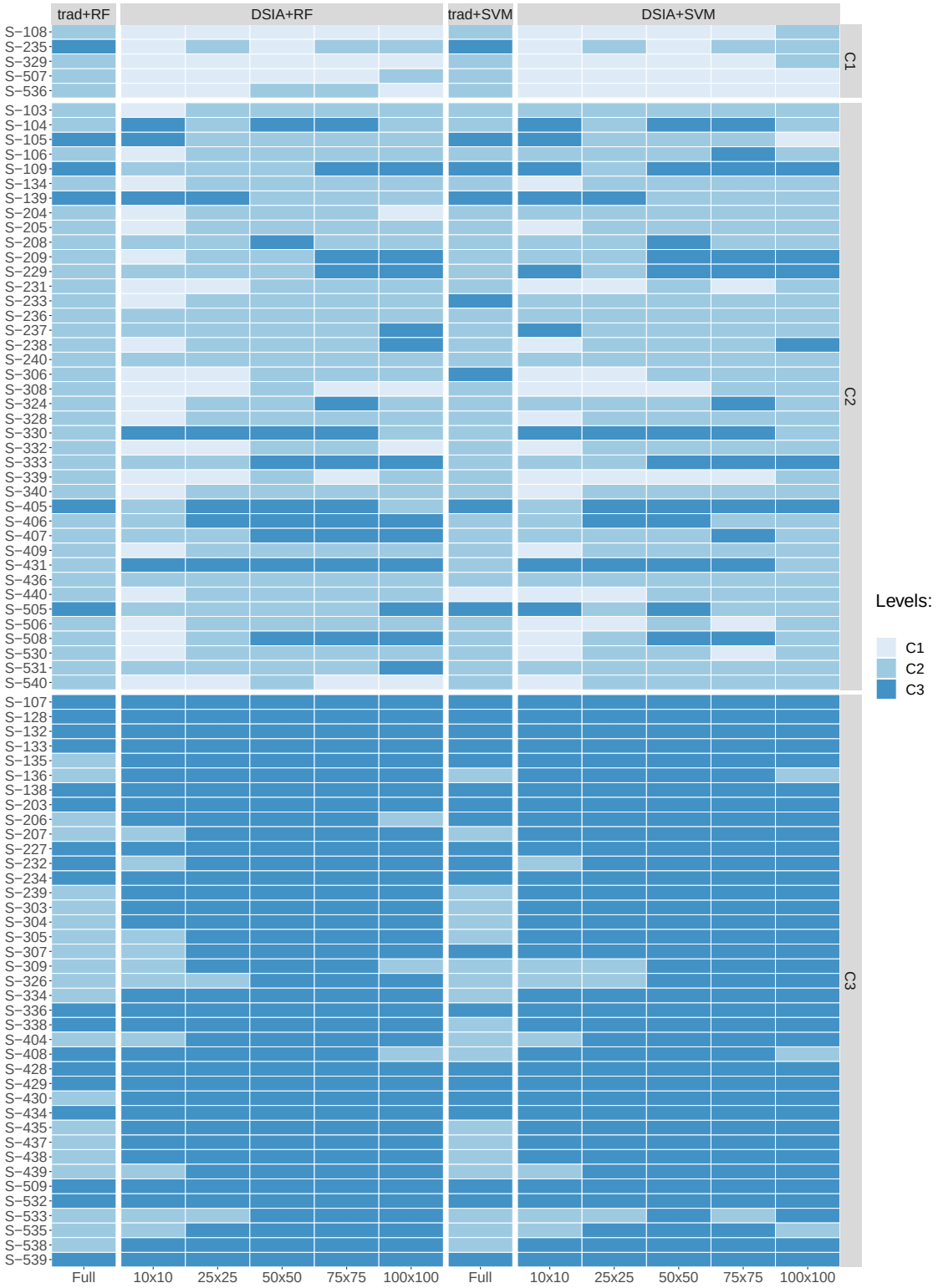


Figure 22 – Accuracy heatmap levels C_1, C_2 and C_3 of RF and SVM over Prediction dataset comparing DSIA technique. Rows correspond the samples each level. Light blue colour represent a C_1 predicted level, while intermediate blue colour C_2, and dark blue the C_3.

Considering Traditional CVS results, a better performance was observed in RF compared to SVM. Most of the incorrectly classified samples belonged to the extreme levels, that is, the levels C_1 and C_3. It is observed the Traditional method was not effective in distinguishing among three classes. The performance was impaired, resulting in a similar classification in its majority in class C_2. Experiments with whole samples induce intermediate classes because they do not have sufficient characteristics to be differentiated among the raw ham levels. The image analysis smooths the characteristics of the vein presence due to the fact the region of interest presents strong texture characteristics in relation to quality control marks and specifics properties of the raw ham as sample S-405 shown in Figure 23. Samples S-330 and S-508 were incorrectly classified by the proposed DSIA approach due to regions that resemble vein structures. These regions increased the prediction ratio of the *veining* defects present in the sample, consequently decreasing the performance capability in the analysis of the level it belongs to.

(a) S-330 (b) S-405 (c) S-508

Figure 23 – Samples of raw ham from C_2 level incorrectly classified.

Despite the fact, DSIA has high precision in distinguishing the levels, including the extreme classes C_1 and C_3. The image dataset with low resolution (10x10 pixels) obtained the worst index of samples which obtained more raw ham incorrectly classified, mainly, in the class C_2. The Figure 22 shows similarity of C_1 and C_2 over prediction set in which samples C_2, were mostly classified as C_1. Thus, as the observed feature, which most describes the problem is related to the co-occurrence matrix entropy, using low-resolution images decrease the detail of the textures, consequently decreases the performance. Therefore, it is possible to observe a significant improvement further datasets used to raw ham prediction. The DS with RF and SVM achievement was similar. In this way, there is an expressive advantage in DSIA classification in comparison to the Traditional CVS. The best-observed result among the prediction of the three levels refers to the dataset with 50x50 pixels resolution per image. As the size of the segmented images has been increased, there is a decline in performance. Likewise, since the low resolution impairs the identification of the composite regions of *veining* and *no veining* defects, large size images decrease the concentration of the presence of *veining* defects per region, thus, decreasing the power of prediction.

6.4 Exploring the influence of sub-regions

According to the exposed results and performance, the DSIA demonstrated the superior capability of prediction in comparison to the applied Traditional CVS. For complex problems which the traditional method is not feasible, the division of the problem into smaller ones increases the detail of information about the problem task. By dividing the images into two stages, regions which display *veining* defects are identified towards the relationship of the regions prediction is analysed by finding a threshold between the samples. The raw ham categorisation levels become efficient in order to reduce the impact of non-relevant sample properties which do not assist in the identification of the regions. Therefore, the DSIA method provides a greater predictive power to distinguish the levels.

Figure 24 summarises at heatmap the description and classification in *veining* and *no veining* defects used to distinguish each level with DSIA method. Analysing each of the datasets differentiated by the size of the resolution made in the segmentation process to divide the problem task, the gradient varies from dark blue (worse prediction) to light blue (better prediction). Observing the performance obtained for the dataset of 25x25 was the best (93%), as a sequence, 50x50 with 88% accuracy. The divergence found between the accuracy of the First Stage method as well as the Second Stage is due to the resolution of the images, which dataset separation in the first step between *veining* and *no veining* defects. The automatic process splits the problem task into *veining* and *no veining* defects classification to analyse the percentage of veining presence in each sub-sample. The categorisation considers percentages above 50% of the resolution according to the sample, exposed in Chapter 4. Moreover, for low order resolution, there is an instability in relation to the identification of the regions with the presence of marked characteristics reflecting the low sub-samples confidentiality.

Concerning exposed results in this section, the segmentation of the images into sub-regions reflects directly on the expected performance. By using "divide and conquer" as well as running on decision trees, very small divisions end up interfering negatively in expected performance. Analysing the data mentioned above, it can be seen through the sub-samples relation of *veining* and *no veining* defects that it becomes unstable for resolutions smaller than 25x25 pixels. Instability is confirmed by comparing the performance of datasets with respect to their respective heatmaps. The 25x25 dataset is more accurate than the other datasets, however, it achieves low performance when related to the second stage to find the threshold between the different classification levels. As the size of the segmented samples decreases, the lower the ratio among sub-regions present from the samples. This consideration occurs due to the fact that very small sub-samples induce errors because it is not possible to use significant resources to identify determining characteristics between regions. Thus, it is noted that only about 2% of the sub-images represent regions containing *veining* defects. This fact corroborates for a greater accuracy

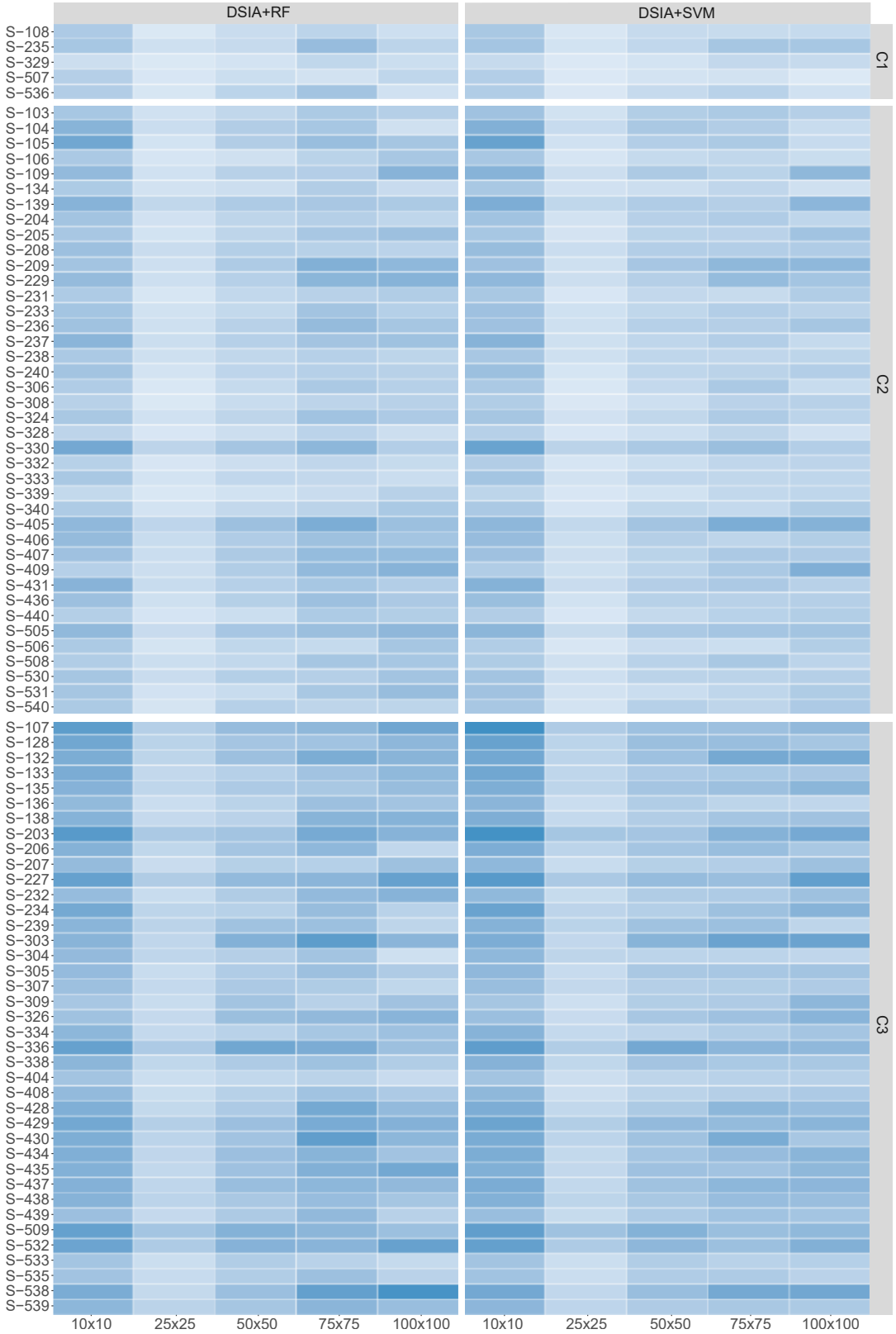


Figure 24 – Accuracy heatmap *veining* and *no veining* defects of RF and SVM over Prediction dataset comparing Dual Stage technique. Rows correspond the samples each level. Light blue colours represent a better performance (higher accuracy), while dark blue colours the opposite.

in the data set. The homogeneity between the sub-samples can be observed according to results previously demonstrated. There is a few difference amog the relationships obtained, which generates a greater concentration of the samples, given the discrepancy between the sub-classes.

Therefore, the performance of the proposed approach demonstrates to be superior to the Traditional CVS for the study case with raw ham samples. The accuracy obtained is higher at 25% which proves to be effective in application tasks where only attribute extraction is not sufficient for prediction. Recognition of complex patterns need robust methods for better identification and distinction among classes. The texture, in this context, becomes paramount for efficient performance. Through the "divide and conquer" strategy, one can obtain satisfactory results by splitting the problem into distinct sub-regions to identify the different levels to be predicted.

Regarding some insights about image features, sub-regions with low resolutions are not able to provide a suitable pattern towards its extraction. For instance, the most important features are related to entropy of the co-occurrence matrix. When extracting these features from low-resolution sub-regions there is a decreasing of texture structure, and consequently the descriptive performance is compromised. More precisely discussion about the relevance of features is made in Section 6.5.

6.5 Feature Importance analysis

As mentioned in a previous Section 5.2, the RF importance propitiates to evaluate the most relevant features in prediction assignments. Regarding the obtained importance values, Figure 25 exposes the relevant features to classify raw ham samples. The descriptions analysis are categorised (Table 4): Border, Colour, Histogram, Intensity and Texture, according to the relevance showed respectively. The highest important features were from texture: entropy of the co-occurrence matrix. The co-occurrence matrix is defined as the tabulation of combinations of different pixel intensity values in grey scale levels that occur in the image. Associating entropy with a degree of dispersion of grey scale; thus, it describes the texture through a set of characteristics considering multiple directions. Sequentially, other important features were from Colour: standard deviation of R channel and V channel (Red from RGB and Value from HSV colour spaces) and Intensity: standard deviation of intensity channel, which has shown the most relevant explaining features with more than 50 of importance. However, all features presented impact for the classification procedure.

Dealing with raw ham levels, the entropy and standard deviation of colour space channels, Red and Green RGB channel and Value HSV channel were the most discriminative image features. Therefore, the values of intensity and FFT of entropy got higher

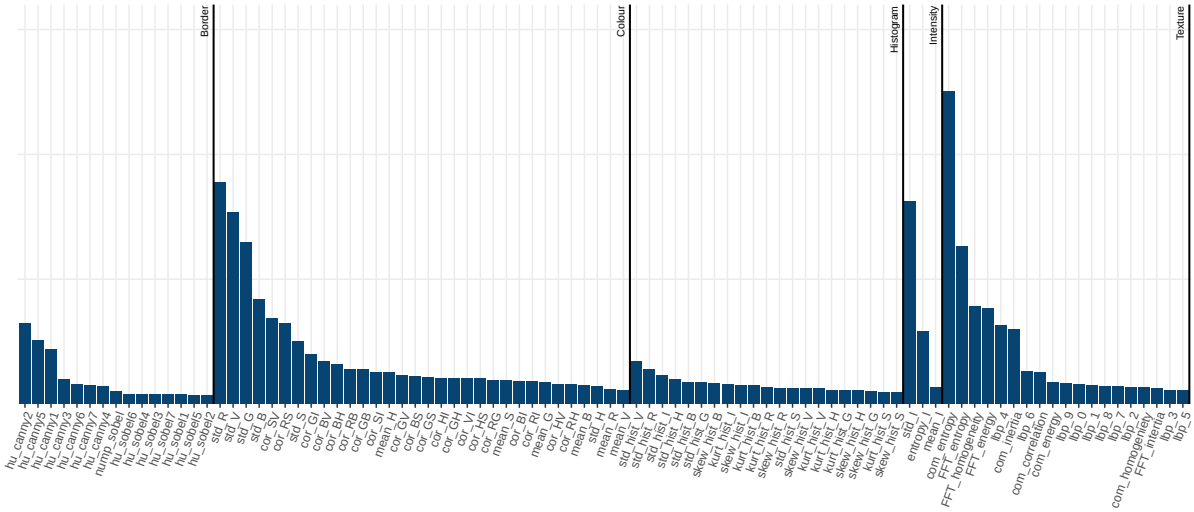


Figure 25 – Features importance per descriptors.

importance, as well as the FFT homogeneity. As expected, Texture features were significant to predict the samples. Indeed, some texture features of LBP metrics were efficient to predict sample varieties, and also could be related to the *veining* defects in raw ham. The presence of texture characteristics can assist in distinguishing *veining* and *no veining* defects sub-samples, consequently, increasing the levels task. Analysing the importance of the features propitiates to obtain the knowledge of the set, including the relevance of samples characteristics which differs the pattern presented in each classified image. Actually, also colour features on the average are almost equally important. Rather than focusing on a single group, observing the fact that useful information is brought by the synergy between different types of features (texture, colour, intensity). Indeed, some texture features of LBP metrics were efficient to predict sample levels, and also could be related to the *veining* defects in raw hams.

6.6 DSIA challenges

Concerning complex image patterns, applying the proposed approach in raw ham samples for distinguishing among three levels has a significant improvement in the predictive performance when using DSIA embedded into a CVS. Summing up the exposed results, the best achieved result refers to the dataset with 50x50 pixels of resolution per sub-region. As the size of sub-region has been increased, a decline in performance was observed. While the low resolution impairs the sub-regions identification between *veining* and *no veining* defects, high size images decrease the concentration of *veining* defects per region; therefore, the classification performance decreases in both cases. Thus, DSIA is highly dependent on a suitable subregion resolution for obtaining promising results.

According to the reported results, DSIA demonstrated a superior capability of pre-

diction, compared to using CVS without the DSIA approach. For complex image problems in which the CVS method is not feasible, the image division into smaller regions increases the detail of information about the object of interest. By dividing the raw hams images into sub-regions, those which display *veining* defects are identified with the usage of a thresholding method to find a suitable number of regions that represent a defect level. On the other hand, the First Stage of DSIA demands images with veining patterns identified to label the sub-regions automatically. This is an additional requirement and disadvantage of DSIA since in the traditional CVS the ROI classification requires just image features and labels concerning the whole sample.

7 CONCLUSION

This work presents a computer vision-based technique called Dual Stage Image Analysis. The DSIA refers to the "divide and conquer" principle being applied in tasks of complex problems that distinguish the samples are made through the quantification of determining characteristics. The proposed approach divides the problem into sub-problems to be predicted in order to provide better understanding and information of the regions. Dividing the samples into subsets using image segmentation techniques provides greater knowledge through fragmentation. Thus, the concentration of specific characteristics becomes evident, providing individual prominence. In general, traditional CVS method uses information extraction and image processing to identify and recognise patterns among samples. However, some tasks require robust methods in which only information obtained from the images render the returns uncertain with significantly unsatisfactory results. Recognition of complex image patterns requires mechanisms that allow the growth characteristics present in specific regions of the samples, which is not always easy to perceive.

Amount of specific properties present in the samples are not often able to induct models for dealing with distinct classes. For this reason, the proposed approach is based on the computer vision system to deal with challenge of improving prediction performance in classification tasks. The proposed technique was tested for predicting raw ham samples among three different levels. The results were compared using the accuracy obtained between Traditional CVS and the DSIA. Through the metrics applied on, it was possible to observe that the DSIA presents predictive performance superior to the traditional method. To find a solution with superior predictive capability, different image resolutions were tested in the segmentation process. Considering the steps in the proposed approach, the DSIA considers local features to be extracted, emphasising the most important properties to improve performance compared to traditional CVS. Moreover, sub-dividing the problem in sub-problems propitiates more accurate results concerning complex image pattern tasks in such a challenge to deal with. Lastly, the best performance shows the high superior ability of prediction across datasets, being 25% upper accuracy then the traditional method compared.

DSIA resulted in a significant step of a CVS for the identification of veining defect since it was capable of improving the classification performance considerably. This alternative is suitable for production lines in the ham industry and does not require a sophisticated environment for image acquisition based on the RGB imaging. A variety of research works dealing with similar tasks were not able to achieve the proper image classification for ROI with the presence of uncontrolled defects such as fire marks, traumatic hematomas, nontraumatic hematomas, scratches or meat stamps. Recognising complex

image patterns is a challenge not restricted to food evaluation, but a claiming from the diverse research fields toward improving the comprehension and performance of automated solutions supported by computer vision. The "divide and conquer" strategy adopted by DSIA opens new opportunities in the analysis of images composed by complex textures.

Future research work aims at comparing the current approach with more recent machine vision systems.

BIBLIOGRAPHY

[1] KLETTE, R. et al. *Computer vision*. [S.l.]: Springer-Verlag New York, 1998.

[2] NIXON, M.; AGUADO, A. S. *Feature extraction and image processing for computer vision*. [S.l.]: Academic Press, 2012.

[3] RUSS, J. C. *The image processing handbook*. [S.l.]: CRC press, 2016.

[4] SONKA, M.; HLAVAC, V.; BOYLE, R. *Image processing, analysis, and machine vision*. [S.l.]: Cengage Learning, 2014.

[5] BURGES, C. J. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, Springer, v. 2, n. 2, p. 121–167, 1998.

[6] SANTOS, E. M. dos. *Teoria e aplicação de support vector machines à aprendizagem e reconhecimento de objetos baseado na aparência*. Tese (Doutorado) — Universidade Federal da Paraíba, 2002.

[7] ISAYED, A. *Classification of Error Related Potential (ErrP) in P300-Speller*. Tese (Doutorado) — Palestine Polytechnic University, 2015.

[8] ANTONELLI, A. et al. Automated evaluation of food colour by means of multivariate image analysis coupled to a wavelet-based classification algorithm. *Analytica Chimica Acta*, Elsevier, v. 515, n. 1, p. 3–13, 2004.

[9] BROSNAN, T.; SUN, D.-W. Improving quality inspection of food products by computer vision—-a review. *Journal of food engineering*, Elsevier, v. 61, n. 1, p. 3–16, 2004.

[10] BARBON, A. P. A. da C. et al. Development of a flexible computer vision system for marbling classification. *Computers and Electronics in Agriculture*, Elsevier, v. 142, p. 536–544, 2017.

[11] LOPES, J. F. et al. Computer vision classification of barley flour based on spatial pyramid partition ensemble. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 19, n. 13, p. 2953, 2019.

[12] MITTAL, G. S. *Computerized control systems in the food industry*. [S.l.]: Marcel Dekker, Inc., 1996.

[13] PU, H. et al. Classification of fresh and frozen-thawed pork muscles using visible and near infrared hyperspectral imaging and textural analysis. *Meat Science*, v. 99, p. 81–88, 2015.

[14] FOCA, G. et al. Classification of pig fat samples from different subcutaneous layers by means of fast and non-destructive analytical techniques. *Food research international*, Elsevier, v. 52, n. 1, p. 185–197, 2013.

[15] MANCINI, R.; HUNT, M. Current research in meat color. *Meat science*, Elsevier, v. 71, n. 1, p. 100–121, 2005.

[16] BARBON, A. P. A. et al. Storage time prediction of pork by computational intelligence. *Computers and Electronics in Agriculture*, Elsevier, v. 127, p. 368–375, 2016.

[17] CUSANO, C.; NAPOLETANO, P.; SCHETTINI, R. Combining multiple features for color texture classification. *Journal of Electronic Imaging*, International Society for Optics and Photonics, v. 25, n. 6, p. 061410, 2016.

[18] ROSENVOLD, K.; ANDERSEN, H. J. Factors of significance for pork quality-a review. *Meat Science*, v. 64, p. 219–237, 2003.

[19] SUN, X. et al. Prediction of pork color attributes using computer vision system. *Meat science*, Elsevier, v. 113, p. 62–64, 2016.

[20] LU, J. et al. Evaluation of pork color by using computer vision. *Meat Science*, v. 56, p. 57–60, 2000.

[21] RUSSO, V. et al. Indagine sul difetto di venatura delle cosce di suino destinate alla produzione del prosciutto di parma. *RIVISTA DI SUINICOLTURA*, Il Sole 24 Ore Edagricole: via Goito 13, I 40139 Bologna Italy: 011 39 051 ..., v. 44, p. 77–82, 2003.

[22] GARAVALDI, A.; ROSSI, A.; FIEGO, D. P. L. Difetti di presentazione della coscia per prosciutto crudo: valutazione sensoriale sul prodotto stagionato. In: FIRENZE UNIVERSITY PRESS. *II Convegno Nazionale della Società di Scienze Sensoriali*. [S.l.], 2008. v. 2, p. 303–308.

[23] ČANDEK-POTOKAR, M.; ŠKRLEP, M. Factors in pig production that impact the quality of dry-cured ham: a review. *Animal*, Cambridge University Press, v. 6, n. 2, p. 327–338, 2012.

[24] O'SULLIVAN, M. et al. Evaluation of pork colour: prediction of visual sensory quality of meat from instrumental and computer vision methods of colour analysis. *Meat science*, Elsevier, v. 65, n. 2, p. 909–918, 2003.

[25] SHIRANITA, K. et al. Grading meat quality by image processing. *Pattern Recognition*, Elsevier, (33), n. 1, p. 97–104, 2000.

[26] KOTSIANTIS, S. B. Supervised machine learning: A review of classification techniques. *An International Journal of Computing and Informatics*, v. 31, p. 249–268, 2007.

[27] MATEO, A. et al. Quality analysis of tuna meat using an automated color inspection system. *Aquacultural engineering*, Elsevier, v. 35, n. 1, p. 1–13, 2006.

[28] DU, C.-J.; SUN, D.-W. Automatic measurement of pores and porosity in pork ham and their correlations with processing time, water content and texture. *Meat Science*, Elsevier, v. 72, n. 2, p. 294–302, 2006.

[29] ULRICI, A. et al. Automated identification and visualization of food defects using rgb imaging: Application to the detection of red skin defect of raw hams. *Innovative Food Science & Emerging Technologies*, Elsevier, v. 16, p. 417–426, 2012.

[30] CHMIEL, M.; SŁOWIŃSKI, M. The use of computer vision system to detect pork defects. *LWT-Food Science and Technology*, Elsevier, v. 73, p. 473–480, 2016.

[31] SÁNCHEZ, A. et al. Control of ham salting by using image segmentation. *Food Control*, Elsevier, v. 19, n. 2, p. 135–142, 2008.

[32] COSTA, L. N. et al. Effect of the slaughterhouse on the incidence of defects in raw pig ham destined to the dry-curing process. *Veterinary research communications*, Springer, v. 32, n. 1, p. 351–353, 2008.

[33] BRADSKI, G.; KAEHLER, A. *Learning OpenCV: Computer vision with the OpenCV library*. [S.l.]: " O'Reilly Media, Inc.", 2008.

[34] SHAPIRO, L. *Computer vision and image processing*. [S.l.]: Academic Press, 1992.

[35] PONCE, J. et al. Computer vision: a modern approach. *Computer*, v. 16, n. 11, 2011.

[36] NIXON M.; AGUADO, A. Feature extraction and image processing. *Electronics & Electrical. Newnes*, 2002.

[37] GONZALEZ, R. C.; WOODS, R. E. *Processamento Digital De Imagens*. [S.l.]: ADDISON WESLEY BRA, 2010. ISBN 9788576054016.

[38] WEEKS, A. R. *Fundamentals of electronic image processing*. [S.l.]: SPIE Optical Engineering Press Bellingham, 1996.

[39] AZEVEDO, E.; CONCI, A.; VASCONCELOS, C. *Computação gráfica: Teoria e prática: geração de imagens*. [S.l.]: Elsevier Brasil, 2018. v. 1.

[40] MARSLAND, S. *Machine learning: an algorithmic perspective*. [S.l.]: Chapman and Hall/CRC, 2014.

[41] CYGANEK, B.; SIEBERT, J. P. *An introduction to 3D computer vision techniques and algorithms*. [S.l.]: John Wiley & Sons, 2011.

[42] SWOKOWSKI, E. W. *Calculus with analytic geometry*. [S.l.]: Taylor & Francis, 1979.

[43] KRIG, S. Image pre-processing. In: _____. *Computer Vision Metrics: Survey, Taxonomy, and Analysis*. Berkeley, CA: Apress, 2014. p. 39–83. ISBN 978-1-4302-5930-5.

[44] EGMONT-PETERSEN, M.; RIDDER, D. de; HANDELS, H. Image processing with neural networks—a review. *Pattern recognition*, Elsevier, v. 35, n. 10, p. 2279–2301, 2002.

[45] CHENG, H.-D. et al. Color image segmentation: advances and prospects. *Pattern recognition*, Elsevier, v. 34, n. 12, p. 2259–2281, 2001.

[46] SZELISKI, R. *Computer vision: algorithms and applications*. [S.l.]: Springer Science & Business Media, 2010.

[47] FOCA, G. et al. Prediction of compositional and sensory characteristics using rgb digital images and multivariate calibration techniques. *Analytica chimica acta*, Elsevier, v. 706, n. 2, p. 238–245, 2011.

[48] ORLANDI, G. et al. Automated quantification of defective maize kernels by means of multivariate image analysis. *Food Control*, Elsevier, v. 85, p. 259–268, 2018.

[49] ORLANDI, G. et al. Electronic eye for the prediction of parameters related to grape ripening. *Talanta*, Elsevier, v. 186, p. 381–388, 2018.

[50] TKALCIC, M.; TASIC, J. F. *Colour spaces: perceptual, historical and applicational background*. [S.l.]: IEEE, 2003. v. 1.

[51] BOLTZMANN, L. The second law of thermodynamics. In: *Theoretical physics and philosophical problems*. [S.l.]: Springer, 1974. p. 13–32.

[52] LYRA, M.; TSALLIS, C. Nonextensivity and multifractality in low-dimensional dissipative systems. *Physical review letters*, APS, v. 80, n. 1, p. 53, 1998.

[53] PUN, T. A new method for grey-level picture thresholding using the entropy of the histogram. *Signal processing*, Elsevier, v. 2, n. 3, p. 223–237, 1980.

[54] HARALICK, R. M.; SHANMUGAM, K.; DINSTEIN, I. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics SMC-3*, (6), p. 610–621, 1973.

[55] WESZKA, J. S.; DYER, C. R.; ROSENFELD, A. A comparative study of texture measures for terrain classification. *IEEE transactions on Systems, Man, and Cybernetics*, IEEE, n. 4, p. 269–285, 1976.

[56] OJALA, T.; PIETIKÄINEN, M.; MÄENPÄÄ, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 24, n. 7, p. 971–987, 2002.

[57] BRIGHAM, E. O.; MORROW, R. The fast fourier transform. *IEEE spectrum*, IEEE, v. 4, n. 12, p. 63–70, 1967.

[58] WATT, A.; POLICARPO, F. *The computer image*. [S.l.: s.n.], 1998.

[59] JAHNE, B. Digital image processing: concepts, algorithms, and scientific applications. *Berlin, Heildelberg: Springer-Verlag*, v. 570, 1997.

[60] SOBEL, I. An isotropic 3x3 image gradient operator. *Presentation at Stanford A.I. Project 1968*, 02 2014.

[61] CANNY, J. F. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 8, p. 679–698, 1986.

[62] RÄTSCH, G. A brief introduction into machine learning. *Friedrich Miescher Laboratory of the Max Planck Society*, 2004.

[63] SUTTON, R. S. Learning to predict by the methods of temporal differences. *Mach. Learning*, (3), p. 9–44, 1988.

[64] MITCHELL, T. *Machine Learning*. [S.l.]: McGraw-Hill Education, 1997. (McGraw-Hill international editions - computer science series). ISBN 9780070428072.

[65] RUSSELL, S.; NORVIG, P.; INTELLIGENCE, A. A modern approach. *Artificial Intelligence. Prentice-Hall, Egnlewood Cliffs*, Citeseer, v. 25, n. 27, p. 79–80, 1995.

[66] BISHOP, C. M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN 0387310738.

[67] Preface. In: THEODORIDIS, S.; KOUTROUMBAS, K. (Ed.). *Pattern Recognition (Fourth Edition)*. Fourth edition. Boston: Academic Press, 2009. p. xv – xvii. ISBN 978-1-59749-272-0.

[68] LIU, H.; MOTODA, H. *Feature selection for knowledge discovery and data mining*. [S.l.]: Springer Science & Business Media, 2012. v. 454.

[69] DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern classification*. [S.l.]: John Wiley & Sons, 2012.

[70] VAPNIK, V. N. *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag New York, Inc., 1995. ISBN 0-387-94559-8.

[71] BREIMAN, L. Random forests. *Machine learning*, Springer, (45), n. 1, p. 5–32, 2001.

[72] DEKA, P. C. et al. Support vector machine applications in the field of hydrology: a review. *Applied soft computing*, Elsevier, v. 19, p. 372–386, 2014.

[73] SUN, D. *Computer Vision Technology for Food Quality Evaluation*. [S.l.]: Elsevier Science, 2016. ISBN 9780128025994.

[74] KWONG, J. N. S.; GONG, S. Learning support vector machines for a multi-view face model. In: CITESEER. *BMVC*. [S.l.], 1999. p. 1–10.

[75] DEKA, P. C. et al. Support vector machine applications in the field of hydrology: a review. *Applied soft computing*, Elsevier, v. 19, p. 372–386, 2014.

[76] BOYD, S.; VANDENBERGHE, L. *Convex optimization*. [S.l.]: Cambridge university press, 2004.

[77] BREIMAN, L. Random Forests. *Machine learning*, v. 45.1, p. 5–32, 2001. ISSN 1098-6596.

[78] BARANAUSKAS, J. *Extração Automática de Conhecimento Utilizando Múltiplos Indutores*. [S.l.], 2001.

[79] BREIMAN, L. et al. *Classification and Regression Trees*. [S.l.]: Taylor & Francis, 1984. (The Wadsworth and Brooks-Cole statistics-probability series). ISBN 9780412048418.

[80] CARVALHO, A. et al. Inteligência artificial–uma abordagem de aprendizado de máquina. *Rio de Janeiro: LTC*, 2011.

[81] HAPFELMEIER, A.; ULM, K. Variable selection by random forests using data with missing values. *Computational Statistics & Data Analysis*, Elsevier, v. 80, p. 129–139, 2014.

[82] GENUER, R.; POGGI, J.-M.; TULEAU-MALOT, C. Variable selection using random forests. *Pattern Recognition Letters*, Elsevier, v. 31, n. 14, p. 2225–2236, 2010.

[83] RUSSO, V. et al. Study of some factors affecting the incidence of veining defect in thighs destined to dry-cured process of parma ham. In: CITESEER. *Proceedings of the 5th International Symposium on the Mediterranean Pig, Tarbes, France*. [S.l.], 2004. p. 16–19.

[84] BOTTACINI, M. et al. Skin lesion monitoring at slaughter on heavy pigs (170 kg): Welfare indicators and ham defects. *PloS one*, Public Library of Science, v. 13, n. 11, p. e0207115, 2018.

[85] PARMA QUALITY INSTITUTE. Dossier 2016. *Langhirano, Italy*, p. 5–38, 2017.

[86] FIEGO, D. P. L. et al. Effect of different stunning methods of pigs on subcutaneous veining defect and meat quality of raw ham. *Italian Journal of Animal Science*, Taylor & Francis, v. 2, n. sup1, p. 370–372, 2003.

[87] FELIN, E. et al. Current food chain information provides insufficient information for modern meat inspection of pigs. *Preventive veterinary medicine*, Elsevier, v. 127, p. 113–120, 2016.

[88] BORRÀS, E. et al. Data fusion methodologies for food and beverage authentication and quality assessment–a review. *Analytica Chimica Acta*, Elsevier, v. 891, p. 1–14, 2015.

[89] SINGH, A.; SINGH, M. L. Performance evaluation of various classifiers for color prediction of rice paddy plant leaf. *Journal of Electronic Imaging*, International Society for Optics and Photonics, v. 25, n. 6, p. 061403, 2016.

[90] CABALLERO, D. et al. Non-destructive analysis of sensory traits of dry-cured loins by mri–computer vision techniques and data mining. *Journal of the Science of Food and Agriculture*, Wiley Online Library, v. 97, n. 9, p. 2942–2952, 2017.

[91] CABALLERO, D. et al. Prediction of pork quality parameters by applying fractals and data mining on mri. *Food Research International*, Elsevier, v. 99, p. 739–747, 2017.

[92] ZAPOTOCZNY, P.; SZCZYPIŃSKI, P. M.; DASZKIEWICZ, T. Evaluation of the quality of cold meats by computer-assisted image analysis. *LWT-Food Science and Technology*, Elsevier, v. 67, p. 37–49, 2016.

[93] XU, J.-L.; RICCIOLI, C.; SUN, D.-W. Comparison of hyperspectral imaging and computer vision for automatic differentiation of organically and conventionally farmed salmon. *Journal of Food Engineering*, Elsevier, v. 196, p. 170–182, 2017.

[94] JR, S. B. et al. Comparison of svm and reptree for classification of poultry quality. *Proceedings of the Modelling, Simulation and Identification/ Intelligent Systems and Control–2016 (MSI 2016)*, p. 840–039, 2016.

[95] MUÑOZ, I. et al. Computer image analysis as a tool for classifying marbling: A case study in dry-cured ham. *Journal of Food Engineering*, Elsevier, (166), p. 148–155, 2015.

[96] ROPODI, A. I.; PANAGOU, E. Z.; NYCHAS, G.-J. E. Multispectral imaging (msi): A promising method for the detection of minced beef adulteration with horsemeat. *Food Control*, Elsevier, v. 73, p. 57–63, 2017.

[97] PREVOLNIK, M. et al. Classification of dry-cured hams according to the maturation time using near infrared spectra and artificial neural networks. *Meat science*, Elsevier, (96), n. 1, p. 14–20, 2014.

[98] BARBUT, S. et al. Progress in reducing the pale, soft and exudative (pse) problem in pork and poultry meat. *Meat science*, Elsevier, v. 79, n. 1, p. 46–63, 2008.

[99] GAJANA, C. et al. Effects of transportation time, distance, stocking density, temperature and lairage time on incidences of pale soft exudative (pse) and the physico-chemical characteristics of pork. *Meat science*, Elsevier, v. 95, n. 3, p. 520–525, 2013.

[100] FIEGO, D. L. et al. Preliminary investigation of the use of digital image analysis for raw ham evaluation. *Italian Journal of Animal Science*, Taylor & Francis, v. 6, n. sup1, p. 693–695, 2007.

[101] FERRARI, C.; FOCA, G.; ULRICI, A. Handling large datasets of hyperspectral images: Reducing data size without loss of useful information. *Analytica chimica acta*, Elsevier, v. 802, p. 29–39, 2013.

[102] CALVINI, R.; FOCA, G.; ULRICI, A. Data dimensionality reduction and data fusion for fast characterization of green coffee samples using hyperspectral sensors. *Analytical and bioanalytical chemistry*, Springer, v. 408, n. 26, p. 7351–7366, 2016.

[103] SOKOLOVA, M.; JAPKOWICZ, N.; SZPAKOWICZ, S. Beyond accuracy, f-score and roc: a family of discriminant measures for performance evaluation. In: SPRINGER. *Australasian joint conference on artificial intelligence*. [S.l.], 2006. p. 1015–1021.

[104] AGGARWAL, C. C. *Data classification: algorithms and applications*. [S.l.]: CRC Press, 2014.

[105] BROSNAN, T.; SUN, D.-W. Inspection and grading of agricultural and food products by computer vision systems—a review. *Computers and electronics in agriculture*, Elsevier, v. 36, n. 2-3, p. 193–213, 2002.

[106] SALAZAR, F. et al. An empirical comparison of machine learning techniques for dam behaviour modelling. *Structural Safety*, v. 56, p. 9–17, 2015.

[107] SCORNET, E. et al. Consistency of random forests. *The Annals of Statistics*, Institute of Mathematical Statistics, (43), n. 4, p. 1716–1741, 2015.

[108] WANG, L. *Support vector machines: theory and applications*. [S.l.]: Springer Science & Business Media, 2005. (177).

[109] LI, D. et al. Pornographic images recognition based on spatial pyramid partition and multi-instance ensemble learning. *Knowledge-Based Systems*, Elsevier, (84), p. 214–223, 2015.

[110] KAPUR, J.; SAHOO, P.; WONG, A. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing*, (29), n. 3, p. 273–285, 1985.

[111] SHEN, H. K.; CHEN, P. H.; CHANG, L. M. Automated steel bridge coating rust defect recognition method based on color and texture feature. *Automation in Construction*, v. 31, n. 0, p. 338–356, 2013.

[112] SOBEL, I. Neighborhood coding of binary images for fast contour following and general binary array processing. *Computer Graphics and Image Processing*, v. 8, p. 127–135, 1978.

WORKS PUBLISHED BY THE AUTHOR

1. ⋆ J. F. Lopes, L. Ludwig, D. F. Barbin, M. V. E. Grossmann, S. Barbon, “Computer Vision Classification of Barley Flour Based on Spatial Pyramid Partition Ensemble”, Sensors, 2019. (Qualis CC 2016, A1)

2. S. Barbon, S. M. Mastelini, A. P. A. Barbon, D. F. Barbin, R. Calvani, J. F. Lopes, A. Ulrici, “Multi-target Prediction of wheat flour quality parameters with near infrared spectroscopy”, Information Processing in Agriculture, 2019. (Impact Factor 3.110)

3. V. G. T. da Costa, E. J. Santana, J. F. Lopes and S. Barbon, “Evaluating the Four-Way Performance Trade-Off for Stream Classification”, Proceedings of the 14th International Conference on Green, Pervasive and Cloud Computing (GPC), Uberlândia, Minas Gerais, 2019. (Qualis CC 2016, B2)