unesp

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
CAMPUS DE GUARATINGUETÁ

FÁBIO HENRIQUE ANTUNES VIEIRA

IMAGE PROCESSING THROUGH MACHINE LEARNING FOR WOOD QUALITY CLASSIFICATION

Guaratinguetá
2016

FÁBIO HENRIQUE ANTUNES VIEIRA

PROCESSAMENTO DE IMAGENS ATRAVÉS DE APRENDIZADO DE MÁQUINAS PARA A CLASSIFICAÇÃO DA QUALIDADE DA MADEIRA

Tese apresentada à Faculdade de Engenharia do Campus de Guaratinguetá, Universidade Estadual Paulista, para a obtenção do título de Doutor em Engenharia Mecânica na área de Materiais.

Orientador: Prof. Dr. Manoel Cléber de Sampaio Alves

Guaratinguetá
2016

V657i

Vieira, Fábio Henrique Antunes
Image processing through machine learning for wood quality classification / Fábio Henrique Antunes Vieira – Guaratinguetá, 2016
119 f .: il.

Bibliografia: f. 94 - 105

Tese (doutorado) – Universidade Estadual Paulista, Faculdade de Engenharia de Guaratinguetá, 2016.

Orientador: Prof. Dr. Manoel Cléber de Sampaio Alves

1. Inteligência artificial. 2. Madeira Serrada. Título

CDU 007.52(043)

unesp **UNIVERSIDADE ESTADUAL PAULISTA**
CAMPUS DE GUARATINGUETÁ

FÁBIO HENRIQUE ANTUNES VIEIRA

ESTA TESE FOI JULGADA ADEQUADA PARA A OBTENÇÃO DO TÍTULO DE
"DOUTOR EM ENGENHARIA MECÂNICA"

PROGRAMA: ENGENHARIA MECÂNICA
ÁREA: MATERIAIS

APROVADA EM SUA FORMA FINAL PELO PROGRAMA DE PÓS-GRADUAÇÃO

Prof. Dr. José Antonio Perrella Balestieri
Coordenador

BANCA EXAMINADORA:

Prof. Dr. MANOEL CLÉBER DE SAMPAIO ALVES
Orientador / UNESP/FEG

Prof. Dr. FÁBIO MINORU YAMAJI
UFSCAR/Sorocaba

Profª. Drª. ANA LÚCIA PIEDADE SODERO MARTINS PINCELLI
FATEC Capão Bonito

Prof. Dr. ANDRÉ LUIS DEBIASO ROSSI
UNESP/Campus de Itapeva

Prof. Dr. CARLOS DE OLIVEIRA AFFONSO
UNESP/Campus de Itapeva

Junho de 2016

ACADEMIC RESUMÉ

FÁBIO HENRIQUE ANTUNES VIEIRA

DATE OF BIRTH	07/31/1970, Sorocaba/SP (Brazil)
FILIATION	José Antonio Vieira Nereide Antunes Vieira
1989-1992	Undergraduate Degree Language College – Major in English Metodista College in Itapeva/São Paulo
1998-1999	Graduate School Specialization in Education Psychology UNINTER in Curitiba/Paraná
2006-2006	Graduate School Specialization in Industrial Management Universidade Tecnológica Federal do Paraná
2007-2008	Master’s Degree Production Engineering Universidade Tecnológica Federal do Paraná
2011-2016	Doctorate Mechanical Engineering Universidade Estadual Júlio de Mesquita Filho Guaratinguetá – Major in Materials

DEDICATION

I would like to dedicate this research to very special people. Throughout my history in education, you teach me how to read, write, learn, study, respect, behave, work in team, and have the correct attitude towards knowledge and participation. Without you, this current accomplishment would not be possible. Thank you to all my teachers, coaches and professors, from the kindergarten to doctorate.

Also I dedicate this achievement to my parents, José Antonio Vieira and Nereide Antunes Vieira, who have consistently supported my initiatives and started, even at an early age, my interest towards knowledge. Dad showed me, through his examples and vast knowledge, the importance of reading; and Mom always with dedication, love and patience in her teachings. I will be forever grateful to you.

ACKNOWLEDGEMENTS

First, I would like to grant special thanks to my wife Marina Fogaça Rodrigues Vieira for loving and supporting me unconditionally, and understanding the importance of this accomplishment.

Not less valuable is the love and active believe in my capacity from my parents and sisters Flavia Helena and Fabiana. Also, I am thankful for my nephew Bruno and niece Laura for the inspiration that always makes me move forward.

To Márcio Teixeira Rodrigues and family, for his continuous interest and concern for education. Moreover, the companionship of Xará, Tati, Rafa, Luísa, and Maria Elizeti.

This thesis would not have been possible if it was not for my advising professor Dr. Manoel Cléber de Sampaio Alves. He always supported and trusted me on my initiatives during the research process.

I would like to thank Eng. Luiz Antonio Sguario Neto who provided an open environment for studying and learning within his company. In addition, I extend the gratitude to the team of Sguario and USP ICMC – UNESP, whose partnerships were essential to the development of this research.

I also want to thank the professors of UNESP and the post-graduation office staff, in special Dr. Carlos de Oliveira Affonso, who was always there to assist me on focusing on my work, contributing intensively to this achievement; and Dr. André Luis Debiaso Rossi for the cooperation and support. My thanks for both of you for the positive friendship.

In conclusion, my special thanks to friends and colleagues who traveled along this journey.

I'm grateful for all of you.

"The measure of intelligence is the ability to change."

— *Albert Einstein*

VIEIRA, F. H. A. Processamento de Imagens Através de Aprendizado de Máquinas para a Classificação da Qualidade da Madeira. 2016. 119f. Tese (Doutorado em Engenharia Mecânica) – Faculdade de Engenharia do Campus de Guaratinguetá, Universidade Estadual Paulista (UNESP), Guaratinguetá, 2016.

RESUMO

A classificação da qualidade da madeira é indicada para indústria de processamento e produção desse material. Essas empresas têm investido em soluções para agregar valor à matéria-prima, com o intuito de melhorar resultados, observando os rumos do mercado. O objetivo deste trabalho foi comparar Redes Neurais Convolutivas, um método de aprendizado profundo, na classificação da qualidade de madeira, com outras técnicas tradicionais de Máquinas de aprendizado, como Máquina de Vetores de Suporte, Árvores de Decisão, Regra dos Vizinhos Mais Próximos e Redes Neurais, em conjunto com Descritores de Textura. Isso foi possível através da verificação do nível de acurácia das experiências com diferentes técnicas, como Aprendizado Profundo e Descritores de Textura no processamento de imagens destes objetos. Foi utilizada uma câmera convencional para capturar as 374 amostras de imagem adotadas no experimento, e a base de dados está disponível para consulta. O processamento das imagens passou por algumas fases, após terem sido obtidas, como pré-processamento, segmentação, análise de recursos e classificação. Os métodos de classificação se deram através de Aprendizado Profundo e por meio de técnicas de Aprendizado de Máquinas tradicionais como Máquina de Vetores de Suporte, Árvores de Decisão, Regra dos Vizinhos Mais Próximos e Redes Neurais juntamente com os Descritores de Textura. Os resultados empíricos para o conjunto de dados das imagens da madeira serrada mostraram que o método com Descritores de Textura, independentemente da estratégia empregada, foi muito competitivo quando comparado com as Redes Neurais Convolutivas para todos os experimentos realizados, e até mesmo superou-as para esta aplicação.

Palavras-chave: Inteligência Artificial, Descritores de Textura, Madeira Serrada.

VIEIRA, F. H. A. Image Processing Through Machine Learning for Wood Quality Classification**.** 2016. 119p. Thesis (Doctorate in Mechanical Engineering) – Faculdade de Engenharia do Campus de Guaratinguetá, Universidade Estadual Paulista (UNESP), Guaratinguetá, 2016.

ABSTRACT

The quality classification of wood is prescribed throughout the wood chain industry, particularly those from the processing and manufacturing fields. Those organizations have invested energy and time trying to increase value of basic items, with the purpose of accomplishing better results, in agreement to the market. The objective of this work was to compare Convolutional Neural Network, a deep learning method, for wood quality classification to other traditional Machine Learning techniques, namely Support Vector Machine (SVM), Decision Trees (DT), K-Nearest Neighbors (KNN), and Neural Networks (NN) associated with Texture Descriptors. Some of the possible options were to assess the predictive performance through the experiments with different techniques, Deep Learning and Texture Descriptors, for processing images of this material type. A camera was used to capture the 374 image samples adopted on the experiment, and their database is available for consultation. The images had some stages of processing after they have been acquired, as pre-processing, segmentation, feature analysis, and classification. The classification methods occurred through Deep Learning, more specifically Convolutional Neural Networks - CNN, and using Texture Descriptors with Support Vector Machine, Decision Trees, K-nearest Neighbors and Neural Network. Empirical results for the image dataset showed that the approach using texture descriptor method, regardless of the strategy employed, is very competitive when compared with CNN for all performed experiments, and even overcome it for this application.

Keywords: Artificial Intelligence, Texture Descriptors, Sawn Wood.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABNT	Brazilian Association of Technical Rules (Associação Brasileira de Normas Técnicas)
AI	Artificial Intelligence
ANN	Artificial Neural Networks
CNN	Convolutional Neural Networks
DBN	Deep Belief Network
DL	Deep Learning
DESTIN	Deep Spatiotemporal Inference Network
DT	Decision Trees
FD	Feature Detecting
GPS	General Problem Solver
IBÁ	Brazilian Tree Industry (Indústria Brasileira da Árvore)
KNN	K-Nearest Neighbors
ML	Machine Learning
MLP	Multilayer Perceptron
NBR	Registered Brazilian Standard (Norma Brasileira Registrada)
NDE	Non-destructive evaluation
NDT	Non-destructive test
NN	Neural Networks
SGLD	Spatial Gray Level Dependence
SVM	Support Vector Machine
TD	Texture Descriptors
VLSI	Very Large Scale Integration

LIST OF SYMBOLS

φ	Activation function specified as sigmoid function
$\sum$	Sum
$\in$	Given element belongs to a certain set
$\mathbb{Z}$	Integer
λ	Probability density
k	Quantity of neighbors
$\mathbb{N}$	Natural numbers
η	Learning rate
k	Cohen’s kappa measure

TABLE OF CONTENTS

1. INTRODUCTION

The quality analyzes of wood are recommended all through the wood chain industry, especially those from the processing and production areas. They have spent time and effort in the attempt of adding value to primary products, with a specific end goal to achieve superior results conforming to the market expectations.

Because of this interest, it is important to improve a few deciding aspects in the productiveness. New advances have risen, particularly those identified with information technology, as an approach to grow the utilization of resources and to support the quality and productivity expansion.

As stated by Almeida (2014), in recent years, the machine vision systems have been advancing quickly, as a result of the development of high technology camera sensors and computing potential of the processors. In the meantime, the price of systems based on cameras has diminished, enabling a cost-efficient classification solution environment for wood quality.

Even so, the automated classification of wood quality is still in its outset. High investment is expected for the few available commercial applications within sawmills. As a rule, the wood quality for adapting the cost and quality precedent is assessed by a human specialist, who must intercede by altering the included parameters (JOHNSON; BODEDE; ADESINA, 2014).

As affirmed by Affonso, Sassi and Barreiros (2015) a normal research dilemma is the way the hand-made estimations and human assessments of wood quality can be supplied by a solid and affordable automated classification system, so that the performance of the wood processing is enhanced. Regularly, it is costly to determine what parameters assume definitive role in acquiring distinguished services or products

In this procedure, computational modeling is frequently used to choose the most suitable product approach to the application. In the quest for better arrangements, there are a few commercial softwares, including modeling through finite elements, among others. However, the software adopted for such action has normally restrictive value for micro and small sized organizations, becoming unfeasible (VIEIRA; AFFONSO; ALVES, 2014).

Considering the focused obstacles, the use of Deep Learning – DL via grayscale pixels or allied with Feature Extraction is suggested for forecasting of these methods, once it gives the essential mathematical tools to manage equivocation in learning acquisition methodology, while that can be executed without the utilization of commercial software kits.

In this setting, the adoption of Artificial Intelligence – AI, in the control and methodology development, has progressed at a quickened rate because of sufficient productivity for an expansive range of modern applications, as image processing. It legitimizes the developing enthusiasm for using deep learning as a part of complex industrial procedures (MISHRA, 2011).

Accordingly, imaging is considered in this thesis as the main technology for programmed wood quality classification, bringing satisfactory result accurateness.

1.1 JUSTIFICATION AND PROBLEM

The most important cost in the improvement of an assessment system is usually the expense of utilizing software designers to compose devoted programming, rather than the equipment itself, i.e. hardware. In any case, considering the significance of quality in commercial centers for finding and maintaining clients, the payback time for automated vision systems can worth the investment (PHAM; ALCOCK, 2003).

According to Zamri et al. (2015) a wood classification system can be created keeping in mind the end goal of providing reasonable support to wood industry. Besides, the employees with expertise in identification on wood defects are limited, concerning amount, accuracy and time.

Researches and advancements in the area of classification systems through images, will supply for the improvement of a high-elaborated, cost benefit equipment for wood sorting. Sawmills are the ones that might receive the greatest advances in automation, as they will have final value-added products and time saving (ALMEIDA, 2014).

As a matter of comparison, study of Huber, McMillin, and McKinney (1985) found that after checking and assessing six employees in a wood sawmill, with the task of sorting wood boards, the accuracy was 68%.

In the same sense, another research with four separated trials on human graders pointed to a performance of 55% accuracy with continuously shaped wood (POLZLEITNER; SCHWINGSHAKL, 1992).

In studies of Kline, Surak, and Araman (2003) with wood classification systems, they reached only 63% accurate, sorting each board individually. Excluding the low accuracy result, it is still higher than the 48% observed on visual graders.

The increasing complexity in the manufacture of wood-made products and the need for greater efficiency, better quality and lower cost have required a change in the wood industry throughout the time.

Hashim, Hashim and Muda (2015) declared that automated inspection system for defects in surfaces has demonstrated high importance in wood industry. Because of the contraction of the forest assets and the rising of wood cost, the choice for a mechanized structure is seen as an answer to enhance resources while preserving a trustworthy quality of the final product.

Besides the economic approach involved in the search for better cost-benefit materials, there is the environmental concern with interest in developing processes that minimize energy consumption (PHAM; ALCOCK, 2003).

Among the available computational tools for the improvement of production processes include the techniques of Artificial Intelligence, due to its generalization and versatility capacity.

As stated by Rosa et al. (2010) the current objective of the research in manufacturing is dedicated to the development of flexible and self-adjusting intelligent systems, aiming to reduce the presence of human operators, so that the control of these processes is done through computer systems, both software and hardware.

These systems are especially important considering the particular characteristics of the wood industry due to hygroscopic, heterogeneous and anisotropic features of the materials produced from it.

In this scenario, Artificial Intelligence has shown to be effective for treating this type of problem (AFFONSO; SASSI; BARREIROS, 2015).

The main disadvantage in the manufacture of high quality products that use wood as raw material is the lack of objectivity in the quality regulation of the final product and process, causing a demand for visual control (TOPALOVA; TZOKEV, 2011).

Image treatments have been used extensively in wood processing, especially in relation to the classification of species. In many applications, such as wooden panels and furniture, the classification of wood can provide relevant information on the appearance, properties and procedures for the preparation of the products. In most cases, trained operators perform the wood classification, but this solution suffers from considerable disadvantages. Thus, the literature presents a variety of studies in which artificial intelligence techniques are used in image processing for identification of wood defects (HE et al., 1997) and (PHAM et al., 2006).

As stated above, the problem of this work is to engineer data acquisition equipment, along with a program to interpret and treat them; generating reliable information in controlling and improving the production processes of the wood industry, via artificial intelligence techniques, more specifically, Machine Learning ones, such as KNN, NN, SVM, DT with texture descriptors, when compared to Deep Learning method, like Convolutional Neural Network - CNN.

Thus, as there are several procedures of Artificial Intelligence applied in distinct industries, there was the idea for applying a combination of those techniques within the wood industry of an emerging center, aiming to provide information to increase the production efficiency.

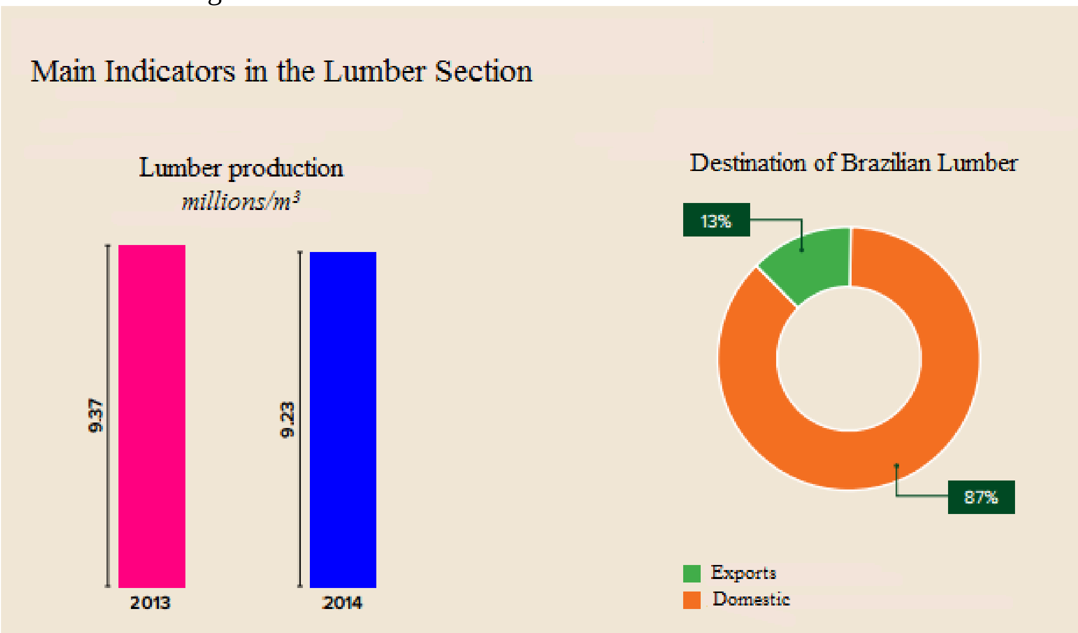
According to IBÁ (2014) the generation of hardboards from planted trees has expanded 5.7% in 2014, compared to the previous year, coming to 2.40 million cubic meters. This amount was certainly impacted by the exportation rate increase of 22.3% of these items, involving 1.32 million cubic meters in this period over 1.02 million cubic meters from the preceding period. In the meantime, hardboards sales within domestic market fell 9.4% relating to the 101.5 million cubic meters in 2013, reaching only 1.08 million cubic meters in 2014.

The deceleration in Brazilian civil construction area has driven to a comedown of 6.1% in domestic utilization of timber, from 9.37 million cubic meters in 2013 to 9.23 million cubic meters in 2014. In this scheme, the domestic result of the amount of timber from planted trees decreased 1.5%, as stated by Brazilian Tree Industry Association – (IBÁ, 2013).

The destination of this lumber has reached 87% within domestic market in 2014, and the exportation of the Brazilian lumber represented about 13% (Figure 1).

But the exportations also included other segments of wood industry, as pulp, paper (tissue, packaging, cardboard, printing and writing paper, special paper, newsprint), and wood panels.

Figure 1 – Main Indicators in the Lumber Section 2013-2014



Source: (IBÁ, 2015)

According to Sochirca, Afonso and Gil (2013) among the sawmills in this country, few of them are considered large companies. Most of them are known as small businesses, which present low budget, and consequently short amount of resources for investment on technologies. This issue is admitted as disabling, once it avoids the improvement of the results, due to ancient production equipment and costly processes.

Thus, it is expected that this work will contribute to the integration of the knowledge areas directed to the wood industry and computing.

That said, the question that this research attempts to answer is: What is the accurateness obtained by Convolutional Neural Network in sawn wood image classification, compared to traditional Machine Learning techniques associated with Texture Descriptors?

The hypothesis are that Texture Descriptors used with traditional techniques may overcome the accuracy classification of wood boards due to the well-behaved data of the objects under investigation, in comparison to Deep Learning, represented here by the Convolutional Neural Networks, because of, among other aspects, computational cost.

This thesis is constituted of the following parts, besides Section 1 with the Introduction, Justification and Problem:

Section 2 brings the General Objective and Specific Objectives.

Section 3 discusses Wood Processing, Wood Defects, Wood Industry Scenario, besides the classification of sawn wood.

Section 4 reviews Artificial Intelligence, from its history, definitions and approaches. Then it explains Artificial Intelligence applied to Image Processing, and Digital Image Processing.

Section 5 specifies the different techniques applied on image classification, such as Artificial Neural Network, Deep Neural Networks, Deep Learning, Convolutional Neural Networks, K-Nearest Neighbors, Support Vector Machines, and Decision Trees.

Section 6 is about Texture Descriptors and Statistical Texture Descriptors.

Section 7 is on charge of Materials and Methods used on the research. It describes the Hardware, Image Acquisition, Pre-processing for image enhancement, distinct Methods for Image Classification, Acquisition Equipment, Softwares, Methods, and Assessment Measures.

Section 8 is the Analysis and Discussion of the data from the research. It discusses the different approaches on classification, such as Convolutional Neural Networks, Neural Networks, K-Nearest Neighbors, Support Vector Machines, and Decision Trees, comparing their performance on the Overall Results.

Section 9 brings the Conclusions based on all information obtained during the study.

2. OBJECTIVES

This section informs the objectives of the research. As stated earlier, some possible options to verify the accuracy level are through experiments with different traditional Machine Learning techniques and Texture Descriptors for image preprocessing of wood boards.

2.1 GENERAL OBJECTIVE

Observing the opportunities for automated systems within wood process, the general objective of this thesis is to compare Convolutional Neural Network, a deep learning method, for wood quality classification to other traditional Machine Learning techniques, namely Support Vector Machine (SVM), Decision Trees (DT), K-Nearest Neighbors (KNN), and Neural Networks (NN). These traditional techniques are induced by a set of texture descriptors extracted from wood board images.

2.2 SPECIFIC OBJECTIVES

i. Assessing the accuracy of wood image classification within the dataset using a Deep Learning method, specifically Convolutional Neural Network;

ii. Verifying the accuracy of traditional Machine Learning techniques, such as K-Nearest Neighbors, Neural Networks; Support Vector Machine and Decision Trees associated with Texture Descriptors, in classifying wood image samples.

iii. Creating a dataset with wood images from a real-world situation, with their possible defects and labels according to an industry;

3. WOOD PROCESSING

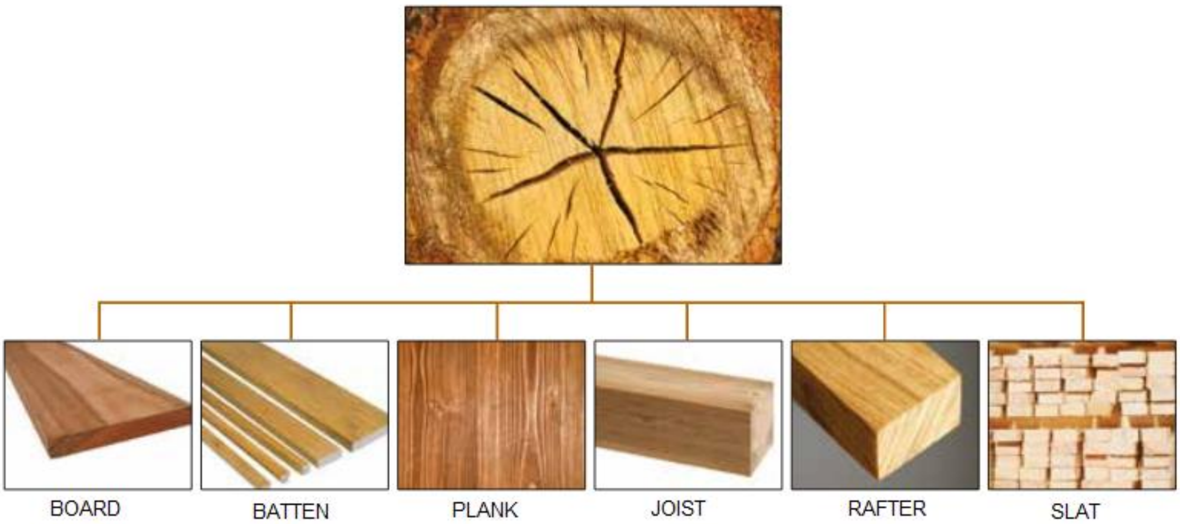
This section enlightens the importance of the wood sector for the country. Additionally, it brings the main products originated from wood and emphasizes sawn wood industry, with normal timber defects and its classification.

Tripoli and Prates (2015) affirm that although there is a significant part of the wood from illegal logging, the activities of the cultivated forests in Brazil have grown in recent years, making the activity an attractive alternative for the economy. In addition, among the wood-base products, raw wood had the worst comparative advantage index, with the lowest growth, indicating that the industry has preferred to negotiate higher value-added goods, such as sawn wood.

Wood represented approximately 5% of National GDP in 2013, including the production of lumber, panels, pulp and paper, and energy. The mechanically processed wood industry participated in the generation of approximately 0.4% of this share (OLIVEIRA, 2014).

The industrial segment of products obtained by processing of wood *in natura*, such as boards, rafters, joists, battens, and slats, are shown on the figure 2.

Figure 2 - Main Products from Sawn Wood Segment



Source: Poyry (2014)

Typically, these products are used in construction, transportation of goods and the production of furniture of all kinds, besides decorative components.

3.1 WOOD INDUSTRY SCENARIO

Traditionally, wood industry has used human expert to classify the quality of wood pieces. However, by only using experts' naked eyes for classification has led to several problems, such as inconsistence quality results, only expert workers who possess a great deal of experience are able to perform the job, and the process of grading the quality is very depending on experts' opinion, which may be subjective (ABDULLAH et al., 2007).

Besides that, the same authors also mentioned the process that depends on human expert usually have problems, especially in managing resources. Examples are the cases of retired workers or high employee turnover. When those examples happen, the company needs to invest on employee training and it takes time to master a new expert in classifying the quality of the wood.

Moreover, there are two other aspects to be considered. On the one hand, traditional methods need a lot of workers and time to make a process done. The other is the low accuracy level obtained by human graders, around 60%, as pointed out in section 1.1.

3.2 WOOD DEFECTS AND CLASSIFICATION OF SAWN WOOD

There are different wood defects induced either by internal or external forces on the material. While some of these imperfections exist because of outer effects and abnormalities of the growth conditions, some of them are due to incorrect storage (BUDAKCI; CINAR, 2004).

It is presented the most common defects found in wood in Table 1. They are caused by growth defects in the tree, improper curing of the rough-cut timbers, or improper treatment/weatherproofing of exposed wooden structures. These flaws are most common in whole, straight grain cuts.

Table 1 – Common wood defects

Bend	The wood curves along the length of an edge. This is rarer and almost never found in panels, only boards.
Bow	Similar to bonding, except instead of curving along of an edge, the piece curves along the length of its face. It is easiest to picture by imagining a long board lying on its face supported on either end but not in the center. If you jump on the center of the board, you are causing it to bow.
Check	The wood splits into a v-shaped channel near an end-grain cut, often following the curved lines of the growth rings.
Cup	The entire length of the board curves along its grain. This is most often seen on decks or patios that were left untreated in the rain. Each board has curved up along its length.
Decay	Disintegration of the wood due to fungi. Also called dote, rot, or unsound wood. It can result in soft spots and discoloration in the piece.
Shake	A lengthwise separation of the wood through the growth rings. This will appear very similar to checking. Basically, a crack that follows the grain lines of a piece.
Wane	Pieces of the timber chip or loosen near the edges of the cut, again following the curved growth rings. Here, two separate branches of the tree grew together over time trapping some layers of bark between them. When the timber is cut these pieces fall loose.
Warp	Similar to cupping, except rather than curving along the grain lines (or the edge of the board), the wood curves on a diagonal.

Source: Adapted from Johnson, Bobede and Adesina (2014).

Most defects on *Pinus* are related to the presence of knots on the boards, which can be initially classified as their consistency, as sound, dry, hole or decayed knots (OLIVEIRA et al., 2008).

Budakci and Cinar (2004) also mentioned knots, deformities related with anatomical structure of the pieces, as the most shocking in manufacturing processes, due to their negative effects on the product.

In spite of the fact that the National Standard Norms (ABNT) created rules related to wood defects as NBR-12297 and NBR-12551 that points out types of knots; and NBR-11700 that describes wood defects and defines their classification, the criteria used in this study to sort wood quality followed the industry classification.

The wood boards were basically classified into 3 categories; class A, which indicates no or very small amount of defects; class B, which suggests a number of defects, but in small size; and class C that accepts higher amount of defects (figures 3-6).

ABNT classification is explained further, in Annex A.

Fig. 3. Class A sample, free of defects

Source: Own Author

The concern of the experts on wood classification in the industry was separating boards, using three categories. As the image on figure 3, one type was the Class A, a piece with no defects. Authors usually refer to this type as clearwood (FUNCK et al. 2003; TODOROKI et al. 2010; PAHLBERG et al., 2015).

Figure 4 demonstrates a wood sample from class B, which presents a knot defect, but in modest size. Visual grader specialists usually classify this kind of problem as a minor defect.

Fig. 4 – Class B sample, defect smaller than 50mm

Source: Own Author

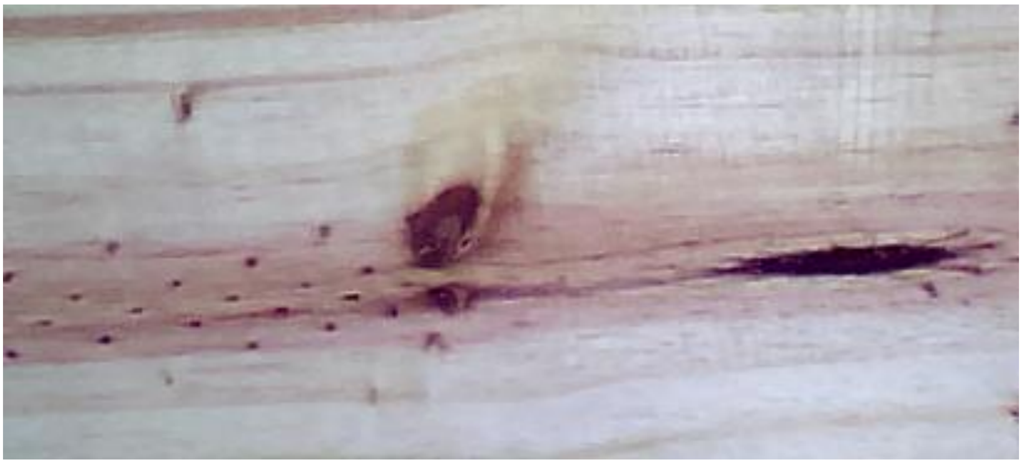
When assessing the quality of the wood, the presence of knot defects influence it negatively. This can be easily confirmed once the wood property would have its strength reduced. A small size knot is already a variance, but group of knots (fig. 5) causes more damage in the piece resistance (HASHIM; HASHIM; MUDA, 2015).

Fig. 5 - Class C sample, group of knots or knots larger than 50mm

Source: Own Author

Figure 6 shows an example of exposed pith, which is not considered a wood defect, as it belongs to the tree. The real-world problem is that the boards with the apparent presence of the pith have to be separated from the ones that are not visible, because during the drying process, they tend to suffer more intensively.

Fig 6. - Class C sample, presence of exposed pith



Source: Own Author

One of the most considerable impediments for viable use of wood is distortion, either for civil construction or furniture production. The losses with this case can become a real issue. Bow and spring are regularly created by pieces that have suffered from any kind of pressure, as compressed wood. On the other hand, spiral grain are responsible for the formation of twists. In order to avoid the absence of straightness, some care must be mandatory, and sawmill employees should separate sawn wood with exposed pith from the others, once they are subject to severe twist after drying (NYSTRÖM, 2002).

In the industry, boards with exposed pith are separated from the others, and then proceed to the next sector, where they will be dried at slow-paced procedure.

4. ARTIFICIAL INTELLIGENCE

In this Section, it will be discussed the Artificial Intelligence - AI, since its history, passing through the chronological advance of its development, its definition, and its approach. Then, it brings AI applied to image and digital image processing, besides the technologies used for imaging.

4.1 HISTORY OF ARTIFICIAL INTELLIGENCE

The initiation of Artificial Intelligence is related to Warren McCulloch and Walter Pitts (1943), when they carried out the primary work that was perceived as so. They drafted on three sources: information of the essential physiology and capacity of neurons within the brain; a formal investigation of propositional sense because of Russell and Whitehead; and Alan Turing's hypothesis of computation. They proposed an artificial neurons model in which every neuron is described as being "on" or "off," with a change to "on" happening in light of incitement by a sufficient number of neighboring neurons. The condition of a neuron was considered as "truthfully comparable to a proposition which introduced its satisfactory stimulus." They demonstrated, for instance, that any computable function could be registered by some network of connected neurons, and that all the consistent connectives, as “and”, “or”, “not”, and so forth, could be executed by basic net structures. McCulloch and Pitts likewise recommended that suitably characterized networks could learn. Donald Hebb (1949) showed a straightforward upgrading rule for adjusting the connection strengths among neurons. His guideline, known as HEBBIAN learning, remains a current persuasive model.

As mentioned by Russell and Norvig (2014), some scientists at that time were doubtful about whether this sort of work ought to be considered science. However, neural networks soon or later would be presented as a field of knowledge. This was the belief of other enthusiastic researchers. Minsky was later to demonstrate persuasive hypotheses showing the constraints of neural network research.

There were various early samples of work that can be portrayed as Artificial Intelligence. Nonetheless, Alan Turing's vision was maybe the most powerful, when searching for solutions to break the Germans’ codes. After the World War II, he pondered the idea of building a computer that was able to think. His paper published in 1950, Computing Machinery & Intelligence, was one of the pioneers to be written describing the theme. In that, he presented the Turing Test, machine learning, genetic algorithms, and support learning. He proposed the Child Program thought, disclosing the idea that as opposed to attempting to create a project to simulate adult minds, he focused on producing one which simulated the child's mind (COPPIN, 2012).

The conception of Artificial Intelligence was Princeton, home of John McCarthy, another powerful figure in AI. When he gets his PhD in 1951 and laboring for a long time as a teacher, McCarthy moved to Stanford and after that to Dartmouth School, which was to turn into the official origin place of the field. McCarthy persuaded Minsky, Claude Shannon, and Nathaniel Rochester to help him unite U.S. specialists involved on automata hypothesis, neural nets, and the investigation of intelligence. They sorted out a two-month workshop at Dartmouth in 1956 (COPPIN, 2004).

The early years of AI were packed with successes in a limited manner. Given the undeveloped computer and programming apparatuses of the time and the truth that just a couple of years prior computers were seen as things that could do math and nothing else, it was impressive whenever a computer did anything remotely intelligent. The scholarly foundation, overall, liked to accept that "a machine can never do X", referring to Turing's list of X's. AI analysts regularly reacted by showing one X at a time (RUSSELL; NORVIG, 2014).

The same authors also observed that Newell and Simon's initial achievement was caught up with the General Problem Solver - GPS. The project was planned from the beginning to imitate human problem-solving protocols. In this manner, GPS was likely the first program to typify the reasoning humanly approach. The accomplishment of GPS and resulting projects as models of cognition drove Newell and Simon (1976) to detail the renowned physical symbol system premise, which expresses that a physical symbol system has the essential and sufficient means for general clever response. They implied that any system (human or machine) demonstrating intelligence must work by controlling information structures made out of symbols. The physical symbol system has the necessary and sufficient means to make a general intelligent action.

The computer scientist Nathaniel Rochester and his associates delivered a portion of the first AI programs at IBM, in which he was the supervisor. Herbert Gelernter (1959) while working in Fortran List Processing Language built the Geometry Theorem Prover, which had the capacity to demonstrate hypotheses that numerous math students would find truly complex. Beginning in 1952, Arthur Samuel created a progression of projects for checkers that, in the end, figured out how to play at a solid beginner level. Along the way, he invalidated the thought that computers can do on1y what they are programmed to: his system immediately learned how to play a superior game than its inventor.

Then McCarthy, a cognitive scientist, created, in the late 1950s, the LISP language that has become one of the most dominant programming languages in AI, which is until used in Artificial Intelligence researches (MCCARTHY, 1959).

Early work expanding on the neural networks of McCulloch and Pitts likewise prospered. The work of Winograd and Cowan, in 1963, demonstrated how an expansive number of components could aggregately represent an individual idea, with a relating increment in vigor and parallelism. The learning strategies initiated earlier were upgraded by Bernie Widrow and Marcian Hoff, who called his networks “adalines”, and by Frank Rosenblatt with his perceptrons. The perceptron convergence theorem by Block, Rosenblatt and Novikoff states that the learning algorithm can conform the association qualities of a perceptron to match any input information (MISHRA, 2011).

According to Coppin (2012) from the 1960s, the focus of artificial intelligence study was not to create an intelligent robot, similar to a human, but use algorithms, heuristics and methodologies based on the way the human brain solves problems.

In 1966, a trouble was the confrontation of large portions of the issues that AI was pursuing to solve. The greater part of the early AI projects tackled problems by going for diverse series of steps until the arrangement was found. This methodology worked at first because microworlds, which are a kind of learning environment, did not contain many objects and consequently could achieve few activities and short solution settings. Prior to the hypothesis of computational complexity was created, it was generally believed that scaling up to bigger issues was just a problem of faster equipment and bigger memories (COPPIN, 2012).

Then, the emphasis on the hardware was switched to the software structure. The reason for that was the hardware evolves more slowly and the focus, until then, was directed most of the time to it. Consequently, the software received less attention. Another reason for the change was that the hardware had reached a point where the evolution did not progress as necessary. Then researchers have alternatively to investigate improvements in software, and thus improve performance and functionality (WALCOTT-JUSTICE, 2014).

From the late 1990s and early 2000s, distinct concepts of AI, as theories and algorithms, in and with online technology, were highlighted, for data retrieval, extraction, classification, summarization of documents, or multi-languages machine translation. Common rules for potential conflict by pair-wise mutual exclusion, web and business intelligence, libraries for e-commerce, and heuristics search methods have emerged (MISHRA, 2011).

Russell and Norvig (2014) announced that the hopefulness that accompanied the improvement of resolution theorem proving, for instance, was soon hosed when specialists neglected to demonstrate hypotheses including more than a couple of dozen truths. The fact that a program can discover an answer on a basic level does not imply that it contains any of the instruments expected to discover this solution in practice.

As observed, since the 1940s, the investigation of Artificial Intelligence has been flavored with significantly more authenticity.

4.2 DEFINITIONS OF AI

As seen on AI history above, every period had its advances and its applicability. New discoveries have been developed as the opportunities have increased. Likewise, authors' definitions of Artificial Intelligence vary accordingly to the time it was created.

At first, the concern of AI was to make the machine think humanly and was concerned with thought processes and reasoning. Besides, it measures the success in terms of human performance.

In this sense, Bellman (1978) delineated it as the mechanization of activities connected to human understanding; and practices like learning, logical decision, and managing.

Throughout 1980s, experts defined AI as the capability of making a computer think literally and rationally, highlighting the intelligence itself and mental faculties (CHARNIAK; McDERMOTT, 1985; HAUGELAND, 1985; MOIGNE, 1988).

Haddock and O'keefe (1990) concluded that AI is just an empowering innovation that will assist on better modelling and examine manufacturing systems, in order to put the method for investigation under the control of those with the issues.

As observed, in less than a decade, the definitions of AI looped from imagining to modelling to acting, as it can be seen on the opinions of Kurzweil (1990); Rich and Knight

(1990); and Schalkoff (1990), who characterized Artificial Intelligence as the ability to reason and act humanly, and perform activities in the correct way.

Shortly, authors started to compare AI to human’s behavior, describing it like the automation of intelligence, and intelligence behavior emulator (DE SILVA, 1994; LUGER, 1995; LIBERATORE, 1997).

Another definition informs that Artificial Intelligence is the study of systems that perform actions as they had intelligence (COPPIN, 2004).

Mishra (2011) describes AI as the method of implementing the characteristics of human intelligence in machines. In the opinion of Coppin (2004, p. 4), a more complex explanation for the term is: “Artificial Intelligence involves using methods based on the intelligent behavior of humans and other animals to solve complex problems”.

Ali et al. (2015) inform that Artificial Intelligence is defined as the capacity of computers or different machines in performing distinct activities that involve human intelligence. Nowadays, it has been extensively used as a part of different applications like medicinal, entertainment games, automation, and process control to suppose unmeasured states, such as quality, temperature, weight, impurity, concentration, among others.

As declared by Chou et al. (2015), most aspects built on AI have concentrated on creating algorithms for enhancing modeling accuracy or the velocity of its construction. Accordingly, those AI-based models are conceivable mechanisms for solving engineering, medical, and management issues.

A more recent definition of AI by Kow et al. (2016) is the one that mentions it is as a kind of human made intelligence based on hardware and softwares, normally used to diminish one’s working burden.

4.3 APPROACHES OF AI

Many AI applications are emerging and they tend to grow more and more. However, regardless of the type of application, the bases of a project are divided into two types: cognitive and connectionist.

For Giere (1992) and Fernandes, Ramos and Rosa (2005), the Cognitive approach emphasizes the cognitive processes that are, in other words, the way the human reasons; and in the Connectionist approach, the emphasis is on working model of the brain, the neurons and neural connections.

Artificial intelligent experts are involved on two of the major issues: learning representation and search. Supposing the learning representation, the focus is on knowledge acquisition, by alluding to capacities of the formal dialect utilized as a part of the procedure for computer manipulation. In the event of search, the emphasis is on such a strategy to solve the issue, which permits to see problem states (LUGER, 2004).

Concerning the primary problems of AI extension and its application fields, it is possible to find the structure of images and operations through which the intelligent arrangement of the issue is performed; and the quest for key rules for those arrangements, which are brought on by specific structures of symbols and operations. The problems of learning representation and search are the fundamental assortment of examination in the issue field of AI (KORNIENKO et al., 2015).

As informed by Russel and Norvig (2014), the processing of learning clusters is the predominant representation of intelligence technologies. The systems generally called intelligent or expert are those that the prevailing basis is a knowledge ground; close to be confused with the natural.

Confronted with the issues of high multifaceted nature, as elucidation of normal dialect, support with problematic decision-making, and understanding of visual message, humans often find assistance in expert/intelligent systems (KORNIENKO et al., 2015).

The same authors also mentioned that those are essential when utilizing low-organized expert knowledge, realistic experience that is sufficiently expansive, and most of the time extremely valuable. Intelligent systems are extremely compelling in the fields where great amount of experience from high-qualified expert professionals is aggregated in numerous branches of knowledge.

For Fernandes, Ramos and Rosa (2005), the objective of AI approaches is to find an explanation for intelligent behavior based on the psychological characteristics algorithmic processes.

Kornienko et al. (2015) also mention that artificial intelligence approaches, or intelligent systems, are separated into expert systems that explain distinctive classes of issues. Among the issues of examination, there are data analysis, diagnostics, and decision-making assist. As for the issues of synthesis, there are control, engineering, and design. Inside of the scope of problems clarified by intelligent systems, associated issues are emphasized, i.e. checking, assessment, and learning.

By the way, whenever the word learning is mentioned, it is common to remember Machine Learning – ML, which is often confused with AI. Some authors claim they are the same thing. Others say that ML is part of AI.

Machine Learning - ML is part of Artificial Intelligence that assists computers with the capacity to learn without being clearly programmed, i.e., lacking human guidance. Machine Learning concentrates on the advancement of computer programs, which are able to instruct themselves to develop and change when exhibited to new information (MEYER et al., 2016).

Artificial intelligence algorithms have turned out to be a promising mechanism for dealing with a few extreme investigative issues. Machine learning, as an expansive subarea of AI, is worried with calculations and systems that permit computers to learn. The machine learning covers principle spaces, for instance, information mining and programming applications (ALAVI; GANDOMI; LARY, 2016).

As described by Franco and Bacardit (2016), the procedure of machine learning is like that of data mining. They both scan through information to search for standard examples. Nonetheless, rather than drawing information out for human understanding, machine learning utilizes that information to enhance the system's own comprehension. Machine learning programs distinguish designs in data and change program activities appropriately.

According to the definition of Ertugrul (2016), ML makes procedures by applying the information necessary for a machine to learn and adjust itself when presented to new data. It is considered as a machine trainer. It relies upon other two techniques by perusing mined information, making an algorithm via Artificial Intelligence, and after that, updating present algorithm appropriately to assimilate another assignment.

The author continues saying that ML is equipped for summing up data from extensive data sets, and afterward identifies and deduces patterns with a specific goal to apply that data to new arrangements and activities. Clearly, necessary parameters must be build up toward the start of the machine learning method so the machine can discover, evaluate, and follow up on new information.

Fernandes, Ramos and Rosa (2005) and Mishra (2011) state that these approaches of the Artificial Intelligence, or Machine Learning models, are distributed, as seen on Table 2:

Table 2 – Artificial Intelligence

Artificial Intelligence	Characteristics
Genetic Algorithm	Model used for machine learning. Moreover, it is a method used by evolutionary algorithms. The objective of this model is to simulate specific genetic operators like crossover, mutation and reproduction.
Fuzzy Logic	Reasoning way that uses logical expressions to demonstrate the degree of membership in fuzzy sets.
Systems Based in rules	Systems that perform intelligent behavior of human connoisseurs.
Genetic programming	Area dedicated for building programs, in order to copy the natural process of genetic using the random search.
Case-based reasoning	It represents and processes knowledge to solve problems through recovery and adaptation of experiences, which are titled as cases.
Neural Networks	Area that focuses the processing and storage of experimental knowledge and make it available. The knowledge obtained by the network according to its environment through learning and the connections between neurons.
Decision Trees	It is one of the most successful techniques for supervised classification learning. It assumes that all of the features have finite discrete domains, and there is a single target feature called the classification.
Support Vector Machine	SVM converts input space to a higher dimensional feature space, where data are separable.
Convolutional Neural Networks	They present multi-layer models that are totally able to be trained, with characteristics of catching nonlinear mappings through inputs/ outputs
K-Nearest Neighbors	KNN classifies neighbor images to the one in case, according to Euclidean distances.

Source: Adapted from Fernandes, Ramos and Rosa (2005) and Mishra (2011)

Machine learning is the computerized procedure of revealing examples in extensive datasets with analytical models based on computer, where an adapted model might then be utilized for expectation reasons on new information, i.e., it is the process of identifying the relationships between predictor and answer variables using computer-based statistical approaches (HEUNG et al., 2016).

AI presents several techniques of Machine Learning for industrial applications such as Genetic Algorithms, Inductive Learning, Tabu Search, Simulated Annealing, Fuzzy Logic, Neural Networks, Decision Tree, among others. They have been used for a while within industrial areas as data mining, classification, design, control, modelling, optimization, diagnosis and prediction machines (PHAM; ALCOCK, 2003).

4.4 ARTIFICIAL INTELLIGENCE APPLIED TO IMAGE PROCESSING

Several authors have used AI applied to image processing, specifically for wood classification, such as Brunner et al. (1990) who performed by utilizing computer tomography with four channels of colors, i.e., RGB and Gray-scale. The results were quite unsuccessful due to the heavy computational pressure on the system, and that the technology of color cameras were not intended to assist image processing.

Other issues, concerning computer tomography scanning technology were the amplitude of samples necessary to achieve acceptable results, and the cost of equipment, often prohibited for the size of most wood processing factories (SCHMOLDT et al., 2000).

On the contrary, empirical study of Marszalec and Pietikainen (1993), also applying the same technique during wood inspection, concluded there were gains in velocity and efficiency throughout the process.

Images acquisition may happen in many ways. Still, the objective is essentially the same, specifically to produce advanced images from sensed data. Most sensors align a consistent voltage wave with amplitude and spatial conduct identified with the physical aspect of the image. At that point, to modify it to digital, it needs to change sensed data into digital structure, with sampling and quantization (GONZALES; WOODS, 2008).

As expressed by Al-Allaf, Abdalkader and Tamimi (2013), images are presented in a continuous form, regarding their x and y coordinates, and their amplitude. Sampling occurs when digitizing the coordinate values, and quantization, the amplitude ones. The idea of using higher number of layers, the multi-layer network, is justified by the method used to train a network to use a specific number of layers in imaging classification. The training rule for a network minimizes the error measure. The required stages for sawn softwood images are acquisition, division, standardization, characteristic extraction and coordinating.

The normalization process is used since the images of wood boards may be captured in different sizes, and with distinct illumination, among other variables. The application of feature extraction is because wood provides enough texture information, and a feature vector is formed to carry the ordered sequence of those features extracted from the different descriptions of the samples. Then, the feature vectors are labelled using ANN models in matching process (HAMID; MEYBODI, 2009).

4.5 DIGITAL IMAGE PROCESSING

There are some differences between image processing and other areas such as image analyses and computer vision. However, it is considered that image processing has as input and as output images. On the other hand, computer vision deals more with AI technology, which is proposed to emulate human attributes, such as learning capacity, and ability to make decision. But they are all under the same field of study.

First, an explanation on image is necessary. It can be defined as a two-dimensional function, $f(x,y)$ which the amplitude of the pair (x,y) is known as the gray level (intensity). When the values of the pair (x,y) are all finite, discrete quantities, it is named digital image. The elements that compose a digital image are called pixels, pels, picture elements, image elements (NAGABHUSHANA, 2006).

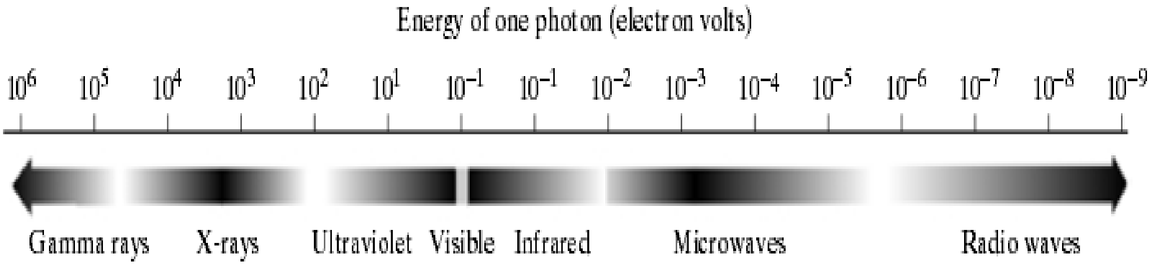
Returning to imaging, according to Gonzales and Woods (2008), it is also possible to characterize the areas of digital image considering three types of computerized processes:

- Low-level – contrast enchantment, reducing noise, image sharpening
- Mid-level – segmentation, reducing objects to a form suitable for computer processing
- High-level – making sense of an ensemble of recognized objects

The first applications of digital images were in the newspaper industry, which major problems were related to the selection of printing procedures and the distribution of the intensity level. As the computer science largely developed, computer and software have got cheaper, allowing the development of digital image processing focusing in a vast number of areas; medicine, astrology, civil and military are some examples of the areas benefited by digital image processing (SONKA; HLAVAC; BOYLE, 1999).

Nowadays, digital image processing influences almost all areas of specialized endeavors, or at least these areas are affected by it somehow, as stated by Nagabhushana (2006). One of the least difficult approaches to build up a fundamental comprehension of the degree of image processing applications is to sort images as indicated by their source, for example, X-rays, ultraviolet, microwave, visual, etc. The major used energy source for images today is the electromagnetic energy spectrum (Fig. 7). Other significant sources consist of ultrasonic, electronic, and acoustic.

Figure 7 – Energy per photon on electromagnetic spectrum



Source: Gonzalez; Woods (2008)

Computers create synthetic images, usually used for modeling and visualization. Images that are dependable of radiation from electromagnetic spectrum are the most ordinary, particularly those in the visible and X-ray bands. The electromagnetic wave propagation can be defined as sinusoidal waves of different wavelengths, or as a massless particle stream, which travels in a wave-shaped pattern and moving at the speed of light. Every massless particle contains certain amount of energy. Each of these quantities is called a photon. If the spectral bands are grouped according to the energy per photon, yields the spectrum shown in Figure 7, they range from gamma rays (highest energy) at one end of the spectrum to radio waves (lower energy) in the other (SHAPIRO; STOCKMAN, 2001).

It is impressive to relate the different types of waves, such as radio, microwaves, infrared, visible, ultraviolet, x-ray, gamma rays to different types of applications, meaning that it is possible to have a better use of digital image processing for specific areas. This fact is explored on Table 3, which presents some usual applications with image processing technologies.

Table 3: Technologies for Imaging

Image Processing Technologies	Usual Applications	Wood Application
Gamma Rays	Nuclear medicine and astronomical observations	Generate tomographic cross section images of logs; Recognize knots and their location. Measure density of wood.
X-rays	Medical and Industrial fields	Detect knot ringlets and heartwood; Gathers information about annual rings; Measure the density of the piece; Detect internal knots and rot.
Ultraviolet	Lithography, industrial inspection, microscopy, lasers, biological and astronomical observations	Cleaning laser fluence for ground wood paper; Inhabit wood fungi and their effect on lignin
Visible + Infrared	Light microscopy, astronomy, remote sensing, industry	Detect edges of logs for diameter and length measurement; Locate knots in wood.
Microwaves	Radars - ability to collect data over virtually any region at any time	Evaluate the moisture of the wood; Detect knots and metallic objects; Determine grain angle by depolarization
Radio Waves	Magnetic resonance in medicine	Differ knots, grain distortions and regular wood; Estimate lumber strength.
Acoustic Sensor	Assess properties from a variety of materials based on their mechanical vibrations through ultrasonic waves	Stress grading of logs; Strength of boards; Recognize harmful internal cracks; Sorting of logs via intrinsic wood quality.
Laser	Surface engineering, texturing, alloying and treatment for diverse materials; Micromachining on optical fiber.	Geometric properties identification, such as bow, crook, twist, warp, holes and wane; Measurement of fiber direction, and width, thickness and length of boards; Measure grain orientation and the strength of boards through vibrations.
Line Scan Color Camera	Detect foreign fibers in cotton lint; industry	Detect knots, pith, blue/red stains, decay, bark pockets, cracks, wane, splits, shakes, holes and grain angle. Classify wood quality; Measure dimensions of wood pieces.

Source: Adapted from Gonzalez; Woods (2008)

There are companies that include several products in their portfolio, which are used to classify wood logs and/or wood boards, with distinct hardware apparatus, such as cameras, lasers, acoustic sensors, LED lighting, scans, line scans, computer tomography, infrared, radio waves, ultraviolet and even gamma rays, as seen in Table 4. Also, a combination of electronic components might magnify the capacity of the application to certain areas or specific problems.

The applicability of each one of these hardware technologies will depend on the necessity of the classification problem faced on different branch of industry. In addition, other issues may contribute for the choice of one or a combination of image processing technologies (Table 4).

Table 4: Available Products on the Market and their Features

Commercial Product / Company	Hardware Equipment	Technical Features
ProfiGrade / Lisker Oy	Laser + Digital matrix cameras	Measure shape of logs and boards; Grade sawn wood; Detect wood defects.
LuxScan / Weinig	Color Cameras + Lasers + X-rays + Acoustic sensor	Detect board defects; Measure shape parameters; Grade board stress.
SprenScan / Sprecher Automation	Camera + Laser	Detect alignment of annual rings; Estimate position of pith; Detect wood defects
Wood-X / Limab Oy	On line Laser + Cameras + X-ray	Measure sawn wood dimension; Measure volume of logs; Detect wood defects.
Optilog / FinScan Oy Decon Ab	Line Color Cameras + Scan	Detect board deformation; Grade and sort board and veneer sheets; Detect wood defects.
Visual Defect Detector / Autolog	CCD camera + LED lighting	Measure profile and dimension of sawn wood; Detect knots; Grade boards by knot properties.
Linear High Grader / USNR	X-rays + Camera + Infrared Laser	Measure geometric profile and density; Detect 3D wood defects; Detect planed bark and pith location.
Visual Defect Analyzer / Mecano Group Oy	Color camera	Detect defects on veneer sheet; Control production, grading, clipping or patching
DynaVision / LMI	3D laser + Scan + Light curtain color vision	Measure dimension of boards and logs; Detect defects; Measure, grade, sort plywood, logs and sawn wood.
WoodEye / Innovativ Vision AB	Camera + Laser	Detect the position, shape and extent of common wood defects; Measure fiber direction and width, thickness and length of the timber
RemaLog Bark / RemaControl	Scanner + Camera + Laser	Detect and classify wood defects; Measure shape of log; Gather information on annual rings; Distinguish bark region;
NV2000 / Ventek	Camera + Lighting	Detect and grade plywood defects; Patch plywood; Measure moisture.
Automatic Quality Sorting / SKS Vision Systems Oy	Camera + Laser	Measure log dimensions; Assess wood quality from color images and cross section surfaces.
Optiside / Microtec	X-rays + Laser + Infrared + Camera + Tomography	Measure the dimensions and shape of pieces; Detect internal and surface defects; Assess the strength and moisture of wood; Estimates cup of board, pith location, distance between annual rings.
GradExpert / Comact Inc.	Line Scan Color Cameras + Laser	Measure geometric properties of board; Detect wood defects.

Source: Adapted from Österberg (2009)

As seen, among the ways of acquiring image data about a given object, the non-destructive test (NDT) or non-destructive evaluation (NDE) devices are commonly used, once they avoid damaging the product during inspection. Sensors can automatically be used as ultrasonic, acoustic, infrared, thermographic, and microwave, among other technologies, as magnetic resonance imaging, x-ray, gamma ray, computed tomography, camera and/or digital camera, visible light imaging, laser technology, electric measurements, and radio frequency scanning.

Combination of those equipment helps to improve performance of measurements and classification. For instance, adding scan sensors to 3D laser machine vision sensors is possible to accurate dimension measurement and defect detection. Besides measuring, it grades and sorts sawn wood, logs and plywood, as it was noticed on Table 4.

In the context of machine vision, image processing is used to enhance, filter, mask and analyze images. Well-known machine vision software packages are available at the market or it is possible to create low-cost ones, are used to create advanced and powerful automation tool that will take an image input and output data based on the content of the image. Ultimately, in commercial machine vision, image processing is used to classify, read characters and noises, and recognize shapes, measure or defects (UKIVA, 2002).

The applicability and execution of the technologies mentioned above commonly have the highest performance for sawn wood boards and the most reduced for logs. For instance, microwaves infiltrate through wood boards more appropriated, when compared to the infiltration through logs. Likewise, system provided with cameras can identify each knot in boards, although it will find difficulties in logs (ÖSTERBERG, 2009).

As observed on Table 4, cameras are the mostly used equipment for classification and measurement of wood boards, veneers and logs, reason why they deserve greater depth of information, and will be discussed on the section Materials and Methods.

5. TECHNIQUES FOR IMAGE CLASSIFICATION

This section demonstrates some machine learning techniques commonly used for object classification. It highlights the Neural Networks, also called Artificial Neural Networks, and then continues to Deep Learning with Convolutional Neural Networks. Then, it proceeds to other traditional Machine Learning techniques, such as K-Nearest Neighbor, Support Vector Machine, and Decision Trees.

5.1 ARTIFICIAL NEURAL NETWORK - ANN

Networks, as their operation methods, were sorted out into control, inference, and characterization. Networks operating in inference are those that after learning, the synaptic weights found are used as parameters for the determination of values of pairs (input, output) not correlating to the initial database. Such networks are otherwise called inference mechanism (AFFONSO; SASSI; BARREIROS, 2015).

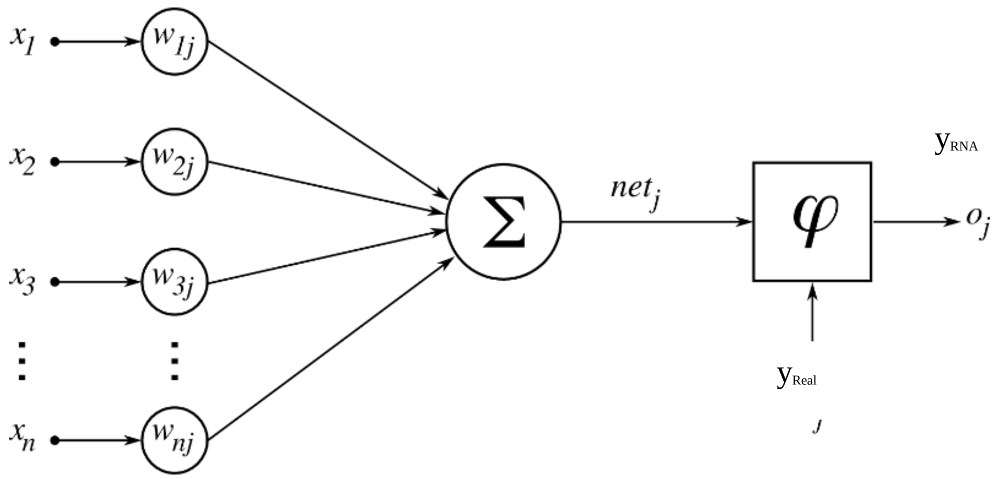
Networks, utilized as a part of the control operation, have the capacity of feedbacking the system by adapting their parameters, to achieve their ideal point or keep up their stability.

In the case of neural networks, they have been applied into various circumstances in the last years, ranging from assessment to processing, classification, prediction, and engineering, among others.

The natural nervous systems motivate ANNs, for example, is the biological brain, which comprises of a substantial number of exceptionally linked components called neurons. This part of the human body gathers and processes the information by adapting the connecting arrangement of the neurons. Naturally, the system capacity is resolved broadly by the associations between those components. A neural system can be prepared to perform a specific action by modifying the estimations of the connections, or weights, among the components. Ordinarily neural systems are adapted or trained, in order that a specific input activate a particular target output. Commonly, countless such input/output sets are utilized as a part of this managed learning, to train a network (KANTA; SANGWAN, 2015).

The beginning of the study on ANNs is the article by Mcculoch and Pitts (1943). That paper proposed that with a sufficient number of artificial neurons is possible to represent logical expressions. After that, other works by Hebb (1949) and Rosenblatt (1958) add contributions in developing the ANN theory. Figure 8 shows a representation of the artificial neuron.

Figure 8. Representation of the artificial neuron



Source: Affonso; Sassi; Barreiros (2015)

An early ANNs, whose architecture is based on the biological neuron, was proposed by Rosenblatt (1958), the Perceptron. The objective of that network is to classify x_i inputs (or stimuli) into two classes by a hyperplane.

In unsupervised learning (or self-supervised), the desired outputs values φ are not known, and learning occurs through the identification of input patterns.

The architecture of an ANN depends on the type of problem in which the network should be used, and it is defined among other factors, by the number of layers, number of nodes in each layer, the type of connection between the nodes (feedforward or feedback) and by its topology (KOVÁCS, 1996).

According to Haykin (2001), one of the most used training algorithms in Multilayer Perceptron - MLP is error backpropagation that works as follows: a pattern is presented to the input layer of the network, then it is processed layer by layer until the output layer provides the processed response, f_{MLP}, according to function (4):

$$f_{MLP}(x) = \varphi(\textstyle\sum_1^n v_l \,.\, \varphi(\sum_1^m w_{lj}\, x_l \;+\; b_{l0}) + b_0) \qquad (1)$$

where v_l e w_{lj} are synaptic weights; b_{l0} e b_0 are the biases, and φ the activation function, commonly specified as the sigmoid function.

The principle of this algorithm for calculating the errors in the intermediate layers is its feedback calculation, allowing the adjustment of weights in proportion to the values φ of the connections among layers (AFFONSO; SASSI; BARREIROS, 2015).

As reported by Haykin (2007) MLP has the following characteristics: non-linear activation function (sigmoidal), one or more layers of hidden neurons and a high degree of connectivity.

Haykin (2001) states that mathematically an ANN is a method for determining a function $f: A \rightarrow B$, where A is a set of data, and some elements of B are known a priori.

The determination of the function f is possible thanks to Andrei Kolmogorov Theorem on representation and approximation of continuous functions, as shown below.

Any continuous function $f\ (x_1)\ i = 1,\ m$ with image In = [0;1] m can be represented as function (5):

$$f(x_1) = \sum_{q=1}^{2m+1} g\ (\sum \lambda_p\, h_p\ (x_p\ ))\ _{i=1,\ m\ |\ p=1,m} \qquad (2)$$

The Kolmogorov Theorem provides an exact representation for a continuous function, but it does not indicate how to obtain these functions, and according to Mitchell (1997) this approach is similar to the study of non-parametric statistical inference.

In the classification operation, also called clustering, unsupervised neural networks map the search space of the problem, clustering elements with similar characteristics. Networks used in the control operation have the function of feedbacking the system by adjusting their parameters, in order to reach their optimum point or maintain their stability.

Neural Networks uses are wide and their application is not restricted to a particular field. In order to illustrate, Table 5 presents scientific papers in the areas of Engineering, Computer Science, Chemistry, Agricultural, Astronomy, Medicine, Environmental Science, Social Sciences, Economics, Imaging, Decision Science, Energy, Materials Science, and Neuroscience.

Table 5 - Different uses of Neural Networks in chronological evolution

Year	Author(s)	Uses of Neural Networks
1990	Optiz, D.	Autonomous Mobile Robots
1991	Sigillito et al.	Radar Signal Processing
1992	Venkatasubramanian, V.; McAvoy, T. J.	Chemical Engineering
1992	Bochereau et al.	Prediction of Apple Quality
1993	Miller, A. S.	Astronomy: Adaptive Telescope Optics, Object Classification and Matching, and Detector Event Filtering
1994	Choi et al.	Optimization of Process Parameters of Injection Molding
1995	Burke, L.; Kamal, S.	Cellular Manufacturing
1995	Guleyupoglu, S.; Smith, R. E.	Prediction of Hypervelocity Orbital Debris Impact Damage to the Space Station
1996	Shustorovich, A.; Thrasher, C. W.	Positioning and Classification of Handwritten Characters
1996	Agrawal, D.; Schorling, C.	Market Share Forecasting
1998	Cramer, C.	Image and Video Compression
1998	Chen et al.	Understanding of Cerebral Processing of Human Pain
2000	Galati et al.	Rainfall Rate Extraction
2001	Vivarelli, F.; Williams, C.	Classifying Segmented Outdoor Images by Comparing Bayesian
2001	Ayrulu, B.; Barshan, B.	Improved Target Differentiation and Localization with Sonar
2002	Venkataram et al.	Optimal Routing Algorithm for Communication Networks
2003	Scardi, M.	Coastal Ecological Modeling
2005	Yalcintas, M.; Akkurt, S.	Building Energy Predictions for Tropical Climate
2006	Lisboa, P. J.; Taktak, A.	Decision Support in Cancer
2007	Van der Merwe et al.	High Dimensional Physics-Based Models in Computational Oceanography
2007	Panagou et al.	Modelling Fungal Growth using Radial Basis
2008	Mazurowski et al.	Training Classifiers for Medical Decision Making / Imbalanced Datasets on Classification Performance
2009	Adeli, H.; Panakkat, A.	Earthquake Magnitude Prediction
2010	Wysoski et al.	Audiovisual Information Processing
2011	Yazdi, M. ; Khorasani, A.; Faraji, M.	Optimization of Coating Variables for Hardness of Industrial Tools
2011	Kaya, B.; Oysu, C.; Ertunc, H. M.	Force-torque based on-line tool wear estimation system for CNC milling
2012	Carrilo et al.	Hardware Implementations using Traffic-Aware Adaptive Network-On-Chip Routers
2013	Saritha et al.	Classification of Magnetic Resonance Imaging of the Brain
2014	Loh, K. J.; Ryu, D.	Multifunctional Materials and Nanotechnology for Assessing and Monitoring Civil Infrastructures
2015	Powar, A.; Date, P.	Modeling of Microstructure and Mechanical Properties of Heat Treated Components
2015	Marsico et al.	Robust Face Recognition after Plastic Surgery using Region-Based Approaches

Source: Own Author

According to Haykin (2001), the use of Neural Networks offers the following useful properties and capabilities: Non-linearity, Input-output mapping, Adaptability, Response to evidence, Contextual information, Fault tolerance, Implementation of VLSI (very-large-scale-integration), and Uniformity of analysis and design.

As reported by Gu, Pan and Wang (2015) there are numerous neurons which compose network to compile and process data in the intellect. Artificial neural network is an extraordinary system which imitates the method the mind works.

We can deduce from this that neural network is a nonlinear framework that is made out of countless basic processing element. At a certain point, it copies the processing and compiling information behavior of the person nerve system. Considering its particular structure, neural network is predicted to figure out what cannot be explained by the conventional techniques.

5.2 DEEP NEURAL NETWORKS

The use of deep neural networks provide state-of-art performance on different sort of objects, such as images. Nowadays, there is a huge adoption of images resources for acquisition, transmission, restoration, enhancement, search, retrieval, recognition, and tagging. For doing so, it is imperative they have skilled and solid methods of image processing. Manap and Shao (2015) highlighted the image quality assessment, a type of image classification, which is able to achieve crucial aspects in several computer vision and image processing applications.

As presented by Hinton et al. (2012), deep neural networks are effective learning models that accomplish exceptional execution on visual and discourse recognition issues. Neural networks carry out great performance since they can explicit self-assertive computation that comprises of an unobtrusive number of parallel nonlinear steps.

In any case, according to Bengio (2009), as the subsequent processing is naturally found by backpropagation by means of supervised learning, it can be hard to clarify and can have irrational properties. Two of these properties are distinct of one another, because of their interest. While one of them deals with the semantic meaning of individual units, the other treat the stability of neural networks regarding small perturbations to their input.

The first property, concerned with the semantic meaning, reviews singular units and makes the verifiable supposition that the units of the last highlight layer form a recognized basis, which is especially valuable for separating semantic data. Past works broke down the semantic significance of different units by discovering the arrangement of inputs that maximally enact a given unit. Arbitrary projections of ϕ (x) are semantically vague from the coordinates of ϕ (x). Additionally, it brings to discussion the guessing that neural networks unravel variety considerations crosswise coordinates (MRAZOV; KUKACKA, 2012).

As stated by Li et al. (2015), for the most part, it appears that it is the whole space of activations, as opposed to the individual units, that contains the majority of the semantic information. A comparable, however significantly stronger conclusion, was showed by Mikolov et al. (2011) for word representations, in which the different directions in the vector space performing the words are demonstrated to offer ascent to a shockingly rich semantic encoding of relations and analogies.

In the meantime, the vector representations are distinct up to the space rotation, so the individual units of the vector representations are unbelievable to contain semantic information.

The second property takes into consideration an ultramodern deep neural network that sums up well on an object recognition assignment. It is expected that this network will be powerful to small input noises, once they cannot change the object classification of an image. Nevertheless, applying an intangible non-arbitrary noise to a test image is conceivable to self-assertively change the network's prediction. These noises are found by improving the input information to boost the expectation error (BENGIO, 2009).

Normally, the exact configuration expectation of the minimal fundamental noises is an arbitrary design of the ordinary variability that emerges in diverse runs of backpropagation learning. Still, adverse samples are moderately powerful, and are distributed by neural networks with diverse number of layers, prepared on distinctive subsets of the trained data. In other words, when a neural network is used to create a group of diverse samples, these are statistically difficult for another one, even when it was prepared with distinctive hyperparemeters or, most shockingly, when it was prepared on an alternate set of samples (ERHAN et al., 2009).

These outcomes recommend that the deep neural networks that are found out by backpropagation have non-instinctive qualities and peculiar blind spots, whose structure is associated with the information dispersion in non-clear ways.

There are some types of Deep Machine Learning Architectures, such as Deep Belief Networks - DBNs, Stacked (Denoising) Auto-Encoders, Hierarchical Temporal Memory, Deep Spatiotemporal Inference Network - DESTIN, and Convolutional Neural Network, each of them with specific characteristics (AREL et al., 2010).

5.3 DEEP LEARNING – DL

The recent revival of Deep Learning models was triggered by the works on learning representations, or more traditional models (HILTON et al., 2012).

Deep learning architectures appear to solve problems that require complex highly varying functions. Besides, usually involve such problems with very large dataset, and in most cases non-labeled.

As stated by Bengio (2009), in order to deal with it, Deep Learning methods learn feature hierarchies with characteristics from higher levels of hierarchy formed by a composition of lower level features.

DL learns complex behaviors with expansive information sets to select effective characteristics automatically by neural network structures in quite profound layers. These models achieve such goals adopting unsupervised layers succeeded by supervised ones, applying learning teaching to signal data (KIM; CHOI; LEE, 2015).

5.4 CONVOLUTIONAL NEURAL NETWORK – CNN

Since Fukushima's Neocognitron (FUKUSHIMA, 1980), computational models built on local connectivity intervening neurons and hierarchically composed changes of the image have been used. The author perceived that when neurons with the same parameters are connected on patches of the preceding layer at diverse areas, a manifestation of translational invariance is acquired.

Convolutional Neural Networks (CNN) models (LECUN; BENGIO, 1995), in addition to neocognitrons (FUKUSHIMA, 1980) referred to as one of naturally propelled models, have been utilized for setup recognition tasks, for example, face perception and manually written numeral recognition.

The network incorporates feature detecting (FD) layers, each of which substituted with a pooling layer to get properties that interpret invariance. Lately, a two-independent

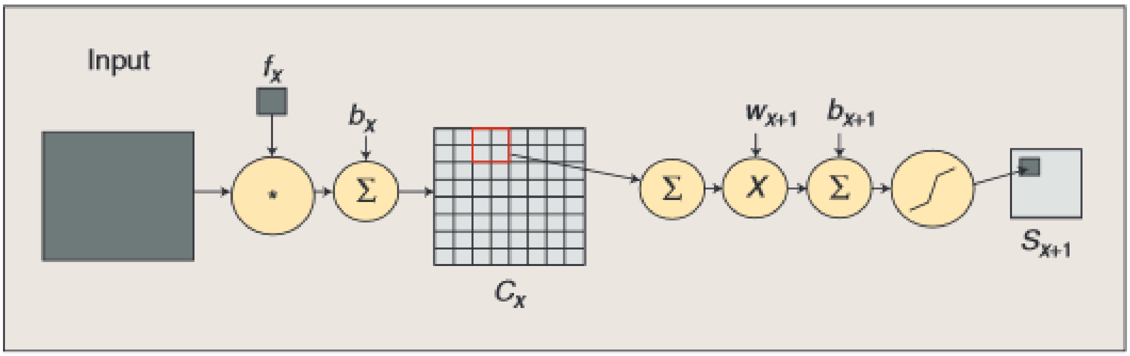
CNN model was proposed, one for face and the other for identity perception, which are joined by a Multilayer Perceptron - MLP (FASEL, 2002).

Despite the fact that deep neural networks were commonly thought excessively troublesome, making it impossible to train well; the Convolutional Neural Networks - CNNs represent a remarkable exception. Hubel and Wiesel (1962) proposed models of structure for visual system in which the convolutional networks were based.

CNN is a kind of multi-layer, completely trainable model that can catch exceptionally nonlinear mappings in the middle of inputs and outputs. These models were initially inspired from computer vision issues and therefore are characteristically suitable for image related applications (KRIZHEVSKY et al., 2012).

As seen on figure 9, CNN possesses two processes, a convolutional and a subsampling. The convolution one consists of convolving an input (the image indicated on the first stage is convolved for next stages) with an instructible filter f_x, subsequently adding a trainable bias b_x to form the convolution layer C_x. The subsampling consists of counting a neighborhood (four pixels), weighting by scalar $w_x + 1$, adding trainable bias $b_x + 1$, and crossing a sigmoid function to form a approximately 2 times smaller feature map $S_x + 1$ (AREL et al., 2010).

Figure 9 – Convolution and Sampling Process in a CNN



Source: (AREL et al., 2010)

LeCun and his associates also classified the layers of convolutional neural networks in two sorts. The first one is called convolutional layers and, the other is referred as sub-sampling layers. Every layer has a topographic structure, or in other words, every neuron is connected with an established two-dimensional position that coincide to an area in the input object, alongside a responsive field, which is the location of the input object that impacts the reaction of the neuron. At every region of every layer, a variety of distinctive neurons is found, each with its weights, connected with neurons in a rectangular area, also known as patches, in the preceding layer. The same weights, although with an alternate input rectangular patch, are connected with neurons at distinctive areas (LECUN et al., 1998a).

Succeeding these findings, other authors lead by LeCun caught up on this thought and prepared such systems utilizing the error gradient, getting high performances on a few vision works (LECUN et al., 1989; LECUN et al., 1998b). Advanced comprehension of the physiology of the visual system is persistent with the processing form discovered by convolutional nets, at least for the recognition of objects in a fast way, or without close attention and feedback connections (SERRE et al., 2007).

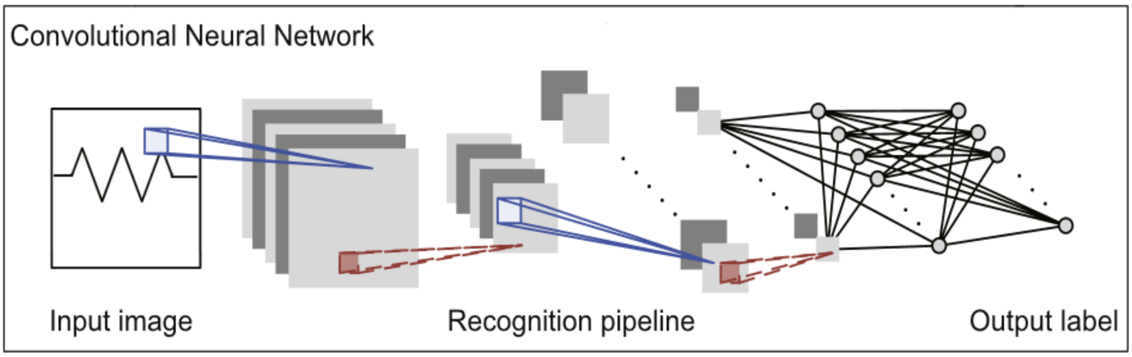
LeCun et al. (2006) state that currently vision systems built on convolutional neural networks achieve the most appropriated results. This has been indicated in LeCun's research on handwritten character recognition (LECUN et al., 1998b), and has been a model for a long time.

Regarding the training of deep architectures, the case of convolutional neural networks is impressive once they normally uses a heavy amount of layers, which makes completely joined neural systems just about difficult to advance appropriately when initialized randomly (SIMARD; STEINKRAUS; PLATT, 2003).

While the majority of classical image processing techniques are based on carefully pre-selected image features, or standardized images, Convolutional Neural Networks (CNN) are, namely, designed to learn local features autonomously. Especially deeper neural networks, not necessarily convolutional neural networks, are known to extract better compact representations of high-dimensional data than shallow architectures (MATSUGU et al., 2003).

Fu and Kara (2015) remarked that the scheme of Convolutional Neural Network basically brings three parts: the object, i.e. the images that will be processed comes first, known as input images; then, the second part carries the recognition pipeline, where the images will be divided into smaller parts; and finally the last part, which is the output label, in which the classification occurs, established through the verification of the database, as it can be observed in Figure 10.

Figure 10 – Convolutional Neural Network scheme



Source: Adapted from Fu and Kara (2011)

Lately, CNN has turned into research center of attraction again, due to its excellent conceptions of shared weights, spatial sub-testing, and local receptive fields. Thus, it can guarantee some level distortion, scale and shift invariance. In addition, this division of neural network can specifically learn features and process images. Because of the numerous convolutional layers whose convolutional weights are prepared by the backpropagation calculations, being generally applied into diverse fields, for example, face detection and traffic sign and document recognition (FU et al., 2015).

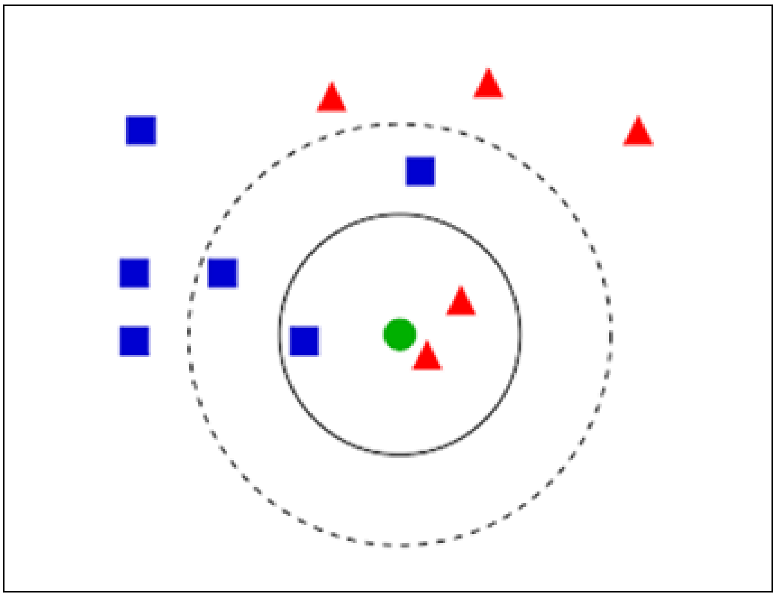
5.5 K-NEAREST NEIGHBORS – KNN

According to Sánchez et al. (2016), the K-nearest neighbors algorithm is a machine learning technique that has been widely used in pattern recognition and classification. The KNN method is a non-parametric algorithm that has been widely used for classification tasks. The algorithm gathers every accessible example as an input and classifies to the new cases, as output, based on an affinity measure.

KNN techniques are a class of multivariate, nonparametric approaches to continuous or categorical prediction. With these techniques, predictions are calculated as linear combinations of observations for population units in a sample that are similar or nearest in a space of auxiliary variables to population units requiring predictions (CHIRICI et al., 2016).

Figure 11 brings the classification through KNN, and it can be observed that two possibilities are available with different results. If one considers k = 3, the output is distinct of k = 5, where k represents the number of neighbors, since the technique recognizes the similarity of the object by its closeness to others. In other words, KNN classifies neighbor images to the one in case, according to their distances.

Figure 11 – K-Nearest Neighbors representation



Source: Adapted from Kim; Kim; Savarese (2012).

As defined by Duda, Hart and Stock (2001), the KNN classifier exploits a training data matrix X, in which every object is part of a class "*n*" out of "*N*" conceivable classes. This classifier allows an unknown object x_t, to the class to which the greater part of the k-nearest neighbors fits. These neighbors are found by suitable metric, normally the Euclidean distance.

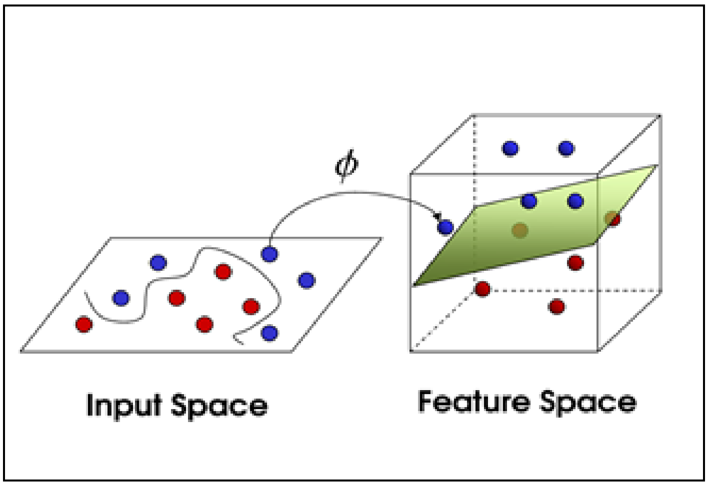
There are a few varieties of the KNN system, assuming the sort of distance utilized or the decision rule applied for classification (VILLA; BOQUÉ; FERRÉ, 2008).

5.6 SUPPORT VECTOR MACHINE – SVM

Support Vector Machines (SVMs) are an arrangement of related regulated learning routines utilized for classification and regression (VAPNIK, 1995). They are part of a group of ambiguous linear classifiers. In other words, Support Vector Machine (SVM) is a prediction and classification engine that uses machine learning hypothesis to amplify prescient exactness while consequently keeping away from over-fit to the data. Support vector machines can be characterized as structures, which utilize speculation space of linear functions in a high dimensional feature space, prepared with a learning algorithm from streamlining hypothesis that actualizes a learning bias device from measurable learning theory. It is a dynamic part of the machine learning research.

It is noticed in figure 12 that SVM converts input space to a higher dimensional feature space, where data are separable.

Figure: 12 – Support Vector Machine representation



Source: Adapted from Gunn (1998).

SVM gets to be popular whenever utilizing pixel maps as input data, it gives exactness equivalent to advanced neural networks with expounded highlights in a handwriting recognition assignment, for instance. It is likewise being adopted for some applications, i.e. hand writing examination, face analysis and so on, particularly for pattern classification and regression based applications. Vapnik (1997) first established Support Vector Machines and picked up prominence because of numerous encouraging features, as better observational performance. Support Vector Machines were created to take care of classification issue, however in recent years they have been stretched out to take care of regression issues, as well (VAPNIK; GOLOWICH; SMOLA, 1997).

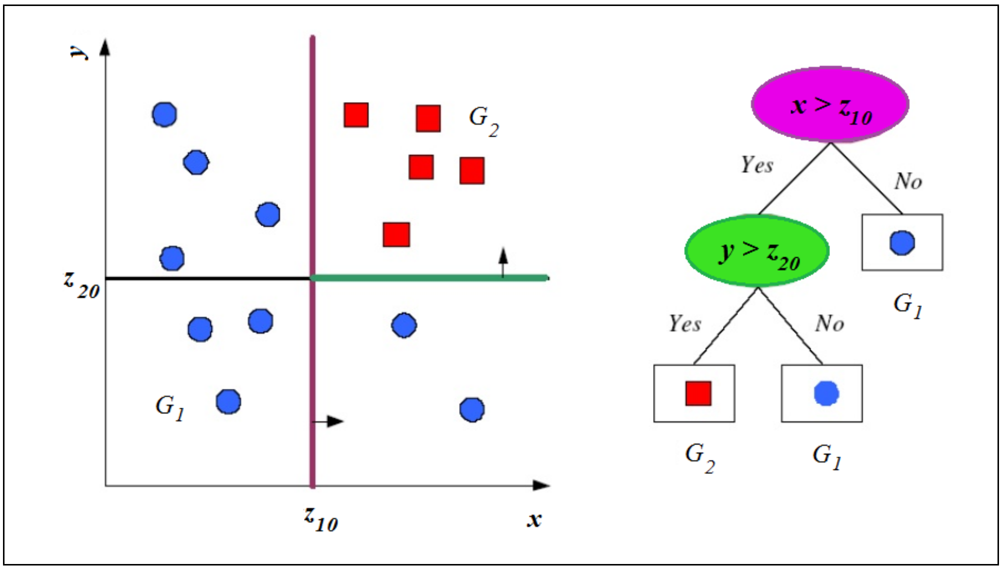
According to Wu et al. (2015) the essential thought of Support Vector Machines is to convert data points from data space to a high dimensional feature space, utilizing a kernel function in order that the data points in the feature space, turn out to be linearly divisible.

5.7 DECISION TREES – DT

DT are common classifiers that induce a rooted tree structure that the terminal nodes, which would be similar to the leaves of that tree, represent class labels, while the branches correspond to features, with associated values leading to the nodes (MA; DESTERCKE; WANG, 2016).

Quinlan (1986); Okach and Maimon (2005) stated that a decision tree is a classifier expressed as a recursive partition of the instance space, and it has been the most widely used approach to represent classification models, mainly due to its comprehensible nature. The structure of a decision tree consists of nodes, as it can be observed in figure 13, which have exactly one incoming edge, except the root node, in the top of the tree. A node with outgoing edges is called an internal or test node. Each internal node splits the instance space into two or more sub-spaces according to an attribute. All other nodes are called leaves. Each leaf represents the most appropriate class. Instances are classified starting at the root and goes down to a leaf node, according to the outcome of the tests along the path.

Figure 13 – Decision Tree representation



Source: Adapted from Lu and Yang (2009)

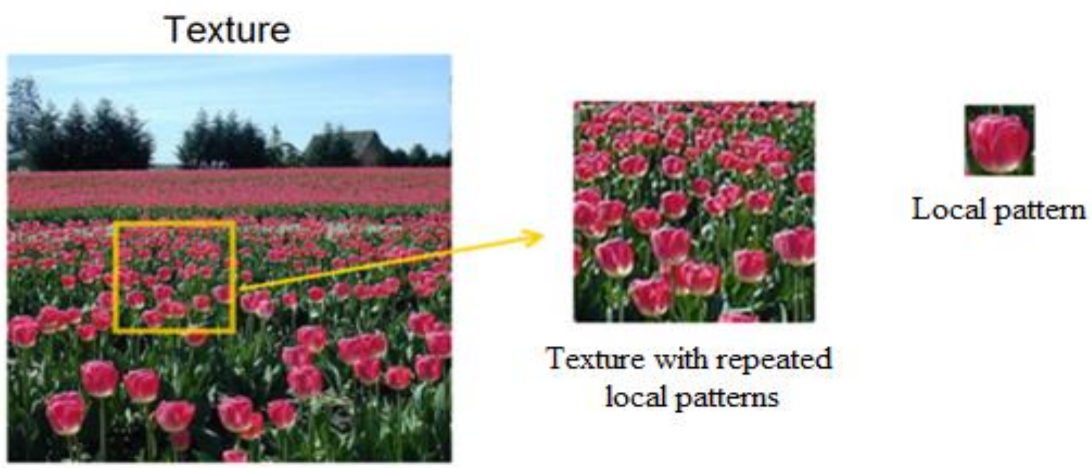
As stated by Ma, Destercke and Wang (2016) decision trees are widely used for classification approaches, due to their favorable learning capacity and clarity of interpretation. Nonetheless, classical decision trees can only execute certain information whose values are accurately known.

6. TEXTURE DESCRIPTORS – TD

This section specifies the image texture definition, which indicates image intensity features and patterns, its differentiation and texture description. Then it describes possible approaches to texture descriptors such as the structural and the statistical, highlighting the second one, as it was used in this thesis.

Texture is an internal property of almost every natural surfaces such as wood, weave of a fabric, patterns in sand, leaves, etc. It contains information about the structural arrangement of surfaces and their relationship to the environment (WEI; HONG-YING, 2016). Texture represents details in an image, as seen on Figure 14.

Figure 14 – Repeated texture patterns



Source: Adapted from Yao et al. (2012)

Texture is an effective signal for describing scene structures via images that show a high level of likeness in their arrangement, namely repeated patterns; allow numerous computer vision assignments, for example, image acquaintance and classification; scene grouping; and visual recovery (QUAN; XU; SUN, 2014).

According to Haralick and Shapiro (1991) the main concern of texture is respected to the spatial arrangement of the image intensities and discrete tonal features. Whenever a limited area of the image has little variety of discrete tonal features, the overwhelming property of that region is gray stress. And whenever a limited area of the image has large variety of discrete tonal features, the overwhelming property of that region is texture.

Three aspects are vital in this differentiation: the extent of the limited areas; the relative proportions of the discrete tonal features; and the quantity of recognizable discrete tonal features. Texture can be explained along ranges of density, regularity, coarseness, roughness consistency, directionality and amplitude.

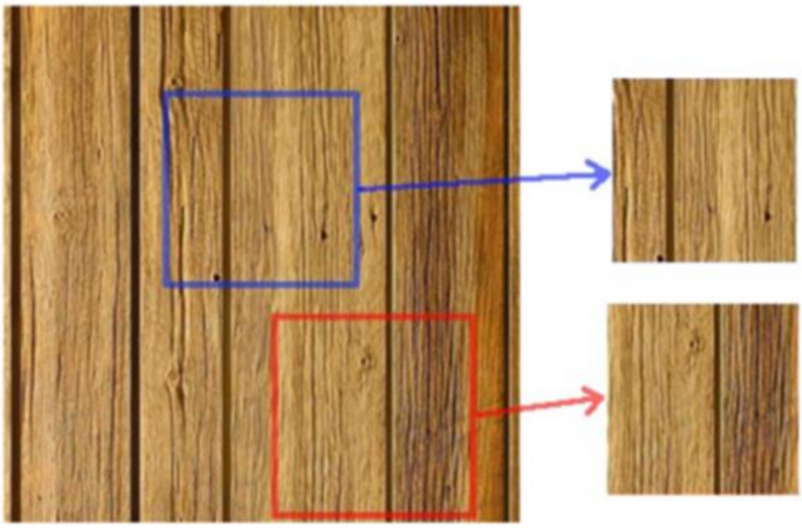
Figure 15 – Texture samples

Source: Adapted from Balas (2006)

There are some textures that contain solid periodicity, so called pseudoperiodic, once they possess even arrangement or distribution of rows (e.g. first two images on the top row, left to right in figure 15). Others are composed of repeated structural elements, but lack solid periodicity or comprehensive structure, for instance, the last image on top row, left to right (Fig. 15). Finally, some of them have asymmetric repetition and may produce lighting effects suggesting depth, such as the first image on the bottom row, left to right of figure 15 (BALAS, 2006). All those types of textures, precisely pseudoperiodic ones, structured ones, and asymmetric ones are handily observed on figure 15.

Figure 16 highlights the last image of figure 15, bottom row, left to right, showing that parts of the same image may have different classification, in agreement with their texture.

Figure 16 – Textures of different parts of the same image



Source: Adapted from Balas (2006)

Texture is universal in numerous natural images and plays a vital part as a standard visual sign for a wide assortment of image examination applications. Texture classification and representation are essential assignments in computational imaging, as well as for pattern recognition. It has been extensively utilized as a part of numerous applications including therapeutic image investigation, automated visual inspection, object recognition, remote detecting, and image recovery, among others (HADIZADEH, 2015).

Yet the texture is easy to be identified by human eyes, it is hard to be defined in mathematical terms. Haralick and his colleagues (1973) specifically define texture on a more rigorous way, considering it as a set of features extracted from spatial domain for a given probability distribution of grayscale on an image.

Classifying textures is a troublesome issue in pattern recognition and computer vision. The undertaking is to relate a class label to its corresponding texture division. In the last years, a diversity of texture description approaches have been proposed. These approaches can be separated into two classes, i.e. sparse and dense representations. The first one works by recognizing feature aspects, either taking into account interest point or dense inspecting points (OJALA; PIETIKAINEN; MAENPAA, 2002).

Then, feature description is performed on these inspecting points. The second system, dense representations, includes separating nearby elements for every pixel in a picture (KHANA et al., 2015).

Other two possible approaches to texture description are the structural and the statistical. In both cases, some requirements must be considered, namely invariance to position, scale and rotation. The main example of structural is the Fourier transform of the image. According to Haralick, Shanmugam and Dinstein (1973), the most usual statistical approach is the co-occurrence matrix, thanks to its best performance.

6.1 STATISTICAL TEXTURE DESCRIPTORS

Suppose a discretized image $I^{m.n} = [i_{x;y}]$ is assumed to be a Gaussian random field, where $i_{x;y}$ denotes the gray level of a pixel at location $x, y \in \mathbb{Z}$, with the quantized pixel $i_{x;y} < 2^8$, $i_{x;y} \in \mathbb{N}$.

The co-occurrence matrix contains elements, which are counts of the elements with the same brightness, according to certain distance and angle.

For n distinct gray level partitions, $b_i = \frac{2^8}{i}$, $i = 1, 2, \ldots n$

So, it is found the Spatial Gray Level Dependence - SGLD matrix = $[p_{i;j}]$

$$p_{i,j} = \sum_1^n \sum_1^n (i_{x,y} \in b_i) . (i_{x,y} \in b_j) \quad (3)$$

and the normalized form: $p_{i,j} = \frac{p_{i,j}}{| p_{i,j} |}$

Considering the averages u_x and u_y and standard deviations s_x and s_y:

$$u_x = \sum_i \sum_j p_{i,j} ;\ u_y = \sum_j \sum p_{i,j} \quad (4)$$

$$s_x^2 = \sum (1 - u_x)^2 \sum p_{i,j};\ s_y^2 = \sum (1 - u_y)^2 \sum p_{i,j} \quad (5)$$

A feature can be extracted from SGLD through its entropy, energy, maximum intensity, correlation, and inverse difference moment (HARALICK; SHAPIRO, 1991):

$$entropy = -\sum p_{i,j}\ log_2\, p_{i,j} \tag{6}$$

$$energy = \sum\sum p_{i,j}^2 \tag{7}$$

$$max = Max\ (p_{i,j}) \tag{8}$$

$$correlation = \sum\sum \frac{(1-u_x)(1-u_y)\, p_{i,j}}{s_x\, s_y} \tag{9}$$

$$IDM = \sum\sum \frac{p_{i,j}}{1+|\,i-j\,|} \tag{10}$$

It is notable that a pair of pixels is counted for both forward and backward directions when calculating the SGLD matrix.

7. MATERIALS AND METHODS

This section explains the methods used to acquire the data for this research. Also, it presents the materials, the techniques, and the measures employed to reach the results, indicating that the problem was based in a real-world situation.

The sciences are characterized by the use of scientific methodology, which is the discipline that teaches the way, the technical standards that must be followed in scientific research (LAKATOS; MARCONI, 2010).

The data for this research was gathered in a medium-sized *pinus* sawmill; as stated by the Brazilian Development Bank - BNDES, which classifies the company size according to its yearly incoming. The *Elliottii pinus,* species used in the research, after they have been deployed into boards, they went through the drying process (BNDES, 2011). The destination of those boards is both domestic and international market. The industry is located in Itapeva region, São Paulo State.

For the learning stage, a set of materials were used as camera, laser, artificial light, computer, and software as well. The programming language used was Scilab and it was written via partnership with USP-ICMC/UNESP, adjusted to this purpose.

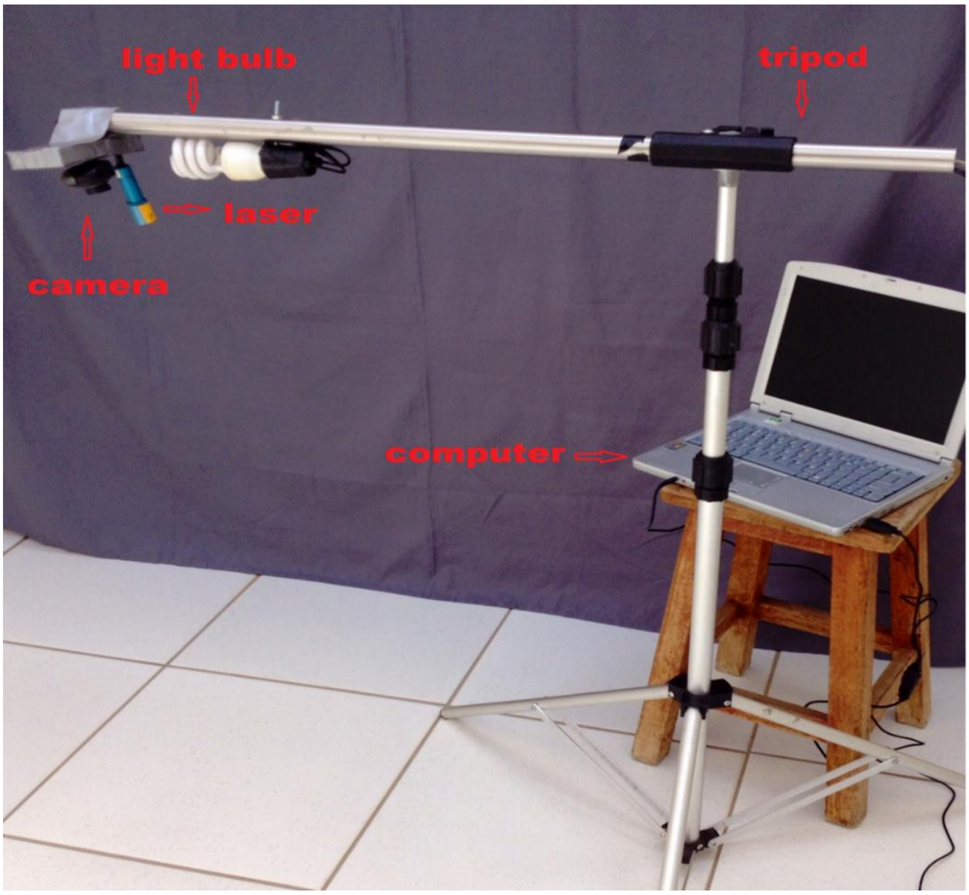
7.1 HARDWARE

The apparatus was constructed by the author to capture the specimens' images. A tripod with an aluminum tube (see Figure 17) was used as well as a camera, a laser beam and a CFL light bulb holder.

For illumination, "a CFL light bulb" (spiral type; 13W – 110-130V Duluxstar Twist by Osram) 40 cm distant from the object, with about 450 lumens artificial and about 1,880 lumens with the combination of natural and artificial, was used for standardization of the specimens lighting. The over-lighting issue happened all the time the natural luminosity was higher than 2,000 lumens, especially around noon. On those situations the artificial lights were switched off.

The photos were captured by the WebCam 1.3 Mega pixel Mobitechplus with Lens F=3.85 mm, Focus Range 30 mm~∞. The mentioned webcam proved to be satisfactory during the image acquisition, and it had the automatic focus turned off once the distances from the objects were always the same, i.e. 400 mm.

Figure 17 – Hardware apparatus for image acquisition



Source: Own Author

Also, a laser beam, model ZM05S9 by Z-laser, was used to help calculate the distances and the limits on the specimens.

The lumens were measured with Digital Lux Meter - MLM-1011 model by Minipa. During the image acquisition, the lumens range from 700 to 1,880 and registered the average of 850 lumens measured at the obtainment site.

7.2 IMAGE ACQUISITION

In this phase, 640 x 480 pixels images of pine wood were obtained. The specimens were often standardized in size (see fig. 18 and 19), with 600 cm in length, 15cm width and 2 cm height.

Figure 18 shows the stack of *Pinus* boards inside sawmill.

Figure 18 – Photo of sawn *Pinus* boards

Source: Own Author

Figure 19 presents collaborators within sawmill during their task of visually classify the *Pinus* boards. At this image is possible to realize the size of the specimens, described previously.

Figure 19 – Visual classification of *Pinus* boards

Source: Own Author

The total of 640 objects have been captured from three different qualities and saved in BITMAP format. However, some of them were discarded due to blurs and/or over-lighting, which made them difficult to understand clearly. The remaining images have qualities graded as A, B or C, according to expert workers. The amount of images were classified as follows: 144 quality A images; 177 quality B; and 53 quality C.

Thus, a total of 374 images were used for this experiment.

The entire database is available for consultation at “https://drive.google.com/drive/folders/0B3sfvwJAURw-fjllMmxoWW44X0RLem9DQktCVWsxSDB2MHVaZmI0M0N3RmVqMEk1ekNVZWM”.

Figure 20 shows an example of an original image, collected on a stand created alongside the visual classification area. This image contains the complete amount of information from the moment it was acquired; specifically it consists of the object itself with noises.

Figure 20 – Original image without any treatment

Source: Own Author

7.3 PRE-PROCESSING (IMAGE ENHANCEMENT)

A pre-processing phase was used to improve the image’s contrast and brightness characteristic, reduce its noise content, or sharpen its details. Some processes of image enhancement were used, as image filtering and image intensification. In this stage, the object, i.e. the whole picture, was cut in order to obtain only the wood image, eliminating the noise, and each specimens measure approximately 400mm length and 150mm width. (Fig. 21).

Figure 21 – Wood image enhanced

Source: Own Author

The images from three different kinds of wood quality (Quality A, B and C) were processed using the Scilab software. In the program, images were cut in standard sizes (630 x 180) selecting only the region of interest, i.e., wood and its possible defects (Fig. 22).

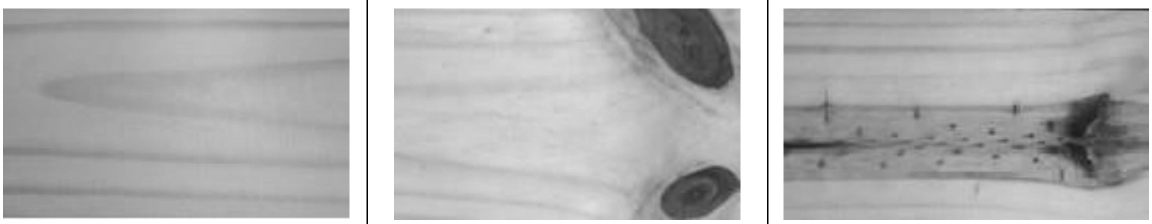
Figure 22 - Classification of the wood samples

 Quality A	 Quality B	 Quality C

Source: Own Author

After accomplish these processes, color image has been converted into gray scale level, which indicates the intensity of image pixels in 256 different shades of gray. This procedure was done to all the 374 images available on the database. The scale of gray ranged from “0”, representing the blackest or totally black to “255”, representing the whitest or pure white. (Fig. 23).

Figure 23 – Wood samples in gray scale



Quality A	Quality B	Quality C

Source: Own Author

After the images have been captured, they were divided into more categories. The Quality A group was the only one with just a category, once it was not necessary to distinguish their images, as seen on table 6.

Group B images were divided into B1 and B2. The first one contains images with knots ≤ 50mm x 50mm. The later, classified images with the presence of pith, not exposed, with length ≤ 150mm and width ≤ 2mm.

Group C was classified as C1, for images with knots ≥ 50mm X 50mm; C2 for images with pith ≥ 150mm in length and ≥ 2mm in width; C3 for those images with the presence of exposed pith, of any size; and C4 for the wood images containing blue stain.

Nevertheless, companies usually discard the specific classifications and adopt broader ones, as A, B and C, grouping the classes B1 and B2, as well as C1, C2, C3 and C4. This strategy was used to assist on the clear classification between classes, once the distinction could be only determined through narrow defect size differences.

Remembering that after the cutting was made, the pixels of the images have been modified to show only shades of gray, ranging on a scale from 0 (totally black) to 255 (pure white). Then, the gray scale processing was made for each pixel of the image and a histogram was originated.

Table 6 - Classification of sawn wood quality from conifers

Defects	Quality Classes A	B1	B2	C1	C2	C3	C4
Firm knots Knot clusters	Not allowed	≤ 50mm X 50mm	-	≥ 50mm X 50mm	-	-	-
Unsound Knot	Not allowed	Not allowed	≤ 30% $l_{1/m}$	≤ 200% $l_{1/m}$	Without restriction	Without restriction	-
Pith	Not allowed	Without restriction	≤ 2mm X ≤150mm	Without restriction	≥ 2mm X ≥150mm	-	-
Exposed Pith	Not allowed	Not allowed	Not allowed	Not allowed	Not allowed	Without restriction	-
Blue stain	Not allowed	Not allowed	Not allowed	Not allowed	Not allowed	Not allowed	Without restriction

Source: Adapted from ABNT NBR 11700

The histogram was divided into nine intervals. It was recorded how many images contained pixels with gray tones corresponding to it, in percentage, for those intervals. These data were inserted in Table 7.

Table 7 – Percentage of images in each interval

	0 - 85/3	85/3 - 170/3	170/3 – 85	85 - 340/3	340/3 – 425/3	425/3 - 170	170- 595/3	595/3 – 680/3	680/3 - 255
Quality A	0.72 %	4.32 %	15.83 %	69.78 %	92.81 %	99.28 %	100 %	52.52 %	5.04 %
Quality B	57.48 %	89.76 %	96.85 %	100 %	100 %	100 %	100 %	90.55 %	40.94 %
Quality C	61.11 %	80 %	90 %	92.22 %	97.78 %	100 %	100 %	90 %	27.78 %

Source: Own Author

The chi-square test to analyze the frequency of occurrence in each interval was used. The standardized residuals data are included in Table 8.

Table 8 – Standardized residuals (chi-square)

	0 - 85/3	85/3 - 170/3	170/3 - 85	85 - 340/3	340/3 – 425/3	425/3 - 170	170- 595/3	595/3 – 680/3	680/3 – 255
Quality A	-15.96	-17.63	-13.96	4.44	10.70	12.28	12.48	0.00	-8.99
Quality B	4.67	7.88	5.78	-1.23	-4.50	-5.45	-5.52	-0.62	6.84
Quality C	7.54	5.54	4.85	-2.17	-3.90	-3.90	-3.97	0.64	0.07

Source: Own Author

Standardization through chi-square benefits the understanding of certain factor occurrence, and in this example indicates that, with a confidence level of 95%, the values above 1.96 are considered occurrence excess; and values below - 1.96, absence of occurrence.

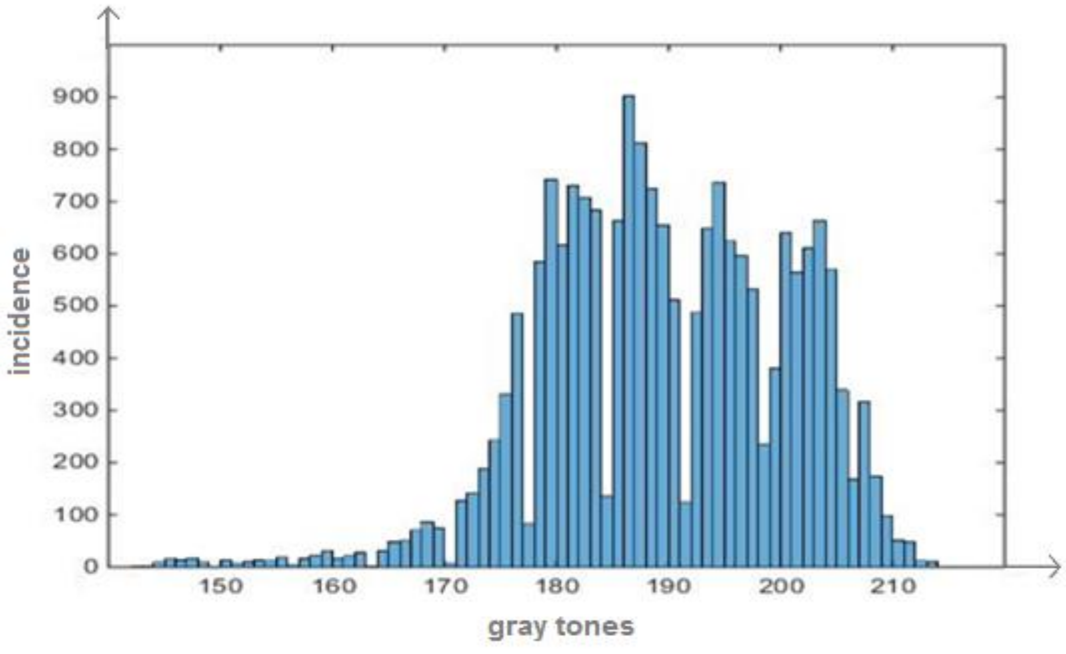
Therewith, it is notable the absence of occurrence of “Quality A” images for the first three intervals (indicated by the negative numbers, larger than -1.96, within those columns), which differs from “Quality B and C” images. The latter ones present excessive occurrence on the same intervals.

As noted, the behavior of grayscale pixels incidence from B-class and C-class are very much the same. The columns that indicate presence of pixels for class B, also demonstrate for class C. Whenever the columns show the absence of pixels for C-class specimens, the same happen for B-class, except in column 85 - 340/3, using chi-square.

This slight difference is unable to create a solid, stable inference to distinguish both classes from one another. In other words, the gray-scale technique for class classification fails on differentiating one or more classes.

Figures 24 to 26 show the histograms into graphics, for better understanding of the grayscale incidence behavior.

Figure 24 – Incidence of the image pixels with gray tones in each interval of A-class

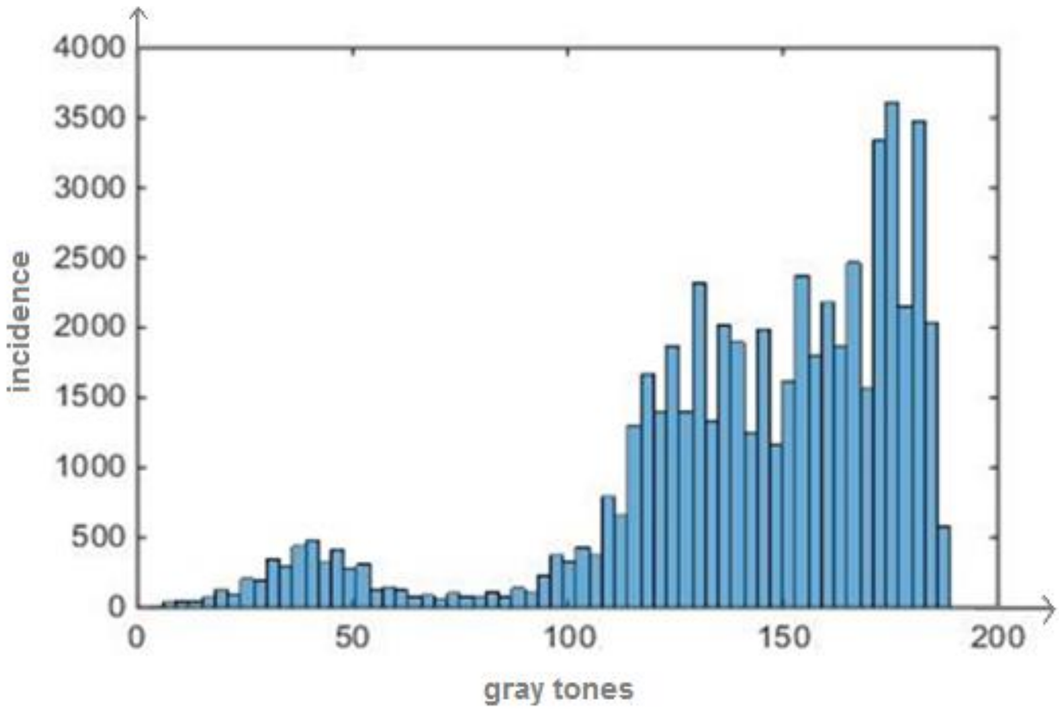


Source: Own Author

As it can observed on figure 24, the occurrence of gray levels for the A-class images indicates the presence of gray shades in some intervals, ranging from approximately 140 to 215. It also shows the absence of occurrence in the other intervals.

However, the obstacle in this type of classification could be effortlessly noticed after the standardization into chi-square, as seen on figures 25 (B-class) and 26 (C-class).

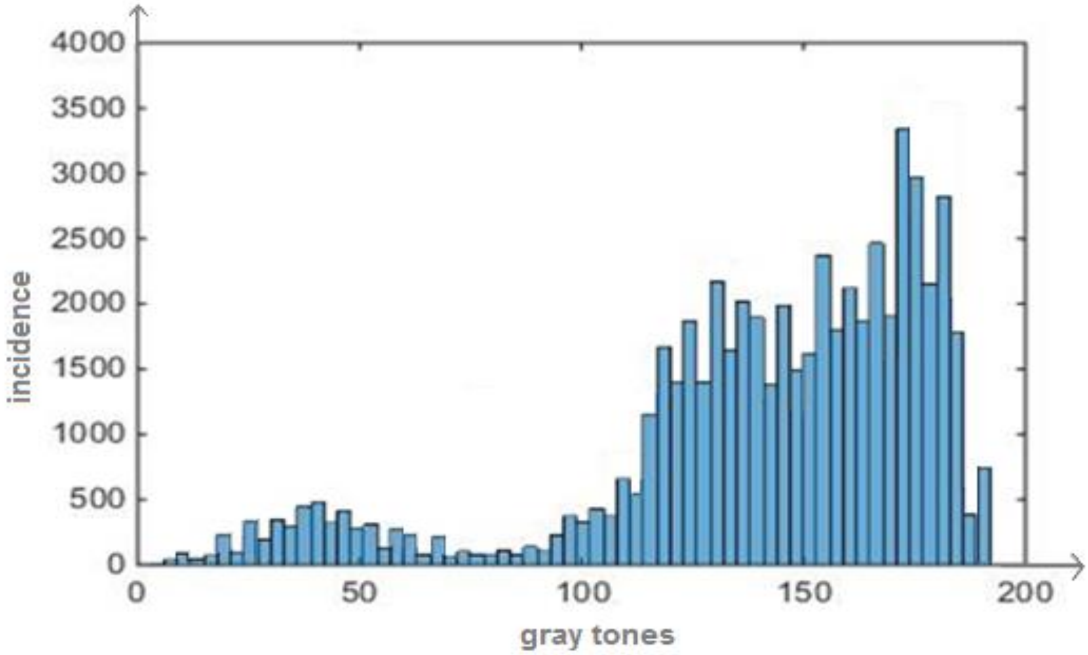
Figure 25 – Incidence of the image pixels with gray tones in each interval of B-class



Source: Own Author

One's can perceive that, at the end of the learning stage, it was quite clear to understand the differences from "A" quality wood boards, represented on the histogram from figure 24; from "B" or "C" ones, via histograms shown on figures 25 and 26. The curve design demonstrated on the histogram from images A quality present a single phenomenon. Images B and C quality show double phenomena, i.e. two curves on their histograms.

Figure 26 – Incidence of the image pixels with gray tones in each interval of C-class



Source: Own Author

As above, it was difficult to state that a given image was from “B” or “C” quality, since both classes developed very similar behavior, creating pretty much identical histograms. This fact can be observed just by comparing visually figure 26 with figure 26.

So, further investigations were necessary at this point. That is the reason the study continued on finding a more appropriate method for the wood image classification under investigation.

7.4 FURTHER METHODS FOR IMAGE CLASSIFICATION

Once the histograms could not determine whether a given wood image was from B or C quality, a new approach was tried. At this time, we used other machine learning techniques to find out the most effective method for wood classification from their images.

As usual, the first step is the image acquisition, which is the starting point for imaging and it plays a vital role. The choice of the combination between software and hardware will depend on the type of the image source and the interface.

As affirmed by Nagabhushana (2006) there is a large number of options, involving video or digital cameras, scanners, etc. Once video cameras lack on quality enhancement due to the speed of capture, providing pixel noises, uniformity and dynamic range; digital ones are able to integrate signals for extended length of time and less noise, besides increased dynamic range.

Therefore, it is worth remembering that the database created earlier through the acquisition, enhancement, segmentation, extraction of features and storage for further classification of the images was used, although new series of actions to achieve the result, i.e., new processes were tested this time.

7.4.1 Acquisition Equipment

Although the scanner image is more appropriated for better image reasoning, due to the uniform lighting, in the on-line application stage at the sawmills, it is extremely heavy. A feasible choice is the use of cameras system, as imaging devices.

As it was observed, this would be closer to the expected real-life approach. Then again, giving sufficient and even lighting for the cameras in the intended applications was considered an issue, once it might be used within environments with natural and artificial lighting acting at the same time, or individually.

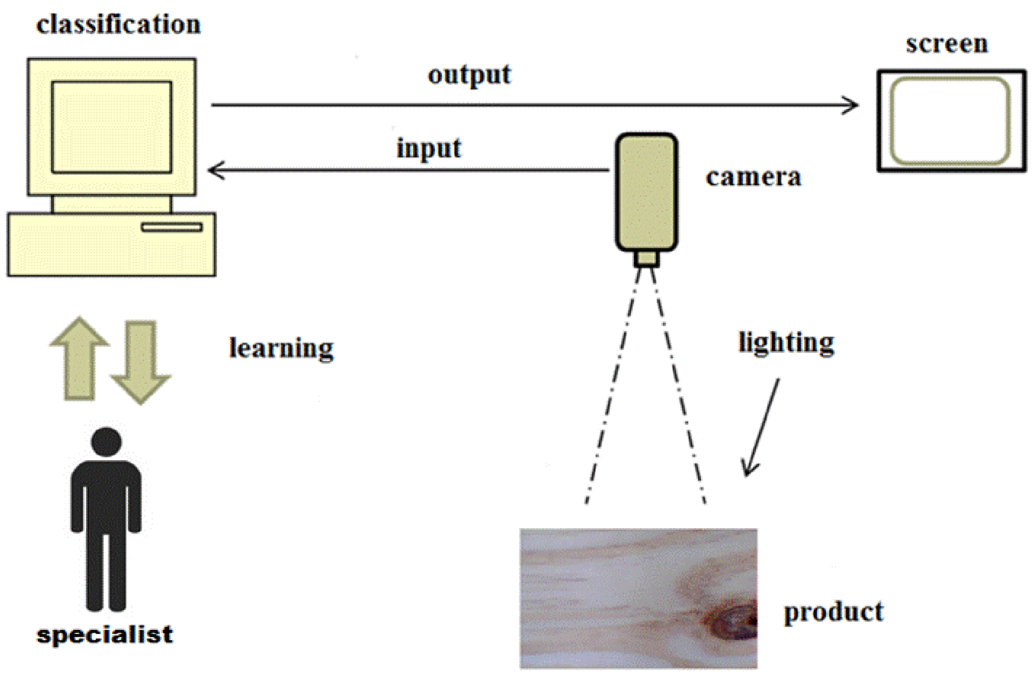
The correct lumens become a challenge since the natural light cannot be controlled. Yet, the digital camera is discovered sufficient for the reasons of image acquisition, because it produces achievable basis for constant applications in the industry processes (ÖSTERBERG, 2009).

The same author cited above explain that cameras are usually employed to acquire images from objects that need to be inspected or classified. Then, the next step is to find areas of interest within the images and classify each one of them through software routines. It is valid to keep in mind that the hardware also plays equivalent importance when dealing with imaging, to process them appropriately.

Pham and Alcock (2003) declare that a common system has some elements, such as computer with a display/monitor, camera, lighting equipment over a conveyor belt, which transports the products to be photographed. After the information is processed, the classification, either rejected or accepted, is displayed on the monitor.

Figure 27 shows exactly that, and this procedure was used to acquire the images. Next they were sorted by specialists, i.e. human graders, who were employees at the sawmill company, hired for this purpose, visual classification of the wood quality.

Figure 27 – Classification through a specialist opinion



Source: Own Author

There are some stages for images processing, after they have been acquired. Abdullah et al. (2007) suggest four steps: image acquisition, pre-processing, segmentation and feature analysis.

Pham and Alcock’s opinion (2003) involves extra steps on processing images, as presented on Table 9. They state that once the image is acquired, it needs to be enhanced, in order to improve its quality. It means that anything else, other than the point of interest of the image needs to be eliminated or filtered, and/or the interest area of the image receives an improvement, through lighting enhancement or blur elimination. Next, it needs to be segmented, and have its parameters calculated, in a stage named Feature Extraction. Only then, the image will be classified, either approved or not approved.

After the image acquisition, the process of enhancement initiated with the help of Scilab software. In this stage, the images needed to be uniform and a process of standardization took place through the referred operating system.

The image processing system operates as follow. The features are extracted from color images by treating each channel of color image (Red-Yellow-Green) as a monochrome image and transforming its shape information into pixels surfaces.

Table 9 - Stages of Image Processing

Image Acquisition	It is used to obtain the image to be processed from a given object.
Image Enhancement	It is necessary to raise the quality of the object image for further process.
Segmentation	In order to divide the image into specific areas of interest.
Feature Extraction	Technique used to calculate the parameters rates or measures, which differentiate each object from one another.
Classification	It is appropriate for the determination of what each object represents.

Source: Adapted from Pham and Alcock (2003)

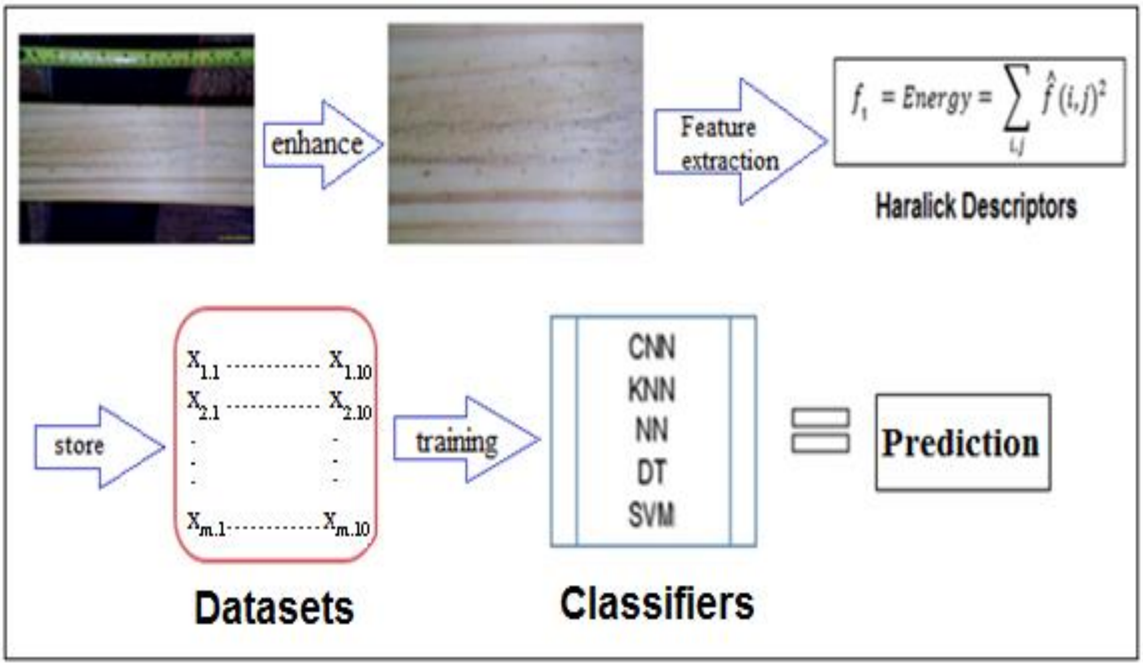
In the second step, a matrix is created with the numeric values of color intensity corresponding to each pixel (between 0 and 255). Later, it performs normalization into the numerical matrix, where it is assigned a unit value for pixels maximum intensity and zero to minimal intensity. These features will be the input for the CNN algorithm.

Each example $e = (x, yb)$ consists of a tuple of p attributes values $x = (x_1, \ldots . x_p)$, where p is the number of features of the image, and a scalar label yb ∈{-1, 0, 1}, corresponding respectively to qualities A, B and C. The set of all 374 examples constitutes a dataset, named here as A-B-C. Another dataset is generated from the same features, but, instead of considering classes A, B and C, the classes B and C are aggregated. Therefore, this dataset has only two classes and it is named A-BC. This is done because some companies or experts may be interested in classifying wood on approved boards (class A) and not approved boards (classes B and C)

7.5 SOFTWARES

Java programming language was used for image acquisition and for selecting the wood image free of noise. In other words, the mentioned software helped on the enhancement task, eliminating parts of the image not interesting for the process, as the background (Fig. 28).

Figure 28 - Machine Learning system used with texture descriptors



Source: Own Author

Then Scilab software, free for download, was applied on the enhanced image, for feature extraction, i.e., the creation of vectors on Haralick Descriptors stage, and later to store in datasets. The classification techniques, after the training period, employed in this thesis were from Weka - Waikato Environment for Knowledge Analysis software, which is free and is a collection of machine learning algorithms for data mining tasks. Weka was able to achieve final predictions.

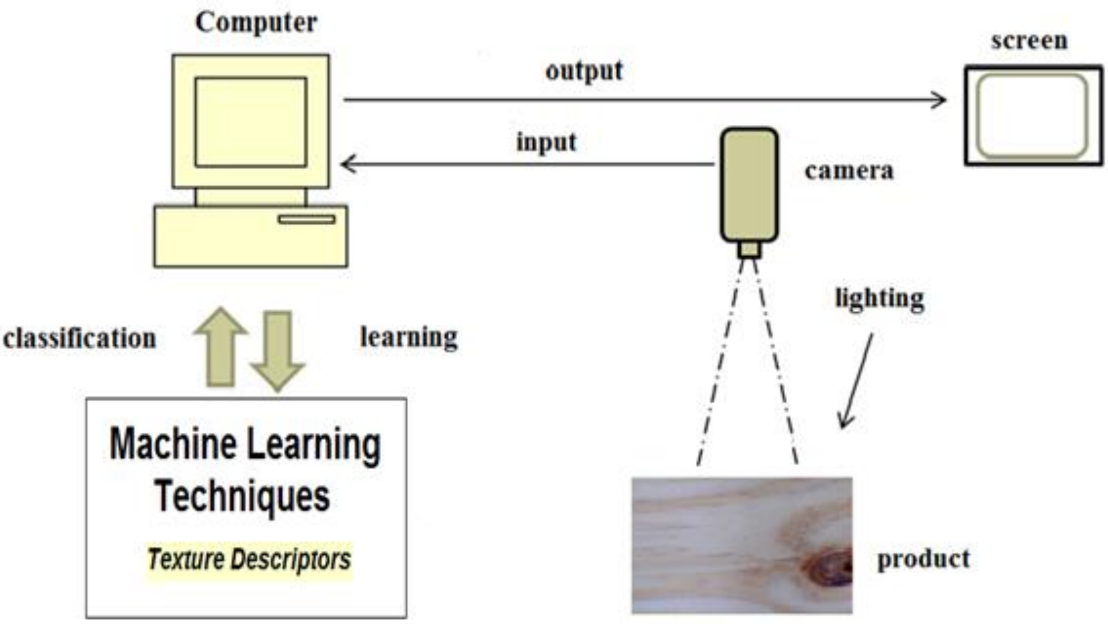
As noted, after the image is acquired, it received an enhancement through Java software in order to eliminate noise. Then, Haralick Descriptors is used with Scilab, in order to extract features. At figure 28, only one type of descriptor is shown, i.e. energy, but during the real process, others were used as well, such as IDM, entropy, correlation, etc. Scilab also organizes the objects into the datasets with a sequence of ten features $(x_{1.1} \ldots\ldots x_{1.10})$; each one of the sequences representing a specific image. Now, Weka is in charge of the classification. Once one determines the classifier, the software presents the full values of each algorithm, providing an output, or prediction, which is the accuracy and kappa result.

7.6 METHODS

In this section, it is presented the experiments carried out to evaluate the image classification performance for CNN and compare it to other Machine Learning (ML) algorithms for the classification of quality of wood boards.

It is distinct from a human classification, because it uses Machine Learning techniques to analyze and reason the images classification. This can be seen on Figure 29.

Figure 29 – Classification through Machine Learning



Source: Own Author

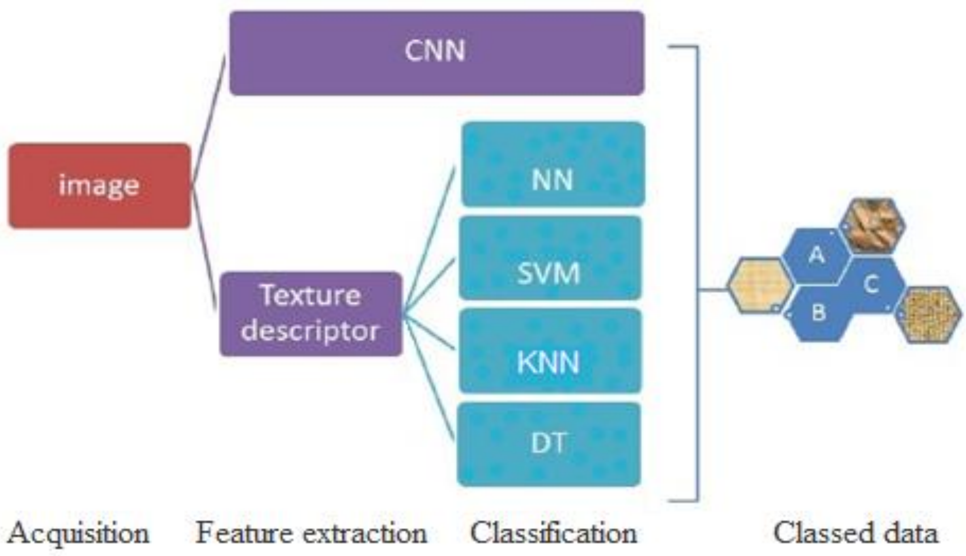
Using the machine learning models associated with the texture descriptor was to find out the most appropriate strategy for sorting wood images, without the need of human visual classification. The principal idea was to discover the most effective on accuracy, but also worrying with computational cost.

The methods used were classification through Convolutional Neural Network, and other traditional machine learning techniques combined with Texture Descriptors - TD, such as Support Vector Machines (SVM), Decision Trees (DT), K-nearest Neighbors (KNN) and Neural Networks (NN).

The texture descriptors listed above define these images properly for these algorithms. In this thesis, the images were described using Haralick's texture features since they have been widely and successfully employed for this purpose. For such, a Spatial Gray Level Dependence Matrix (SGLD) from the image and Haralick's texture features are computed from the SGLD.

The main difference of CNN is that it is able to extract features from the raw image pixels while more traditional classification algorithms expect as input a set of features describing an image. Therefore, in this thesis, Haralick texture descriptors were also used to pre-process an image before training a model by different Machine Learning (ML) algorithms, namely KNN, NN, DT and SVM. This strategy is shown on Figure 30.

Figure 30 – Strategy used to classify wood images



Source: Own Author

As discussed earlier, the hypothesis is that some or all these traditional ML algorithms will generate more accurate models than CNN due to the characteristics of this problem, namely small number of instances and well defined domain. CNN is known as an architecture that helps on solving problems that require complex highly varying functions, and it usually involves non-labeled problems with very large dataset.

7.6.1 Assessment measures

When the topic is Machine Learning, several different measures have been used to evaluate the performance of predictive models in classification issues (ROSSI, 2014). To affirm which one is the most appropriated model, one must take into consideration that it may vary with the measures used in the evaluation. Several studies have investigated and described measures commonly used by Machine Learning community, as Fawcett (2003), Ferri et al. (2009) and Japkowicz and Shah (2011). However, the most appropriate measure depends primarily on the characteristic issues under analysis and the final assessment goal.

In classification problems, the error rate of a model for a test set with m examples is defined by Equation 14, where $\check{y}_i$ is the predicted class, y_i is the observed class; and L is a function of loss or function of cost that is defined by Equation 15.

$$Error\ rate = \frac{1}{m} \sum_{i=1}^{m} L\ (y_i\,, \check{y}_i) \tag{11}$$

$$L\ (\check{y}, y) = \begin{cases} 0,\ if\ \check{y}_i = y_i \\ 1,\ if\ \check{y}_i\ \neq\ y_i \end{cases} \tag{12}$$

Other measures to assess the performance can be calculated from a contingency table with absolute frequencies, known as confusion matrix, shown in table 11 to a problem of two classes: A class and BC class, for instance. In this table 11, the lines indicate the true class and columns indicate the predicted class. The TA, TBC, FA and FBC acronyms correspond to:

- True A class (TA): Total classified as A class examples and really are A;
- True BC class (TBC): Total classified as BC class examples and really are BC;
- False A class (FA): Total classified as A class examples but are really BC;
- False BC class (FBC): Total classified as BC class examples but are really A.

The total true examples of A class and BC class are given by Y_A and Y_{BC}, respectively, and the total predicted examples as A class and BC class are given by $\hat{Y}_A$ and $\hat{Y}_{BC}$. The value sum of the main diagonal of Table 10 represents the total number of correctly classified examples; and the value sum of the secondary diagonal represents the total number of incorrectly classified examples.

Table 10: Confusion matrix for a problem of two classes: A class and BC class

		PREDICTED CLASS		
		A class	BC class	Total
TRUE CLASS	A class	TA	FBC	$Y_A = TA + FBC$
	BC class	FA	TBC	$Y_{BC} = FA + TBC$
	Total	$\hat{Y}_A = TA + FBC$	$\hat{Y}_{BC} = FA + TBC$	m

Source: Adapted from Rossi (2014)

The measures used to assess the efficiency of the machine learning techniques in this thesis were accuracy and Cohen's kappa statistic. The first one was indicated in percentage and represents the capability of the machine to predict the correct result. In other words, accuracy attributes the nearness of a measured value to an established or standard value (NANNI; MELUCCI, 2016).

The accuracy is defined by Equation 16 from Table 11 (JAPKOWICZ; SHAW, 2011).

$$Accuracy = \frac{TA+TBC}{m} \tag{13}$$

The second one, Cohen's kappa statistic (k), means the capacity of a correct hunch compared to a random guess, and it is used as another measure in which the unbalance classes' problem, i.e., when there is a considerable difference in the number of samples belonging to each class, is minimized, since it considers the distribution of classes in calculation. This measure is defined according to Equation 17, where p_0 is the accuracy of the classifier and p_c is the probability due to chance, which can be obtained from the confusion matrix. Where m is the total number of samples, p_c is defined by Equation 18.

The value of κ ranges from -1 to 1, where -1 means the total disagreement, 0 is a random classifier and 1 full agreement (ROSSI, 2014).

$$k = \frac{p_o - p_c}{1 - p_c} \tag{14}$$

$$p_c = \frac{Y_A}{m} \text{ x } \frac{\check{Y}_A}{m} + \frac{Y_{AB}}{m} \text{ x } \frac{\check{Y}_{AB}}{m} \tag{15}$$

Cohen's kappa statistic helps to determine if the scheme output is better than random allocation (FLESIA et al., 2013).

Both accuracy and Cohen's kappa were measured for the assessment of the Machine Learning techniques results in this thesis.

8. ANALYZES AND DISCUSSION

This section describes the results of the experiments on wood boards image classification. It begins with the classification using Deep Learning, via Convolutional Neural Network – CNN. Then, different types of classifiers, called traditional machine learning techniques, were used to verify the most accurateness of image sorting. Texture Descriptors were also used with each one of them, namely Neural Networks - NN, K-Nearest Neighbors - KNN, Support Vector Machines - SVM, and Decision Trees - DT.

By the end of this section, the performance of the recognition tasks via Deep Learning and traditional Machine Learning with Texture Descriptors were evaluated, indicating the most appropriated for pine wood boards.

8.1 CLASSIFICATION WITH CONVOLUTIONAL NEURAL NETWORKS

Some computational experiments were performed to optimize the CNN topology as shown on Table 11 with three different orders for the A-B-C dataset. The topology is defined by the number of feature maps, patch width, patch height, pool width and pool height, respectively, and separated by a hyphen. That's the reason of five numbers in the sequences. Besides this parameter, the number of maximum training iterations (epochs) was set to 100.

Table 11: Convolutional Neural Network parameters and performance

Order	Topology	Accuracy	k
1	20-3-3-2-2, 25-3-3-2-2	0.7766	0.6201
2	**20-5-5-2-2, 25-5-5-2-2**	**0.8191**	**0.6876**
3	20-10-10-4-4, 25-10-10-4-4	0.8085	0.6656

Source: Own Author

As seen on table 11, the topology with best performance was order 2 (in bold), with higher accuracy and Cohen's kappa.

Two databases were tested, one with images with resolution 96 x 27 and the other with 472 x 135 resolution. The worst result with higher resolution pictures indicates an overfitting problem, i.e. an excess of non-relevant input information to the network.

It is important on classification problems, to compare the class tendencies errors, as presented on Table 12.

Table 12. CNN Confusion Matrix: Classes A, B and C

real / predicted	A	B	C
A	**30**	7	1
B	6	**39**	1
C	1	7	**2**

Source: Own Author

In considerations of comparing the class tendencies errors, the CNN was running over 2 dataset configurations, namely clusters A-B-C, and clusters A-BC. The confusion matrix for the 2-class configuration A-BC is presented on Table 13, and in this case, value found was k = 0.650 and accuracy = 83.0 %.

Table 13. CNN Confusion Matrix: Classes A and BC

real / predicted	A	BC
A	**31**	7
BC	9	**47**

Source: Own Author

These experiments were performed with CNN, with its trainable multi-layer model.

8.2 CLASSIFICATION WITH TEXTURE DESCRIPTORS

On the opposite of the CNN, the features extracted directly from the images are not appropriate for traditional Machine Learning algorithms, such as NN, as explained earlier (section 5.1).

Figure 31 presents an example of Inverse Difference Moment - IDM, which is one of the features with best outcome, extracted from Spatial Gray Level Dependence - SGLD, for all instances in the dataset.

The inputs to the function are:

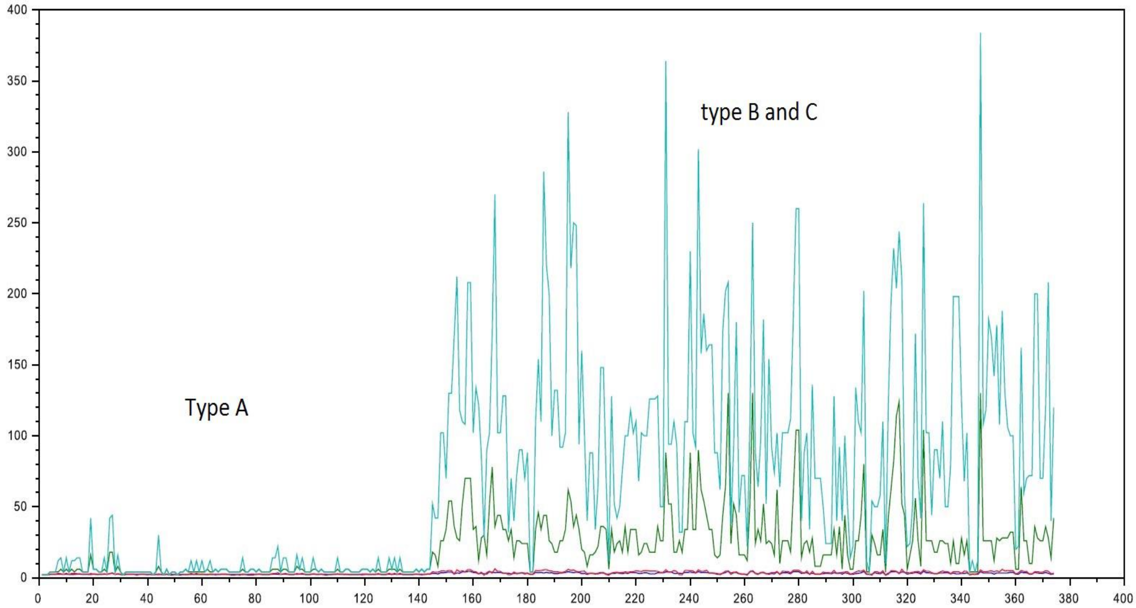
Image - the input gray scale image;

Input bits - the gray level resolution of the input image, according to equation 6;

Distance between pixels for calculating the SGLD matrix - it can take values between 0 and the minimum size of image and;

Theta - the angle, at which the SGLD matrix is calculated, which can take on values of 0º, 45º, 90º or 135º (degrees).

Figure 31 - Haralick descriptor IDM for classes type A, B and C



Source: Own Author

The outputs are either the SGLD matrix, as well as the energy, entropy, maximum probability, correlation, or IDM. Therefore, the dataset consists of the five Haralick's texture features extract from the images, specifically: entropy, energy, maximum intensity, correlation, IDM, and the target (class A, B or C).

Thus, each example using texture descriptor, e = (x,yb), consists of the five Haralick's texture features extracted from the images, x={entropy, energy, maximum intensity, correlation, IDM} and the target, yb ∈ {A,B,C}. e = (x,yb). Just as for the CNN, another dataset is created using the same features, x, but considering only two classes: A and BC, where BC is the aggregation of classes B and C.

As it can be verified on the Table 14, Haralick descriptors had a good performance to extract features from the image dataset and to cluster it on approved parts (class A) and not approved parts (class B and C). This table also shows the average value for each characteristic extracted and the standard deviation.

Table 14: Haralick descriptor for test and training dataset

Class	entropy	energy	intensity	correlation	IDM
A	0.51 ± 0.11	1.31 ± 0.27	0.63 ± 0.12	2.30 ± 0.16	4.43 ± 1.37
B	0.48 ± 0.07	1.51 ± 0.23	0.61 ± 0.09	3.19 ± 0.64	34.58 ± 28.96
C	0.38 ± 0.10	1.87 ± 0.38	0.50 ± 0.12	3.22 ± 0.66	26.60 ± 21.08

Source: Own Author

Considering the IDM Texture descriptor, for instance, it is possible to verify that image type-A forms a cluster; however, its descriptor is not enough to classify the images type B and C.

8.3 OTHER CLASSIFICATION TECHNIQUES

Concerning the process phase, after the images were captured, enhanced, and have their features extracted, they were put into classification.

The classification techniques employed in this research were used from Weka, which is a collection of machine learning algorithms for data mining tasks. The algorithms used are Neural Networks - NN via Multilayer Perceptron – MLP, K-Nearest Neighbor - KNN, Support Vector Machine - SVM, Convolutional Neural Networks- CNN, and Decision Tress - DT.

8.3.1 Classification with Neural Networks

The choice of MLP, an Artificial Neural Network model, parameters were: initial learning rate $\eta \in [0.5; 0.7]$; the stop criteria was maximum 100 epochs; bath size =100; momentum $\in [0.5; 0.9]$, and two hidden layers. The configuration of the number of neurons and layers is accomplished by choosing the architecture that has the lowest permissible error or running time. Some computational experiments were done to optimize the NN parameters optimization, such as bath size, learning rate, and number of hidden units, as shown on Table 15.

Table 15. MLP parameters and predictive performance

η	momentum	topology	k	accuracy	run time
0.7	0.7	4-12	0.7475	0.8617	3.57
0.7	0.7	4-10	0.7475	0.8617	1.96
0.7	0.7	4-8	0.7475	0.8617	1.88
0.7	0.7	4-12	0.7475	0.8617	2.88
0.5	0.7	4-12	0.7475	0.8617	2.38
0.9	0.7	4-12	0.7475	0.8617	2.17
0.7	0.5	4-12	0.7475	0.8617	2.62
0.7	0.9	4-12	0.6470	0.7950	2.60

Source: Own Author

It is imperative to compare the class tendencies errors based on confusion matrix, for the case of NN, as given on Table 16.

Table 16. NN Confusion Matrix: Classes A, B, and C

real / predicted	A	B	C
A	**37**	1	0
B	2	**44**	0
C	2	8	**0**

Source: Own Author

As it can be seen on Table 16, the Neural Networks have not classified any samples as class C. This might be for the reason that the samples are pretty much alike B-class samples, and/or it would be necessary a classifier with higher dimensionality.

The confusion matrix for the 2-class configuration A-BC is granted on Table 17, and the value found was $k = 1$ and accuracy = 100.0 %.

Table 17. NN Confusion Matrix: Classes A, and BC

real / predicted	A	BC
A	**38**	0
BC	0	**56**

Source: Own Author

On the other hand, the predicting of A –BC classes achieved 100%, classifying all samples correctly. It might indicate that B and C are similar.

8.3.2 Classification with K-Nearest Neighbors

KNN method is among the simplest machine learning algorithm. It is used to classify objects based on the closest training examples in dataset feature space.

In this procedure, the unlabeled query point is classified by votes among the label of its K-nearest neighbors. After some trials, it was determined that “k” (from the k-nearest neighbors) would be equal to three (k = 3), i.e. the number of selected neighbors for this test.

In this case, the value found was k = 0.6434 and accuracy = 79.8 %, according to Table 18.

Table 18. KNN Confusion Matrix: Classes A, B, and C

real / predicted	A	B	C
A	**33**	5	0
B	2	**40**	4
C	0	8	**2**

Source: Own Author

Once more, the critical part of KNN, for differentiating among 3 classes, was the errors in the classification of samples of class C, due to dimensionality of the classifier, and to the fact that the samples, automatically chosen by the classifier, might have been selected in an unsuccessful way.

The confusion matrix for the 2-class configuration A-BC is shown on Table 19. On the contrary of the previous classification of KNN, separating the classes among themselves, the value found was k = 1 and accuracy = 100.0 %, i.e., it had no misclassification between the two classes.

Table 19. KNN Confusion Matrix: Classes A, and BC

real / predicted	A	BC
A	**38**	0
BC	0	**56**

Source: Own Author

As noticed for this case, the KNN for 2-class prediction reached excellence in results. It could mean that B-class samples were considerably equivalent to C-class ones for this technique.

8.3.3 Classification with Support Vector Machines

Support Vector Machines - SVM is a learning technique based on Statistical Learning Theory (VAPNIK, 1995). This technique is robust to high-dimensional data and has been shown high generalization ability in different domains (STATNIKOV; WANG; ALIFERES, 2008).

The principle of SVMs in classification problems is to find an optimum hyperplane that satisfactory split the input data. According to Haykin (2001) the optimum hyperplane is defined such that the separation margin among classes is maximized. The support vectors used by SVMs are examples that fit the decision surface. Therefore, these examples are more difficult to classify and directly influence the decision boundary.

After a running time of 0.01s, the SVM model mapped the example space and returns a classification, which achieves the figures; k = 0.7694 and accuracy = 87.2 %, according to Table 20.

Table 20. SVM Confusion Matrix: Classes A, B, and C

real / predicted	A	B	C
A	**38**	0	0
B	3	**43**	0
C	2	7	**1**

Source: Own Author

As observed with other traditional machine learning techniques (NN and KNN), the classification with SVM also brought haziness in predicting class C. Despite the fact that there has been an improvement, as it classified one correct sample, it could have had mischance when grouping the samples.

The confusion matrix for the 2-class configuration A-BC is given on Table 21, where k = 0.9780 and accuracy = 98.9 %.

Table 21. SVM Confusion Matrix: Classes A, and BC

real / predicted	A	BC
A	**38**	0
BC	1	**55**

Source: Own Author

Otherwise, SVM delivered 98.9% accurateness for sorting only 2 classes. It might indicates that, once again, the classes B and C were considered pretty much alike for this machine learning technique.

8.3.4 Classification with Decision Trees

A decision tree is a classifier expressed as a recursive partition of the instance space and has been the most widely used approach to represent classification models, mainly due to its comprehensible nature.

Quinlan (1986) mentions that the structure of a decision tree consists of nodes that have exactly one incoming edge, except the root node, which is in the top of the tree. A node with outgoing edges is called an internal or test node. Each internal node splits the instance space into two or more sub-spaces according to an attribute. All other nodes are called leaves. Each leaf represents the most appropriate class.

Instances are classified starting at the root and goes down to a leaf node, according to the outcome of the tests along the path (ROCACK; MAINON, 2005).

The J48 algorithm available in Weka was employed in this paper to generate the tree model. This is a version of the C4.5 decision tree algorithm proposed by Quinlan (1993). This model achieved k = 0.734 and accuracy = 85.11 %, according to Table 22.

Table 22. DT Confusion Matrix: Classes A, B, and C

real / predicted	A	B	C
A	**36**	2	0
B	3	**42**	1
C	2	6	**2**

Source: Own Author

It seems that for the classification with traditional learning machines boosted with texture descriptors, the same difficulty was found among the techniques. They all fail to predict class C in a proper way, as noticed in table 22 with Decision Trees. Another time, this event might be due to the method of grouping chosen for this research, separating 70% of the samples for training, and 30% of them for testing.

With only 2 rules, shown on Table 23, k = 0.955 and accuracy = 97.9 %.

Table 23. DT Confusion Matrix: Classes A, and BC

real / predicted	A	BC
A	**36**	2
BC	0	**56**

Source: Own Author

In the other way, the 2-class classification (A-BC) scored 100% in accurateness.

8.4 OVERALL RESULTS

So far, we have compared two strategies, i.e. one using Texture Descriptor as feature extractor, associated with KNN, NN, SVM and DT as regressor techniques; and other directly with CNN without taking any pre-processor.

In order to provide a more comprehensive comparison between these strategies, Tables 24 and 25 exhibit our main results, demonstrating the differences in an order way, from lowest to highest accuracy. Each box summarizes the accuracy values from each method for the validation dataset.

Table 24. Dataset: A-B-C average accuracy by classification process

Model		*k*	Accuracy
CNN		0.563	0.755
Texture Descriptor	KNN	0.642	0.798
	DT	0.734	0.851
	NN	0.748	0.862
	SVM	**0.770**	**0.872**

Source: Own Author

Considering the data summarized on the table 24, it is possible to conclude that the accuracy presented by the Convolutional Neural Networks - CNN, with grayscale pixels, was the lowest among other Machine Learning techniques, with 75.5% accuracy, and Cohen's kappa 0.563.

On the other hand, the Machine Learning technique that provided the most accuracy for this process, which is the classification of the quality of *Pinus* boards, was the Texture Descriptors through Support Vector Machine - SVM, with 87.2% accuracy, and Cohen's Kappa 0.770.

Table 25. Dataset: A-BC average accuracy by classification process

Model		*k*	Accuracy
CNN		0.650	0.830
Texture Descriptor	DT	0.955	0.979
	SVM	0.978	0.989
	NN	**1.000**	**1.000**
	KNN	**1.000**	**1.000**

Source: Own Author

In view of compiled data from Table 25, it is suitable to deduce that the accuracy found via CNN was also the lowest among other techniques, with 83% accuracy, and Cohen's kappa of 0.650. Nevertheless, the ML technique that provided the most accuracy for this wood quality classification was the Texture Descriptors through NN and KNN, both with 100% accuracy, and consequently, Cohen’s kappa of 1.000.

9. CONCLUSIONS

The main contribution of this research was the finding that Deep Learning methods, as Convolutional Neural Network, greatly and successfully used for high-complexity problems with large databases, did not have similar performance with lower well-behaved images from the dataset under analysis, which means with less complexity issues.

On the matter of the investigation, this work has been centered on the image based assessment techniques, which can be connected all through the quality chain, and that essentially classify the wood quality.

In this thesis, a comprehensive evaluation with a Convolutional Neural Network compared to traditional Machine Learning techniques using Texture Descriptors were extended and performed to classify wood board samples.

In order to evaluate this method, experiments were carried out using a dataset taken from a real-world problem. Empirical results for the image dataset showed that the texture descriptor method proposed, regardless of the strategy employed, is very competitive when compared with Convolutional Neural Networks for all performed experiments.

The objective of this study was to compare Convolutional Neural Network, a deep learning method, for wood quality classification to other traditional Machine Learning techniques, namely Support Vector Machine (SVM), Decision Trees (DT), K-Nearest Neighbors (KNN), and Neural Networks (NN), associated with Texture Descriptors. The traditional techniques were able to improve the general performance of the classification systems for the problems under analysis.

Considering the dataset and the experiments achieved during this research, the Machine Learning technique boosted with Texture Descriptors with higher performance was Support Vector Machine, when classifying wood qualities from A-B-C, which is 3-class quality.

When the datasets were split in only two classes, i.e., wood qualities A-BC, both K-Nearest Neighbors and Neural Networks were highlighted, with no mistakes achieved during tests.

The best accomplishment of the Texture Descriptor (TD) method could be caused by the nature of the image dataset, once it was captured from well-behaved images, such as wood board samples.

In cases of higher dimensionality, as face or handwritten recognition system, there are plenty examples on literature pointing out CNN as the best performance solution.

Some suggestions for future works are listed from this point. An option could be the attempt to change the method for sampling the image, so that it could have different representation. A logical suggestion found on literature is the 10-fold statistic method, a cross-validation technique to evaluate the generalizability of a model, from a dataset.

Another possibility is to adopt features selection algorithms, such as reliefF, after features extraction, for the texture descriptors; once this research has used them randomly.

Finally, an additional alternative could be the increase of the amount of images in the dataset, with an even distribution of quality samples, especially during the training period.

BIBLIOGRAPHY

ABDULLAH, A. ISMAIL, N. K. N. KADIR, T. A. A., ZAIN, J. M., JUSOH, N. A., ALI, N. M. **Agar Wood Grade Determination System Using Image Processing Technique.** Proceedings of the International Conference on Electrical Engineering and Informatics Institut Teknologi Bandung, Indonesia June 17-19, 2007.

AFFONSO, C.; SASSI, R. J.; BARREIROS, R. M. Biological image classification using rough-fuzzy artificial neural network. **Expert Systems with Applications**, Vol. 42, Issue 24, p. 9482-9488, 2015.

AL-ALLAF, O. N. A.; ABDALKADER,S. A. A.; TAMIMI, A. Pattern Recognition Neural Network for Improving the Performance of Iris Recognition System. **Int'l Journal of Scientific & Engineering Research**, Vol. 4, Issue 6, p. 661-667, 2013.

ALAVI, A. H.; GANDOMI, A. H.; LARY, D. J. Progress of machine learning in geosciences: Preface. **Geoscience Frontiers**. Vol. 7, Issue 1, p. 1-2, 2016.

ALI, J. M.; HUSSAIN, M. A.; TADE, M. O.; ZHANG, J. Artificial Intelligence techniques applied as estimator in chemical process systems – A literature survey. **Expert Systems with Applications** 42, p. 5915–5931, 2015.

ALMEIDA, O. C. P**. Classificação de Tábuas de Madeira usando Processamento de Imagens Digitais e Aprendizado de Máquina**. Universidade Estadual Paulista "Júlio de Mesquita Filho" Faculdade de Ciências Agronômicas - Campus Botucatu. Botucatu, 2014.

AREL, I.; ROSE, D. C.; KARNOWSKI, T. P. **Deep Machine Learning -** A New Frontier in Artificial Intelligence Research. IEEE Computational Intelligence Magazine, 2010.

ABIMCI - Associação Brasileira da Indústria de Madeira Processada Mecanicamente. **Estudo Setorial 2013**. 2013. Available at: < http://www.abimci.com.br/abimci-lanca-estudo-setorial-2013/> Accessed in Aug. 2014.

BALAS, B. J. Texture synthesis and perception: Using computational models to study texture representations in the human visual system. **Vision Research** 46, p. 299–309, 2006.

BELLMAN, R. E. **An Introduction to Artificial Intelligence**: Can Computers Think? Boyd & Fraser Publishing Company, 1978.

BENGIO, Y. Learning Deep Architectures for AI. **Foundations and Trends in Machine Learning**, Montreal, vol. 2, pp 2-127, 2009.

BRAZILIAN ASSOCIATION FOR MECHANICALLY PROCESSED TIMBER – ABIMCI. **Improving processes in the production chain**. Curitiba – PR, 2008.

BRAZILIAN ASSOCIATION FOR TECHNICAL STANDARDS - ABNT (NBR 11700) - **Conifers sawn wood from plantation for general uses** - Grading rules - Classification. Rio de Janeiro, 1991a.

BRAZILIAN ASSOCIATION FOR TECHNICAL STANDARDS - ABNT (NBR 12297) - **Conifers sawn wood from plantation for general uses -** Measurement and quantity of defects – Procedure. Rio de Janeiro, 1991b.

BRAZILIAN ASSOCIATION FOR TECHNICAL STANDARDS - ABNT (NBR 12551) - **Sawnwood** – Terminology. Rio de Janeiro, 2002.

BRAZILIAN ASSOCIATION OF PLANTED FORESTS PRODUCERS – ABRAF. **Anuário estatístico ABRAF 2013.** Brasília: 2013. Available at: <http://www.ipef.br/estatisticas/relatorios/anuario-ABRAF13-BR.pdf> Accessed in Oct. 2014.

BRAZILIAN DEVELOPMENT BANK – BNDES. Circular Nº 34, de 06 de setembro de 2011. Available at <http://www.bndes.gov.br/SiteBNDES/bndes/bndes_pt/Institucional/Apoio_Financeiro/porte.html> Accessed in Sept. 2015.

BRUNNER, C. C.; SHAW, G. B.; BUTLER, D. A.; FUNCK, J. Using Colour in Machine Vision Systems for Wood Processing. **Wood and Fibre Science**, v. 22, p. 413-428, 1990.

BUDAKCI, M.; CINAR, H. Colour effects of stains on wood with knots, cracks and rots. **Progress in Organic Coatings**. Vol. 51, Issue 1, p. 1-5, 2004.

CHARNIAK, E.; McDERMOTT, D. **Introduction to Artificial Intelligence**. Addison-Wesley, 1985.

CHIRICI, G.; MURA, M.; McINERNEY, D.; PY, N.; TOMPPO, E. O.; WASER, L. T.; TRAVAGLINI, D.; McROBERTS, R. E. A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data. **Remote Sensing of Environment** 176, p. 282-294, 2016.

CHOU, J.; LIN, C.; PHAM, A.; SHAO, J. Optimized artificial intelligence models for predicting project award price. **Automation in Construction** 54, p. 106–115, 2015.

COPPIN, B. **Inteligência Artificial**. Translated from Artificial Intelligence Illuminated by Jorge Duarte Pires Valério. Ed. Gen/LTC. Rio de Janeiro, 2012.

__________. **Artificial Intelligence Illuminated**. Jones and Bartlett Publishers Inc, Boston, 2004.

DE SILVA, C. Automation Intelligence. **Engineering Application of Artificial lntelligence**. Vol. 7, No. 5, p. 471-477, 1994.

DUDA, R.O.; HART, P.E.; STOCK, D. G. **Pattern Classification**. John Wiley and Sons Inc., New York, 2001

ERHAN, D.; BENGIO, Y.; COURVILLE, A.; VINCENT, P. **Visualizing higher-layer features of a deep network.** ICML - Workshop on Learning Feature Hierarchies, Montreal, Canada, 2009.

ERTUGRUL, Ö. F. Forecasting electricity load by a novel recurrent extreme learning machines approach. **International Journal of Electrical Power & Energy Systems**. vol. 78. p. 429-435, 2016.

FAWCETT, T. **ROC Graphs**: Notes and Practical Considerations for Data Mining Researchers. Intelligent Enterprise Technologies Laboratory. Palo Alto, 2003.

FASEL, B. **Robust face analysis using convolutional neural networks**. Proceedings of International Conference on Pattern, 2002.

FERNANDES, C.; RAMOS, V.; ROSA, A. C. Self-Regulated Artificial Ant Colonies on Digital Image Habitats. **Int. Journal of Lateral Computing**, v. 2, p. 1-8, 2005.

FERRI, C.; HERNÁNDEZ-ORALLO, J.; MODROIU, R. An experimental comparison of performance measures for classification. **Pattern Recognition Letters**, v. 30, n. 1, p. 27–38, 2009.

FLESIA, A. G.; BAUMGARTNER, J.; GIMENEZ, J.; MARTINEZ, J. Accuracy of MAP segmentation with hidden Potts and Markov mesh prior models via Path Constrained Viterbi Training, Iterated Conditional Modes and Graph Cut based algorithms. **Computing Research Repository**. 1307.2971. v1. Cornell University, New York, 2013.

FRANCO, M.A.; BACARDIT, J. Large-scale experimental evaluation of GPU strategies for evolutionary machine learning. **Information Sciences**. vol. 330, p. 385-402, 2016.

FU, L.; KARA, L. B. From engineering diagrams to engineering models: Visual recognition and applications. **Computer-Aided Design**, 43. p. 278–292, 2011.

FU, M.; XU, P.; LI, X.; LIU, Q. YE, M.; ZHU, C. Fast crowd density estimation with convolutional neural networks. **Engineering Applications of Artificial Intelligence**, Vol.43, p. 81-88, 2015.

FUKUSHIMA, K. Neocognitron: a self-organizing neural networks for a mechanism of pattern recognition unaffected by shift in position. **Biological Cybernetics**, 36, p. 193 – 202, 1980.

FUNCK J.; ZHONG, Y.; BUTLER, D.; BRUNNER, C.; FORRER, J. Image segmentation algorithms applied to wood defect detection. **Comput Electron Agric**. Vol 41, p. 157–179, 2003.

GELERNTER, H. **Realization of a Geometry Theorem Proving Machine**. IFIP Congress, 1959.

GEORGE, D. **How the brain might work**: A hierarchical and temporal model for learning and recognition, Ph.D. Dissertation, Stanford University, Stanford, CA, 2008.

GIERE, R. N. Cognitive models of science. **Minnesota studies in the philosophy of science**; v. 15, University of Minnesota Press, 1992.

GOMES, R. C. Desenvolvimento de uma Base de Dados de Imagens Digitais de Madeira Serrada de Coníferas. **Dissertação**. Faculdade de Ciências Agronômicas da UNESP. Botucatu, 2013.

GONZALES, R. C.; WOODS, R. E. **Digital Image Processing**. 3rd. Ed. NJ: Prentice Hall, 2008.

GU, J.; PAN, Y.; WANG, H. Research on the improvement of image edge detection algorithm based on artificial neural network. **Optik** 126. p. 2974–2978, 2015.

GUNN, S. R. **Support Vector Machines for Classification and Regression**. Technical Report. University of Southhampton, 1998.

HADIZADEH, H. Multi-resolution local Gabor wavelets binary patterns for texture description. **Pattern Recognition Letters**. Vol. 65. p. 163-169, 2015.

HADDOCK, J.; O'KEEFE, R. M. Using artificial intelligence to facilitate manufacturing systems simulation. **Computers & Industrial Engineering**. Vol. 18, Issue 3, p.275-283, 1990.

HAMID, B.; MEYBODI, R. M. A learning automata-based algorithm for three-layer neural networks. **Int'l Journal of Systems Science**. Vol. 40, No.1, p. 101–118, 2009.

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

HARALICK, R. M.; SHAPIRO, L. G. Glossary of Computer Vision Terms. **Pattern Recognition**. Vol. 24. No. 1. p. 69-93, 1991.

HASHIM, U.R.; HASHIM, S. Z.; MUDA, A. K. Automated vision inspection of timber surface defect: a review. **Jurnal Teknologi.** v. 77, p.127–135, 2015.

HAUGELAND, J. **Artificial Intelligence**: The Very Idea. MIT Press, 1985.

HAYKIN, S. **Neural Networks:** A Comprehensive Foundation. New York: Willey & Sons, 2001.

__________. **Redes Neurais**: Princípios e prática. Trad. Paulo Martins Engel. 2.ed. Porto Alegre: Bookman, 2007.

HE, J. A. Comparison of Artificial Neural Network Classifiers for Analysis of CT Images for the Inspection of Hardwood Logs. **Thesis** submitted to the Faculty of the Bradley Department of Electrical and Computer Engineering of Virginia, 1997.

HEBB, D. O. **The Organization of Behavior**: a Neuropsychological Theory. John Wiley & Sons, Inc. New York, 1949.

HEUNG, B.; HO, H. C.; ZHANG, J.; KNUDBY, A.; BULMER, C. E.; SCHMIDT, M. G. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. **Geoderma**. vol. 265. p. 62–77, 2016.

HINTON, G. E.; DENG, L.; YU, D.; DAHL, G. E.; MOHAMED, A.; JAITLY, N.; SENIOR, A.; VANHOUCKE, V.; NGUYEN, P.; SAINATH, T. N.; KINGSBURY, B. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. **IEEE Signal Process**, p. 82–97, 2012.

HUANG, F. J.; LECUN, Y. **Large-scale learning with SVM and convolutional nets for generic object categorization**. In Proc. Computer Vision and Pattern Recognition Conf. (CVPR'06), 2006.

HUBEL, D. H.; WIESEL, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. **J. Physiol**. 160 (1), 106, 1962.

HUBER, H. A.; MCMILLIN, C. W.; MCKINNEY, J. P. Lumber Defect Detection Abilities of Furniture Rough Mill Employees. **Forest Products Journal**, v. 35, p. 79-82, 1985.

IBÁ – Indústria Brasileira da Árvore. **Publicação do Relatório do Desempenho do Setor 2014**. Available at: <http:// http://iba.org/images/shared/iba_2015.pdf.> Accessed in: December, 2015.

JANG J. R.; SUN, C. T.; MIZUTANI, E. **Neuro-Fuzzy and Soft Computing**: A Computational Approach to Learning and Machine Intelligence. Prentice Hall, 1996.

JAPKOWICZ, N.; SHAH, M. **Evaluating learning algorithms**: A classification perspective. New York, NY, USA: Cambridge University Press, 2011.

JOHNSON, D. M. V.; BODEDE, O. R; ADESINA, F. Wood Processing Industries in Nigeria: Problems and the Way. **Innovative Systems Design and Engineering**. Vol.5, No.6, 2014.

KAMAL, M. W., AHMED, M. A. A Proposed Framework for Use Case based Effort Estimation using Fuzzy Logic: Building upon the outcomes of a Systematic Literature Review. **IJNCAA**. p. 953-976. Hong Kong, 2011.

KANTA, G.; SANGWAN, K. S. Predictive Modeling for Power Consumption in Machining using Artificial Intelligence Techniques. 12th Global Conference on Sustainable Manufacturing. **Procedia CIRP** 26, pp 403 – 407, 2015. Available at <http://ac.els-cdn.com/S2212827114008853/1-s2.0-S2212827114008853-main.pdf?_tid=aa19 0b2a-feef-11e4-bb51-00000aab0f6b&acdnat=1432126783_a4f717dd1928cb798041843b621c2dd8> Accessed in: May 2015.

KHANA, F. S.; ANWERB, R. M.; VAN DE WEIJERC, J.; FELSBERGA, M.; LAAKSONENB, J. Compact color–texture description for texture classification. **Pattern Recognition Letters**, Vol. 51. p. 16–22, 2015.

KIM, J.; KIM, B. S.; SAVARESE, S. Comparing Image Classification Methods: K-Nearest-Neighbor and Support-Vector-Machines**. Applied Mathematics in Electrical and Computer Engineering**. p. 133-138. Cambridge, 2012.

KIM, S.; CHOI, Y.; LEE, M. Deep learning with support vector data description. **Neurocomputing** 165. 2015. p. 111 – 117. Available at: <http://www.sciencedirect.com/science/article/pii/S092523121500380X>. Accessed in: January 2016.

KLINE, D. E.; SURAK, C.; ARAMAN, P. A. Automated hardwood lumber grading utilizing a multiple sensor machine vision technology. **Computers and Electronics in Agriculture,** v. 41, n. 1–3, p. 139 – 155, 2003. Available at: <http://www.sciencedirect.com/science/article/pii/S0168169903000486>. Accessed in: October, 2014.

KORNIENKO, A. A.; KORNIENKO, A. V.; FOFANOV, O. B.; CHUBIK, M. P. Knowledge in artificial intelligence systems: searching the strategies for application. **Procedia - Social and Behavioral Sciences** 166. p. 589 – 594, 2015.

KOW, K. W.; WONG, Y. W.; RAJKUMAR, R. K.; RAJKUMAR, R. K. A review on performance of artificial intelligence and conventional method in mitigating PV grid-tied related power quality events. **Renewable and Sustainable Energy Reviews**. Vol. 56, p. 334-346, 2016.

KOVÁCS, Z. L. **Redes Neurais Artificiais**: Fundamentos e Aplicações. Ed. Collegium Cognitio, 1996.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. Imagenet classification with deep convolutional neural networks. In: BARTLETT, P.; PEREIRA, F.; BURGES, C.; BOTTOU, L.; WEINBERGER, K. **Advances in Neural Information Processing Systems**. p. 1106 – 1114, 2012.

KUMARI, R.; KUMAR, V.; KUMAR S. Adaptive Neural Fuzzy Inference System for Employability Assessment. **International Journal of Computer Applications Technology and Research,** Vol. 3 – Issue 3, p. 159 - 164, 2014.

KURZWEIL, R. **The Age of Intelligent Machines**. MIT Press, 1990.

LAKATOS, E. M. MARCONI, M. A. **Fundamentos de Metodologia Científica** - 7ª Ed. São Paulo: Atlas, 2010.

LECUN, Y.; BENGIO, T. **Convolutional networks for images, speech, and time series**. In M. A. ARBIB (Ed.). The handbook of brain theory and neural networks. p. 255 – 258. Cambridge, MA: MIT Press, 1995.

LECUN, Y.; BOSER, B.; DENKER, J. S.; HENDERSEN, D.; HOWARD, R. E.; HUBBARD, W.; JACKEL, L. D. Backpropagation Applied to Handwritten Zip Code Recognition. **Neural Computation**. 1, p. 541-551. 1989.

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. **Gradient-based learning applied to document recognition.** Proc. of the IEEE, vol. 86 (11), p. 2278–2324, 1998.

LECUN, Y.; BOTTOU, L.; ORR, G. B.; MULLER, K. R. **Efficient BackProp**. In: ORR, G. B.; MULLER, K. R. (Eds). Neural Networks: Tricks of the Trade. LNCS 1524, p. 09-50, 1998.

LECUN, Y.; CHOPRA, S.; HADSELL, R.; RANZATO, M. A.; HUANG, F. J. **A tutorial on energy-based learning.** In BAKIR, G.; HOFMAN, T.; SCHOLKOPF, B.; SMOLA, A.; TASKAR, B. (Eds.). Predicting Structured Data. MIT Press, 2006.

LEE, A. H. I., KANG, H. Y., HSU, C. F., HUNG, H. C. A green supplier selection model for high-tech industry. **Expert Systems with Applications** 36. p. 7917–7927, 2009.

LI, P.; YAN, F.; GE, C.; WANG, X.; XU, L.; GUO, L. A simple fuzzy system for modelling of both rate-independent and rate-dependent hysteresis in piezoelectric actuators. **Mechanical Systems and Signal Processing**. Volume 36, Issue 1, p. 182–192, 2013.

LI, Y.; PO, L-M.; XU, X.; FENG, L.; YUAN, F.; CHEUNG, C-H.; CHEUNG, K-W. No-reference image quality assessment with shearlet transform and deep neural networks. **Neurocomputing.** Volume 154, p. 94–109, 2015.

LIBERATORE, M. J. Automation, AI and OR: in search of the synergy and publication priorities. **European Journal of Operational Research**. Vol. 99, Issue 2, p. 248-255, 1997.

LORENA, A. C. Investigação de estratégias para a geração de máquinas de vetores suporte multiclasses. **Tese** (Doutorado) — Instituto de Ciências Matemáticas e de Computação - ICMC - USP, Janeiro 2006.

LU, K. C.; YANG, D. L. Image Processing and Image Mining using Decision Trees. **Journal of Information Science and Engineering** 25. p. 989-1003. Taiwan, 2009.

LUGER, G. F. **Computation & Intelligence**: Collected Readings. The AAAI Press / The MIT Press, 1995.

LUGER, G. F. **Artificial intelligence**: Structures and strategies for complex problem solving. 5th ed. Boston: Addison Wesley, 2004.

MA, L.; DESTERCKE, S.; WANG, Y. Online active learning of decision trees with evidential data. **Pattern Recognition** 52. p. 33–45. 2016.

MANAP, R. A.; SHAO, L. Non-distortion-specific no-reference image quality assessment: A survey. **Information Sciences**. Volume 301. p. 141–160, 2015.

MAMDANI, E. H.; ASSILIAN, S. An experiment in linguistic synthesis with a fuzzy logic controller. **International Journal Man Mach**. Studies vol. 7, p. 1-13, 1975.

MANIYAR, D.; CORNFORD, D. **Dealing with large complex models in emulators:** Guided Bayesian committee machine. Managing Uncertainty and Complex Models. Neural Computing Research Group. Aston University, Birmingham, UK, 2006.

MARSICO, M.; NAPPI, M.; RICCIO, D.; WECHSLER, H. Robust face recognition after plastic surgery using region-based approaches. **Original Research Article Pattern Recognition,** v.48, Issue 4, p. 1261-1276, 2015.

MARSZALEC, E.; PIETIKAINEN, M. Colour Analysis for Automated Visual Inspection of Pine Wood. **SPIE**, v. 1907: Machine Vision Application in Industrial Inspection, San Jose, p. 80-94, 1993.

MATSUGU, M.; MORI, K.; MITARI, Y.; KANEDA, Y. **Subject independent facial expression recognition with robust face detection using a convolutional neural network**. Canon Research Center, Atsugi, Japan, 2003.

MCCARTHY, J. **Programs with Common Sense**. In Proceedings of the Teddington Conference on the Mechanization of Thought Processes, p. 756-791. London, 1959.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**. v. 5, Issue 4, p. 115-133, 1943.

MEYER, H.; KÜHNLEIN, M.; APPELHANS, T.; NAUSS, T. Comparison of four machine learning algorithms for their applicability in satellite-based optical rainfall retrievals. **Atmospheric Research**. vol. 169, part B, p. 424-433, 2016.

MITCHELL, T.M. **Machine Learning**. New York: McGraw-Hill, 1997.

MIKOLOV, T.; DEORAS, A.; POVEY, D.; BURGET, L.; CERNOCKY, J. Strategies for Training Large Scale Neural Network Language Models. In Proceedings of **ASRU,** 2011.

MISHRA, R. B. **Artificial Intelligence.** PHI Learning Private Limited. New Delhi, 2011. Available at: <https://books.google.com.br/books?id=Mishra> Accessed in: Mar, 01 2015.

MOIGNE, J. L. Advances in artificial intelligence. Seventh European Conference on Artificial Intelligence. **European Journal of Operational Research**. Vol. 36, Issue 2, p.274-285, 1988.

MRAZOV, I.; KUKACKA, M. Can Deep Neural Networks Discover Meaningful Pattern Features? **Procedia Computer Science** 12, p. 194 – 199, 2012.

NAGABHUSHANA, S. **Computer Vision and Image Processing**. New Age International Publisher, New Delhi, 2006.

NANNI, L.; MELUCCI, M. Combination of projectors, standard texture descriptors and bag of features for classifying images. **Neurocomputing** 173, p. 1602–1614, 2016.

NAUCK, D. D.; SPOTT, M.; AZVINE, B**. Fuzzy Methods for Automated Intelligent Data Analysis**. Fuzzy Systems, 2004. Proceedings. 2004 IEEE International Conference. p. 487-492, vol.1, 2004.

NEWELL, A.; SIMON, H. **Human problem solving**. New York, Prentice Hall, 1976.

NILSSON, N. J. **Artificial Intelligence**: A New Synthesis. Morgan Kaufmann, 1998.

NYSTRÖM, J. Automatic measurement of compression wood and spiral grain for the prediction of distortion in sawn wood products. **Thesis**. 142 p. Luleå University of Technology Department of Engineering Sciences and Mathematics. Wood Technology. Luleå, Sweden, 2002.

OJALA, T.; PIETIKAINEN, M.; MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **PAMI**. Vol. 24. p. 971–987, 2002.

OLIVEIRA, L.E.S.; CAVALIN, P.R.; BRITTO JR, A. S.; KOERICH, A. L. Inspeção Automática de Defeitos em Madeiras de Pinus usando Visão Computacional. **RITA**. Vol. 15, No. 2, 2008.

OLIVEIRA, L. J. G. G. **Secagem Convencional de Madeira Serrada de *Pinus taeda l.* buscando Atender a Indústria Moveleira do Sul do Brasil**. Universidade Federal de Santa Catarina Curso de Graduação em Engenharia Florestal - Campus Curitibanos. Curitibanos, 2014.

OPTIZ, D. Neural Network Ensemble. **Connection Science**, v. 12, p.212-225, 1990.

ÖSTERBERG, P. Wood Quality and Geometry Measurements Based on Cross Section Images. **Thesis** (Doctorate Degree). Tampere University of Technology. Julkaisu 807 – Publication 807, 2009.

PAHLBERG, T.; JOHANSSON, E.; HAGMAN, O.; THURLEY, M. Wood fingerprint recognition using knot neighborhood K-plet descriptors. **Wood Science and Technology**. Vol. 49, Issue 1, p. 7–20, 2015.

PHAM, D. T.; ALCOCK, R. J. **Smart Inspection Systems**: Techniques and Application of Intelligent Vision. Academic Press: London, 2003.

PHAM, D. T.; SOROKA A. J; GHANBARZADEH, A.; KOC, E.; **Optimising Neural Networks for Identification of Wood Defects Using the Bees Algorithm**. IEEE International Conference on Industrial Informatics, p. 1346-1351, 2006.

POLZLEITNER; W.; SCHWINGSHAKL, G. Real-Time Surface Grading of Profiled Wooden Boards. **Industrial Metrology**. Part 2, Issue 3-4, p. 283-298, 1992.

POOLE, D.; MACKWORTH, A. K.; GOEBEL, R. **Computational intelligence**: A logical approach. Oxford University Press, 1998.

POWAR, A.; DATE, P. Modeling of microstructure and mechanical properties of heat treated components by using Artificial Neural Network. **Original Research Article Materials Science and Engineering**, v.628, p. 89-97, 2015.

POYRY - International Consulting and Engineering Company. **Financial Statements 2014.** Available at: <http://www.poyry.com/sites/default/files/imce/poyry_fs2014_eng.pdf> Accessed on: December 2015.

QUAN, Y.; XU, Y.; SUN, Y. A distinct and compact texture descriptor. **Image and Vision Computing** 32, p. 250–259, 2014.

QUINLAN, J. R. Induction of decision trees. **Machine Learning** (1) p. 81-106. 1986.

QUINLAN, J. R. **Programs for Machine Learning**. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA. 1993.

RICH, E.; KNIGHT, K. **Artificial Intelligence**. 2nd ed. McGraw-Hill, 1990.

ROKACH, L.; MAIMON, O. Top-down induction of decision trees classifiers: a survey. **Trans. Sys. Man Cyber Part C 35** (4) p. 476-487. 2005.

ROSA, E. S.; SALGADO, R. M.; OHISHI, T.; MASTELARI, N**.** Performance comparison of artificial neural networks and expert systems applied to flow pattern identification in vertical ascendant gas–liquid flows. **International Journal of Multiphase Flow** 36, p. 738–754, 2010.

ROSENBLATT, M. The Perceptron: A probabilistic model for information storage and organization in the Brain. **Psychological review**, vol. 65, n.6, p. 386-408, 1958.

ROSSI, A. L. D. Meta-aprendizado aplicado a fluxos contínuos de dados. **Thesis.** 189 p. Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2014.

RUSSELL, S., NORVIG, P. **Artificial Intelligence**: A Modern Approach. 3rd Ed. Pearson New International Edition, 2014.

SÁNCHEZ, A. S.; IGLESIAS-RODRÍGUEZ, F. J.; FERNÁNDEZ, P. R.; JUEZ, F. J. C. Applying the K-nearest neighbor technique to the classification of workers according to their risk of suffering musculoskeletal disorders. **International Journal of Industrial Ergonomics** 52. p. 92-99. 2016.

SCHMOLDT, D. L.; OCCENA, L. G.; LYNN ABBOTT, A.; GUPTA, N. K. Nondestructive Evaluation of Hardwood Logs: CT Scanning, **Machine Vision and Data Utilization Nondestructive Testing and Evaluation**. Newark, 2000.

SERRE, T., KREIMAN, G., KOUH, M., CADIEU, C., KNOBLICH, U., POGGIO, T. A quantitative theory of immediate visual recognition. **Progress in Brain Research,**

Computational Neuroscience: Theoretical Insights into Brain Function, 165, p. 33–56, 2007.

SCHALKOFF, R. J. **Artificial Intelligence**: An Engineering Approach. McGraw Hill, 1990.

SHARMA, C.; SABHARWAL, S.; SIBAL, R. **A Survey on Software Testing Techniques using Genetic Algorithm**. IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 1, 2013.

SHAPIRO, L. G.; STOCKMAN, G. C. **Computer Vision**. Prentice Hall, Upper Saddle River, New Jersey, 2001.

SIGILLITO, V. G.; HOTTON, L. V.; CHAUMETTE, E.; NOUALS, C.; BESSON, O. Radar Signal Processing: Case Study II. **Neural Network PC Tools.** 1991.

SIMARD, P. Y.; STEINKRAUS, D.; PLATT, J. **Best practices for convolutional neural networks**. In Proc. of ICDAR, 2003.

SOCHIRCA, E.; AFONSO, Ò; GIL, P. M. Technological-knowledge bias and the industrial structure under costly investment and complementarities. **Economic Modelling**. Vol. 32, p.440-451, 2013.

SONKA, M.; HLAVAC, V.; BOYLE, R. **Image Processing, Analysis and Computer Vision**. PWS Publishing, NY, 1999.

STATNIKOV, A.; WANG, L.; ALIFERES, C. A comprehension comparison of random forests and support vector machines for microarray-based cancer classification. **BMC Bioinformatics** 9 (1). p. 1-10, 2008.

STUCKENS, J.; COPPIN, P. R.; BAUER, M. E. Integrating Contextual Information with per-Pixel Classification for Improved Land Cover Classification. **Remote Sensing of Environment**, Vol. 71, Issue 3, pp 282-296. 2000.

TAHMASEBI, P. HEZARKHANI, A. Application of Adaptive Neuro-Fuzzy Inference System for Grade Estimation: Case Study. **Australian Journal of Basic and Applied Sciences,** 4(3): 408-420, 2010.

TODOROKI, C.; LOWELL, E.; DYKSTRA, D. Automated knot detection with visual post-processing of Douglas-fir veneer images. **Comput Electron Agric**. Vol. 70, p.163–171, 2010.

TOPALOVA, I; TZOKEV, A. **Modular Structure of Neural Networks for Classification of Wooden Surfaces with PLC Industrial Implementation.** ICAS 2011: The Seventh International Conference on Autonomic and Autonomous Systems, p.13-17, 2011

TRIPOLI, A. C. K.; PRATES, R. C. Certificação ambiental e internacionalização: uma análise do setor madeireiro brasileiro. **Desenvolvimento em Questão**. Vol. 13, n. 31, 2015.

UKIVA. **Twenty-one Financial Justifications for using Machine Vision**. UK Industrial Vision Association. Herts, UK. 2002. Available at <http:// www.ukiva.org/image-processing-software.html> Accessed on Jan. 2015.

VAPNIK, V.; GOLOWICH, S.; SMOLA, A. Support vector method for function approximation, regression estimation, and signal processing. In: MOZER, M.; JORDAN, M.; PETSCHE, T. **Advances in Neural Information Processing Systems** 9, p. 281– 287, Cambridge, MA, 1997.

VAPNIK, V. **The Nature of Statistical Learning Theory**. Springer, N.Y., 1995.

VENKATASUBRAMANIAN, V.; MCAVOY, T. J. Neural network applications in chemical engineering. **Computers & Chemical Engineering**, v. 16, Issue 4, 1992.

VIEIRA, F. H. A.; AFFONSO, C.; ALVES, M. C. S. Application of Neuro-Fuzzy Inference System on Wood Identification. **Applied Mechanics and Materials**. Trans Tech Publications Inc., 2014.

VILLA, J. L.; BOQUÉ, R.; FERRÉ, J. Chemometr. Intelligence. **Lab. System** 94 p. 51–59, 2008.

WALCOTT-JUSTICE, K. R. Exploiting Hardware Monitoring in Software Engineering. In: HURSON, A. R.; MEMON, A. M. **Advances in Computers**, Vol. 93, 2014. p. 53-101.

WEI, L.; HONG-YING, D. Real-time Road Congestion Detection Based on Image Texture Analysis. **Procedia Engineering** 137. p. 196 – 201, 2016.

WINSTON, P. H. **Artificial Intelligence**. 3rd Ed. Addison-Wesley, 1992.

WU, C. H.; LAI, C. C.; CHEN, C. Y.; CHEN, Y. H. Automated clustering by support vector machines with a local-search strategy and its application to image segmentation. **Optik** 126, p. 4964–4970, 2015.

YAO, L.; SURYANARAYAN, P.;·QIAO, M.; WANG, J. Z.;·LI, J. OSCAR: On-Site Composition and Aesthetics Feedback Through Exemplars for Photographers. **International Journal Computer Vision**, 96, p. 353–383, 2012.

ZAMRI, M. I. P.; CORDOVA, F.; KHAIRUDDIN, A. S. M. MOKHTAR, N.; YUSOF, R. Tree species classification based on image analysis using Improved-Basic Gray Level Aura Matrix. **Computers and Electronics in Agriculture**, Vol. 124, pp. 227-233, 2016.

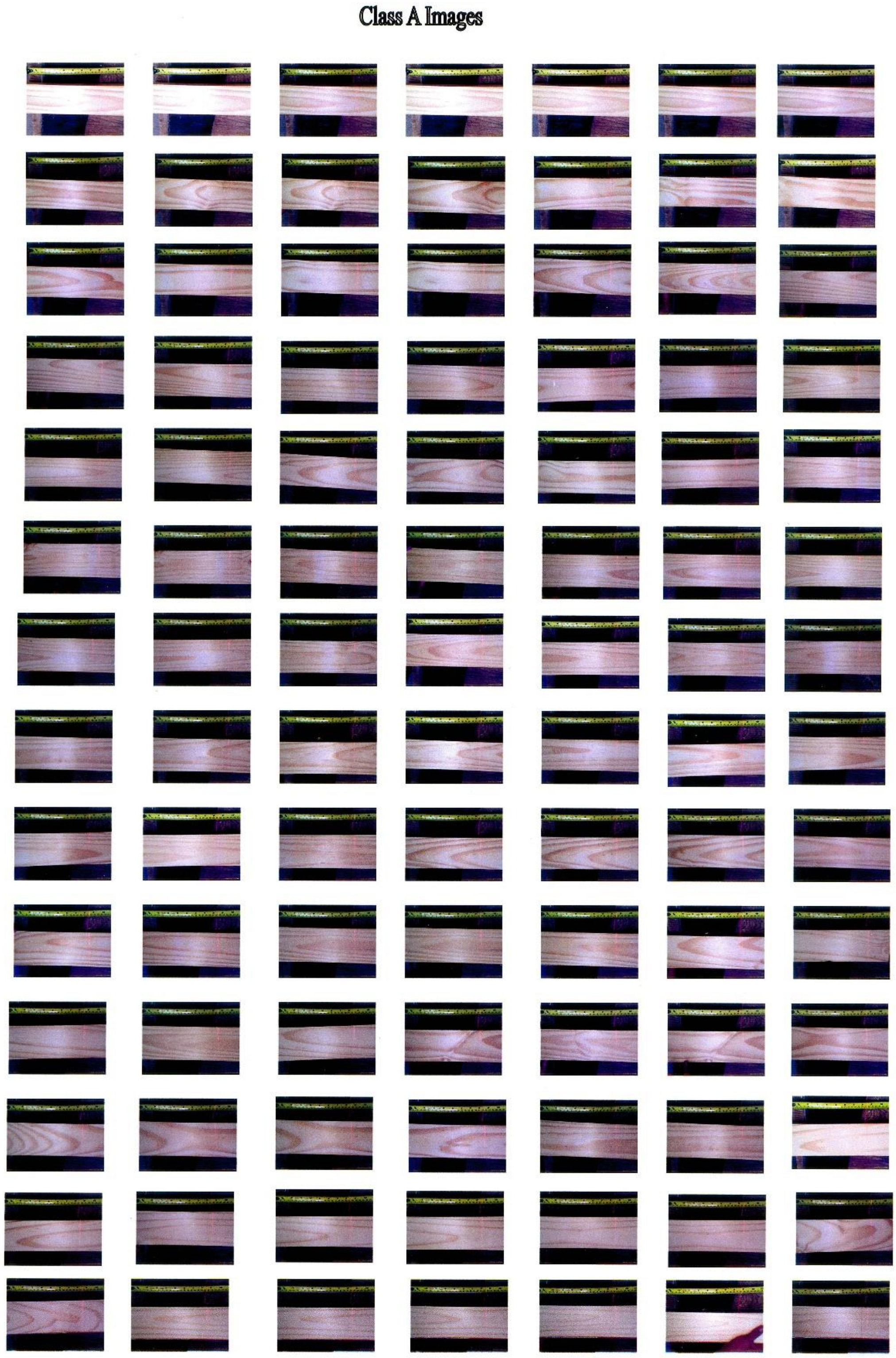
ANNEX A – Classification of Sawn Softwood from Reforestation

Table : Classification of Sawn Softwood from Reforestation

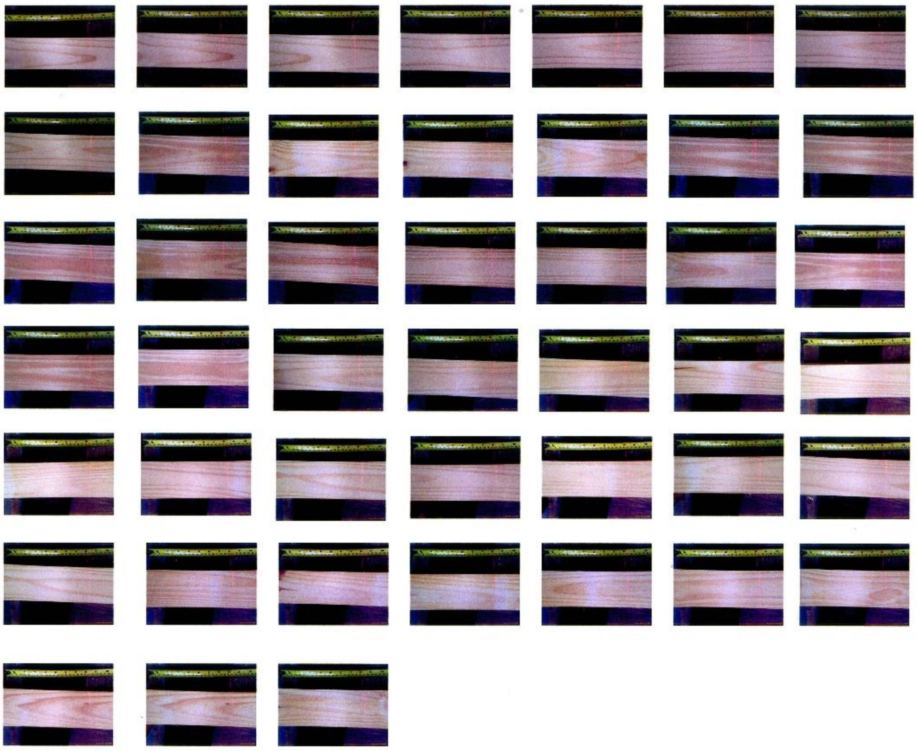
Defects		Characteristics	Quality Classes: Super	Extra	First	Second	Third
Knots	Firm. Group of knots	Sum of the Diameters	≤ 50% $l_{l/m}$	≤ 70% $l_{l/m}$	≤ 90% $l_{l/m}$	No Restrictions	All defects exceeding the limits of the second class, except those that impede the use of the piece, such as advanced rot, insects holes (active), etc.
	Loose and Decayed	Sum of the Diameters	Not Allowed	Not Allowed	≤ 30% 1 1/m	≤ 200% 1 1/m	
	Open				≤ 10% $l_{l/m}$	≤ 30% $l_{l/m}$	
Pitch Pockets		Sum of the widths	Not Allowed	≤ 5% l_l	≤ 10% l_l	≤ 15% l_l	
		Sum of the Lenghs		≤ 25% L_l	No Restrictions		
Pith		Widths	Not Allowed	≤ 12 mm	Allowed without Restrictions		
		Sum of the Lenghs		≤ 15% l_l			
Shakes		Sum of the Lenghs	Not Allowed	≤ 10% L_l	≤ 20% L_l	≤ 50% L_l	
		Individual Lengh		0,15 m	0,20 m	0,50 m	
Cracked		Depth	Not gauged planing		Allowed without Restrictions		
Wane		Thickness	≤ 15% e_l	≤ 15% e_l	≤ 30% e_l	≤ 30% e_l	
		Width	≤ 3% l_l	≤ 5% l_l	≤ 10% l_l	≤ 20% l_l	
		Sum of the Lenghs	≤ 20% L_l	≤ 20% L_l	≤ 30% L_l	≤ 40% L_l	
		Number of Corners	1	1	1	2	
	Brown Spot	Affected Area	Not Allowed	≤ 10% A_l	Allowed without Restrictions		
	Blue Spot			≤ 10% A_l	≤ 25% A_l	≤ 50% A_l	
	Incipient Rot			Not Allowed	≤ 10% A_l	≤ 15% A_l	
	Insect Holes (inactive)			≤ 15% A_l	≤ 30% A_l	≤ 50% A_l	
Gallery, Insect Holes (active), Rot		Presence	Not Allowed				
Inclination, spurs, exudation and crisscrossed grain		Presence	Allowed without Restrictions				
All of them		Quantity	No Restrictions	≤6	≤7	No Restrictions	

Source: Brazilian Association for Technical Standards - NBR 11700 (own translation)

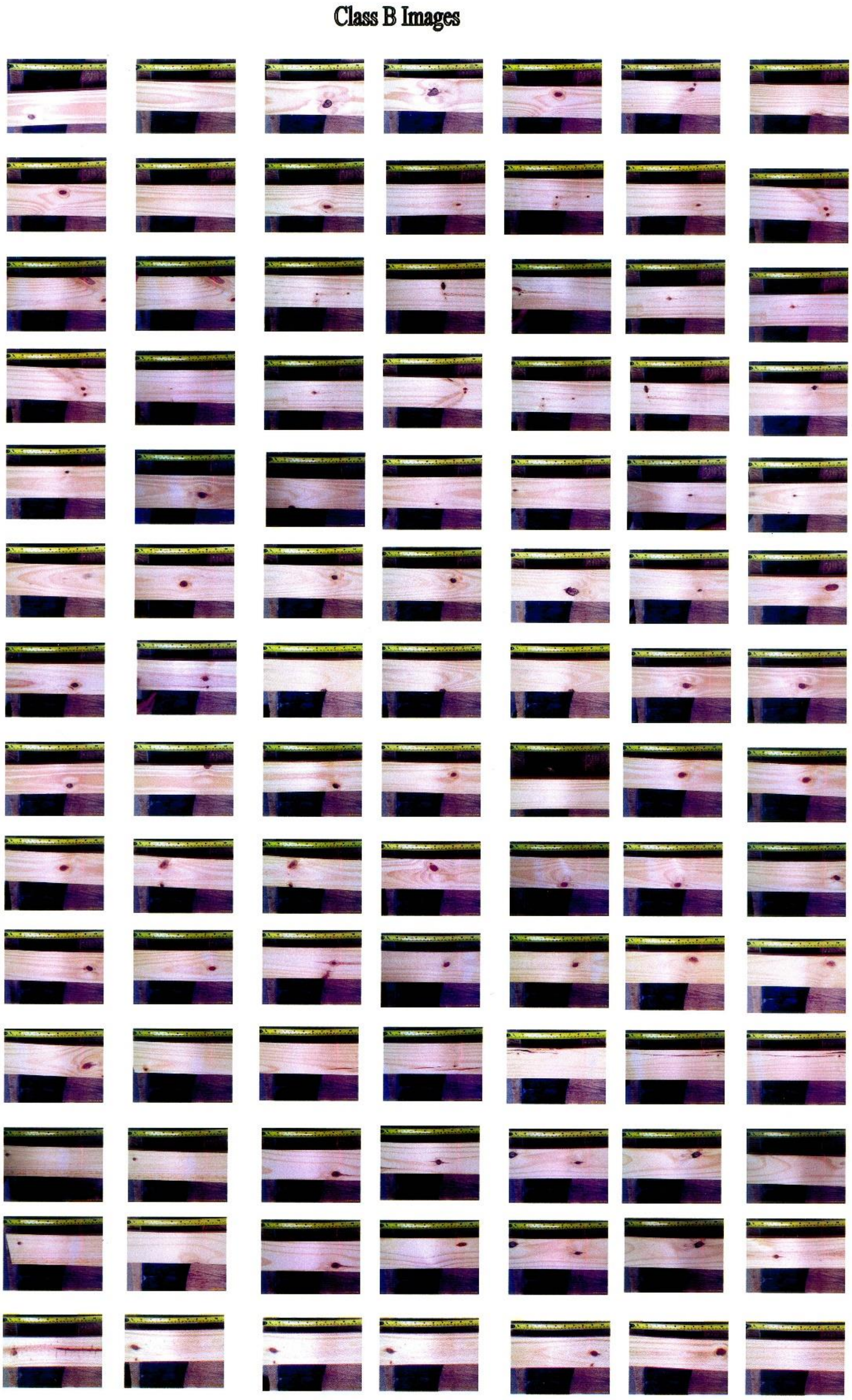
APPENDIX A – A-Class board images



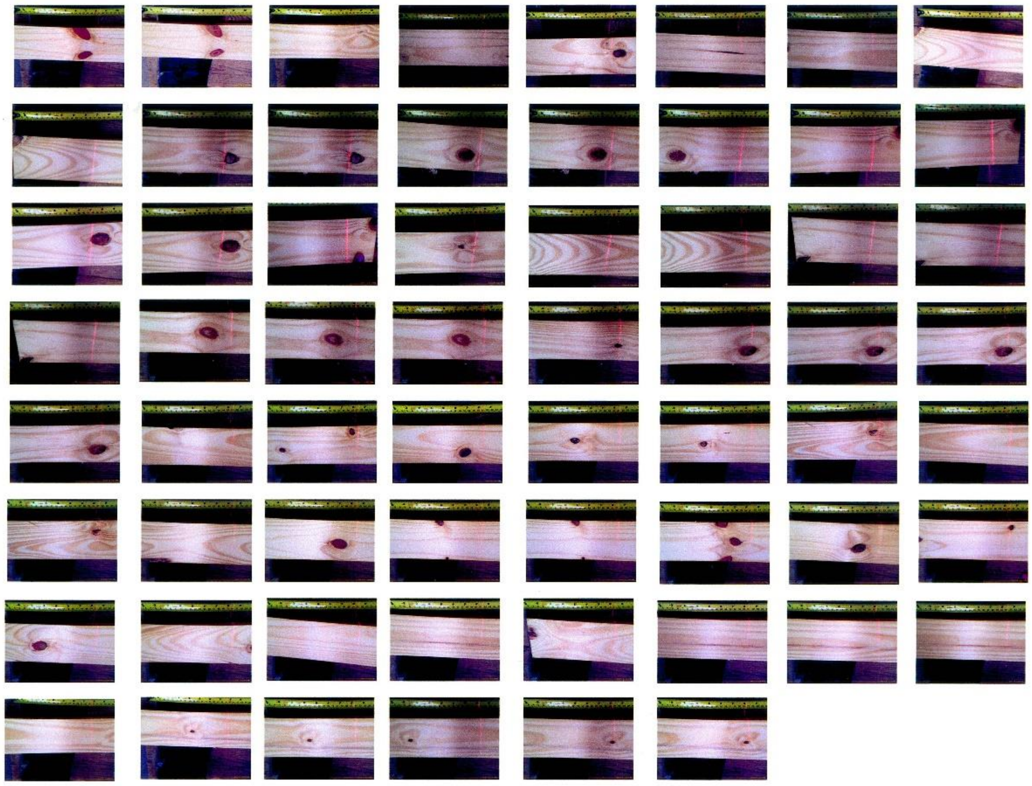
Class A Images



APPENDIX B - B-Class board images

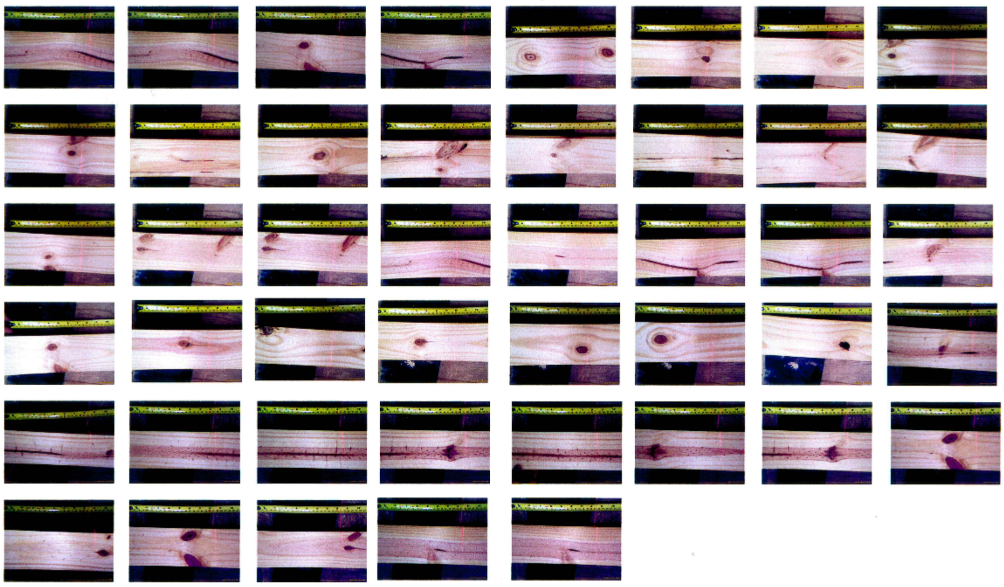


Class B Images



APPENDIX C - C-Class board images

Class C Images



APPENDIX D – Haralick Texture Feature

```
// Haralick Texture Feature

//

// Author:  may 15th 2015

//
====================================================================

// haralick(theta)

====================================================================
//

// Haralick Texture Features Scilab

//

// Haralick(image,input_bits,output_bits,d,theta)

//

// This function generates a Spatial Gray Level Dependence Matrix
(SGLD) of

// a gray scale image and calcuates 5 Haralick's texture features
from it.

//

// The Haralick's texture features calculated are:

// 1) Energy

// 2) Entropy Correlation

// 3) Maximum probility Inertia

// 4) Correlation

// 5) Inertia

//

// The inputs to the function are

// a) image: the input gray scale image

//

// b) input_bits: the gray level resolution of the input image,

//                 eg. for a 255 gray level image input_bits=8
```

```
//

// c) output_bits: the gray level resolution of the SGLD matrix,

//                 eg. for an output_bits=0, the SGLD matrix will be
at a
resolution of input_bits

//                 eg. for a output_bits=1, the SGLD matrix will be
at a resolution of input_bits-1 etc.

//

// d) d: the distance between pixels for calculating the SGLD
matrix, it can
 take values between 0 and the min(size(image))-1

//

// e) theta: the anlge at which the SGLD matrix is calculated,

//           can take on values of 0 or 45 or 90 or 135 degrees

//

// The outputs of the function are:

// a) SGLD: the SGLD matrix

// b) energy,entropy,maximum probability,correlation,inertia

//


=====================================================================

// haralick(theta,d)

//
=====================================================================


function texture=haralick(theta,d)



// Input parameters

    input_bits=3;

    output_bits=0;
```

```
    // getting the size of the input image

    im_final=floor(double(image)/(2^output_bits));

    [im_final_x im_final_y]=size(im_final);


    // setting the size of co-occurence matrices depending on the
gray level depth

    CO_size=2^input_bits/(2^output_bits);

    SGLD=zeros(CO_size,CO_size);


    select theta

       case {0}

         for i=1:im_final_x

           for j=1:(im_final_y-d)

SGLD(im_final(i,j)+1,im_final(i,j+d)+1)=SGLD(im_final(i,j)+1,im_fina
l(i,j+d)+1)+1;

           end

         end


       case {45}

          for i=(1+d):im_final_x

            for j=1:(im_final_y-d)

              SGLD(im_final(i,j)+1,im_final(i-
d,j+d)+1)=SGLD(im_final(i,j)+1,im_final(i-d,j+d)+1)+1;

            end

          end


       case {90}

          for i=(1+d):im_final_x

            for j=1:im_final_y
```

```
            SGLD(im_final(i,j)+1,im_final(i-
d,j)+1)=SGLD(im_final(i,j)+1,im_final(i-d,j)+1)+1;

          end

        end


      case {135}

        for i=(1+d):im_final_x

          for j=(1+d):im_final_y

            SGLD(im_final(i,j)+1,im_final(i-d,j-
d)+1)=SGLD(im_final(i,j)+1,im_final(i-d,j-d)+1)+1;

          end

        end

    end


    // make the SGLD matrix symmetric by adding it's transpose to it

    SGLD=SGLD+SGLD';


    //  normalize the SGLD matrix to values between 0 and 1

    SGLD=SGLD/sum(sum(SGLD));


    // *****************************************************

    // Calculating the texture features from the SGLD matrix

    // *****************************************************


    // Energy:

    energy=sum(sum(SGLD.*SGLD));


    // maximum probability

    maxprob = max(SGLD);
```

```
// Correlation

[m,n]=size(SGLD);

px=sum(SGLD,2);

[i,j,v]=find(px);

mu_x=sum((i-1).*v);

sigma_x=sum((((i-1)-mu_x).^2).*v);


py=sum(SGLD,1);

[i,j,v]=find(py);

mu_y=sum((j-1).*v);

sigma_y=sum((((j-1)-mu_y).^2).*v);


[i,j,v]=find(SGLD);

corre=(sum(((i-1)-mu_x).*((j-1)-
mu_y).*v))/sqrt(sigma_x*sigma_y);


// Entropy

entrop=0;

for i = 1:m

    for j = 1:n

        if SGLD(i,j)>.0 then

            entrop = SGLD(i,j)*log2(SGLD(i,j))+entrop;

        end

    end

end

entrop=-entrop;
```

```
    // Inertia:
    [i,j,v]=find(SGLD);
    inertia=sum((((i-1)-(j-1)).*((i-1)-(j-1))).*v);
    //texture=[energy,entrop,maxprob,corre,inertia];
    texture=[corre,inertia];
    return texture;
endfunction

// Input parameters

sizeA = 144  //total = 144    155
sizeB = 177  //total = 177    170
sizeC = 53   //total = 53     53
band= 32
th1=0
dist1=1
th2=45
dist2=2

// Imput image type A
for k=1:sizeA,
    num=string(k);
    adress = "C:\Users\Carlos\ImageSource\A"+num+".bmp"
    image=imread(adress);
    // Normalize image gray scale from 0 to 8
    image=rgb2gray(image);
    image =  image/band;
    // Haralick descriptor for theta = 0 and distance = 1
```

```
    texture=haralick(th1,dist1);
    //descriptor(k,1:5)=texture;
    descriptor(k,1:2)=texture;
    // Haralick descriptor for theta = 90 and distance = 2
    texture=haralick(th2,dist2);
    //descriptor(k,6:10)=texture;
    descriptor(k,3:4)=texture;
    alvos(k)=-.5;
end

// Imput image type B
for k=1:sizeB,
    num=string(k);
    adress = "C:\Users\Carlos\ImageSource\B"+num+".bmp"
    image=imread(adress);
    // Normalize image gray scale from 0 to 8
    image=rgb2gray(image);
    image =  image/band;
    texture=haralick(th1,dist1);
    //descriptor(k+sizeA,1:5)=texture;
    descriptor(k+sizeA,1:2)=texture;
    // Haralick descriptor for theta = 90 and distance = 2
    texture=haralick(th2,dist2);
    //descriptor(k+sizeA,6:10)=texture;
    descriptor(k+sizeA,3:4)=texture;
    alvos(k+sizeA)=0;
end
```

```
// Imput image type C
for k=1:sizeC,
    num=string(k);
    adress = "C:\Users\Carlos\ImageSource\C"+num+".bmp"
    image=imread(adress);
    // Normalize image gray scale from 0 to 8
    image=rgb2gray(image);
    image =  image/band;
    texture=haralick(th1,dist1);
    //descriptor(k+sizeA+sizeB,1:5)=texture;
    descriptor(k+sizeA+sizeB,1:2)=texture;
    // Haralick descriptor for theta = 90 and distance = 2
    texture=haralick(th2,dist2);
    //descriptor(k+sizeA+sizeB,6:10)=texture;
    descriptor(k+sizeA+sizeB,3:4)=texture;
    alvos(k+sizeA+sizeB)=.5;
end


plot(descriptor)



save('haralick_dados.dat','descriptor','alvos');
```