An Initial Study in Wood Tomographic Image Classification using the SVM and CNN Techniques

Antonio Alberto Pereira Junior\textsuperscript{a} and Marco Antonio Garcia de Carvalho\textsuperscript{b}

School of Technology, University of Campinas, R. Paschoal Marmo 1888, Limeira/SP, Brazil

Keywords: Ultrasound Tomography, Wood Internal Pathologies, Image Interpolation, Data Augmentation, Classification.

Abstract: The internal analysis of wood logs is an essential task in the field of forest assessment. To assist in the identification of anomalies within wood logs, methods from the Non-Destructive Testing area can be used, as the acoustic methods. The ultrasound tomography is an acoustic method that allows to evaluate the internal conditions of wood logs, through the analysis of wave propagation, without being necessary to cause damage to the specimen. The images generated by ultrasound tomography can be improved by spatial interpolation, i.e., estimating the values of wave propagation not measured in the initial examination. In this paper we present an initial study of classification techniques in order to identify tomographic images with anomalies. In our approach we consider three different classifiers: k-Nearest-Neighbor (k-NN), Support Vector Machine (SVM) and Convolutional Neural Network (CNN). Experiments were conducted comparing them by means of metrics obtained from the confusion matrix. We build a dataset with 5000 images using data augmentation process. The quantitative metrics demonstrate the effectiveness of CNN when compared with k-NN and SVM classifiers.

1 INTRODUCTION

Non-Destructive Tests (NDTs) are studied as they allow the evaluation of the specimen while maintaining its integrity. In woods, the use of non-destructive testing is interesting because it allows to decide which information is needed to characterize each wood and to know how to use the information to explain the behavior of the wood (Bucur, 2006).

One of the most used NDT technique is the ultrasound tomography. The ultrasound tomography allow the evaluation of the internal condition in woods by measuring the propagation time of the ultrasonic pulse, which does not cause any damage to the object. Usually, the images obtained by the ultrasound tomography don’t describe exactly the image regions associated with the anomaly. For this reason, the presence of an expert is necessary to identify an exact region of the defect based on the tomography image. A strategy used in order to improve the quality of generated images by ultrasound tomography is data interpolation. Some of the main techniques are Inverse Distance Weighting (Shepard, 1968), Ellipse Based Spatial Interpolation (Du et al., 2015), and Path Contextual Analysis (Strobel, 2017).

Alternatively to the visual analysis of tomographic images, there is an increase in the use of digital image processing techniques and data classification (Du et al., 2019)(Mu et al., 2015). However, there are still few works in the literature that address the classification of image regions in wood logs.

The goal of this paper is to propose a classification approach in order to identify images with defects in woods from the images generated by ultrasound tomography. To achieve this goal, we accomplish an initial study using three different techniques: k-Nearest-Neighbor (k-NN), Support Vector Machine (SVM) and Convolutional Neural Network (CNN). The main contributions of this initial study are: (i) enable the use of CNN using a dataset with few images and obtained with data augmentation support; (ii) compare different image classification strategies. Therefore, this work compares the classification techniques by means of metrics calculated from a confusion matrix. Also, as a secondary objective, this work intends to gather data from ultrasound experiments on wooden logs to generate a dataset and make it publicly available.

\textsuperscript{a} https://orcid.org/0000-0002-2304-3068
\textsuperscript{b} https://orcid.org/0000-0002-1941-6036
The reminder of this paper is organized as follows: In Section 2 we address some related works and basic concepts. We describe how a tomographic image is obtained by means of an ultrasound inspection test. A description of our proposed method is addressed in Section 3. We present briefly each one of the classification techniques used in this work. The experiments and results are shown in Section 4. Finally, in Section 5 we presented our conclusions.

2 BACKGROUND AND RELATED WORKS

In this section we introduce basics concepts of ultrasound tomography on woods, as well as we briefly describe data interpolation techniques used in this work. Finally, some related works on wood classification are presented.

2.1 Concepts of Ultrasound Tomography

Tomography is an acoustic non-destructive technique used to analyze the structure and composition of objects (Grangeat, 2010). The ultrasound tomography allows to evaluate the internal conditions of wood logs by means of an ultrasound test, illustrated in Figure 1.

The ultrasound test measures the wave propagation time in the wooden logs, known as Time Of Flight (TOF). Therefore, the acoustic velocity can be determined as the distance between the transducers (path length) is also known.

In addition, the ultrasound test is performed according to a wood representation scheme, which provides a guide for positioning the transducers. The main existing schemes, or inspection grids, are diffraction and straight grids (Secco, 2011). The inspection grids define routes or paths through which wave propagation occurs and are important attributes in the generation of tomographic images.

2.2 Interpolation and Image Creation

In this work, spatial interpolation will be used to estimate the acoustic velocity values of points that do not belong to a diffraction grid route. We select two different methods in order to interpolate the data: the Inverse Distance Weighting and the Path Contextual Analysis.

Inverse Distance Weighting - IDW

The IDW interpolation technique is simple and computationally efficient. The value of a point to be interpolated is given by the weighted average of the known values of its neighborhood (Shepard, 1968). In this case, the weight refers to the distance between the points.

Path Contextual Analysis

In this method, the interpolation is accomplished according to the position of the unknown point, considering two different regions to be analyzed: near to the bark and to the pith wood (Strobel, 2017). The Path Contextual Analysis uses an Ellipse Based Spatial Interpolation Algorithm in order to estimate the velocity values in unknown grid cells that represent unknown points (Du et al., 2015). In this method it is considered that the propagation of the ultrasonic pulse has an elliptical influence zone around the route generated by the positioning of the transducers.

2.3 Related Works

There are several works related to the application of ultrasound tomography in woods. Usually, the works intends to solve problems regarding to the generation of images through ultrasound tomography, improvement of interpolation methods and evaluation of tomographic image quality.

The works related to the investigation of the ultrasound tomography in wood seeks to understand the relationships between: the coupling of transducers in wood, problems related to the anisotropy, signal attenuation, number of measurement points, frequency used in the test, arrangements or meshes for testing and type of transducers ((Socco et al., 2004), (Bucur, 2005), (Lin et al., 2008), (Palma et al., 2018)).

There are works which the authors report the use of algorithms to reconstruct the internal characteristics of the woods. For those studies, spatial interpolation algorithms are studied. In (Zeng et al., ) is presented an approach using affected areas through an ellipse with the same eccentricity. This method is improved by the authors (Du et al., 2015) considering
ellipses with different eccentricities, giving rise to the Ellipse Based Spatial Interpolation (EBSI) method. Finally, (Du et al., 2018) and (Du et al., 2019) present some improvements to the EBSI method. As a way of evaluating wood considering different axis, the author (Feng et al., 2018) performs an experiment with the Radial and Longitudinal axis, and also proposes the interpolation method Velocity Correction Interpolation (VCI).

In general, methods are not quantitatively evaluated in order to verify tomographic image quality. In (Strobel et al., 2018) is presented a quantitative evaluation approach, using the confusion matrix, comparing the tomography image with a ground-truth image.

There are also works that use artificial intelligence (AI) methods to identify defects in wood (Zhu et al., 2009), (Mu et al., 2015), and to reconstruct the internal characteristics of wood (Effendi et al., 2019), (Hansson et al., 2015). However, in these works other inspection techniques are considered, such as CT and X-Ray.

There’s an alternative for the reconstruction of tomography images using the technique of deep learning and contour constraint (DLCC) (Du et al., 2019). In this work the deep learning algorithm was applied to detect the size, texture and limit the contour of defects in the wood.

Studies on ultrasound tomography are taking advantage of AI concepts as a tool to identify defects in wood. However, there are opportunities to be developed due to the low amount of works that approach these techniques. In order to cover this gap, this work intends to implement classification techniques for the identification of tomographic images of wood logs with defects.

3 PROPOSED APPROACH

This section discuss the approach proposed in this paper. The workflow shown in the Figure 2 illustrates a general view of each step of our work: (i) image acquisition; (ii) feature extraction; (iii) classification using k-NN, SVM and CNN; and (iv) performance evaluation.

![Figure 2: Workflow of the proposed approach.](image)

The following subsections explain each step of our proposed approach.

3.1 Dataset and Evaluation Protocol

The image were obtained from the Non-Destructive Laboratory (LabEND) of the University of Campinas - UNICAMP. Table 1 show the wood species and provides a brief description of the wood logs used in this work.

<table>
<thead>
<tr>
<th>Species</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lonchorcarpus</td>
<td>It has artificial circular hollows 5cm in diameter.</td>
</tr>
<tr>
<td>Liquidambar styriicflu</td>
<td>Small area with an early stage of fungal decomposition near to the pith, plus lateral cracks from pith to bark.</td>
</tr>
<tr>
<td>Platanus sp</td>
<td>Most of the wood shows signs of fungus attacks and there are some empty areas caused by termites.</td>
</tr>
<tr>
<td>Centrolobium sp</td>
<td>Contains central hollow due to termite attack.</td>
</tr>
</tbody>
</table>

Due to the limited amount of images, a process of data augmentations is also required to be able to increase the size and the quality of data (Shorten and Khosgoftaar, 2019). This process consists of manipulating an existing image by submitting it to geometric transformations, i.e., rotation and resizing. The addition of noise is another technique used to increase the amount of images, one of the most known techniques is the interference of Salt & Pepper, which consists of adding white and black pixels in a sparse occurrence. Labelling is the last step to create the dataset.

First, it’s necessary to determine an area in the ultrasound tomography image and associate to the “anomaly” class. An initial labeling process had been done in previous work (Strobel, 2017) and served as the reference for the dataset annotation.

Finally, three metrics were employed to evaluate the performance of the classification technique, comparing its results with the ground-truth provided by the labelling process: (1) recall - R; (2) precision - P; and (3) accuracy - Acc. The equations are defined as follows:

\[
R = \frac{TP}{TP + FP} 
\]

\[
P = \frac{TP}{TP + FN} 
\]

\[
Acc = \frac{TP + TN}{TP + TN + FP + FN} 
\]
where TP, TN, FP and FN represents, respectively, true-positive, true-negative, false-positive, and false-negative values.

### 3.2 Feature Extraction

The purpose of this step is to obtain descriptors that are able to represent the texture characteristics of the tomographic image. The following subsections describes the texture descriptor we use in this work: the Gray Levels Co-occurrence Matrix (GLCM) and the Local Binary Pattern (LBP).

#### 3.2.1 Gray Levels Co-occurrence Matrix

Gray Level Co-occurrence Matrix (GLCM) is a method of extracting texture features from grayscale image (Robert et al., 1973). Through each vector of characteristics obtained, another 14 characteristics can be obtained from the generated matrix, among them: Mean, Variance, Entropy, Dissimilarity, Correlation, Homogeneity, Cluster Shade, Cluster Prominence, Sum Entropy, Sum Mean, Entropy Difference, Sum Variance.

#### 3.2.2 Local Binary Pattern

Local Binary Pattern (LBP) is a method to describe the local image pattern (Ojala et al., 1996). LBP assumes that textures can be described by measurements: local spatial patterns and gray level contrast. This method extracts local texture information and sets a threshold for a number of neighbors, in the value of the central pixel in that defined local neighborhood.

### 3.3 Classification Techniques

Classification techniques are use to classify datasets with known definitions of groups. In this case it’s presented to the model samples of desired inputs and outputs, previous defined by an human. The goal of this technique is to learn a general rule that map the inputs and outputs to the overall model.

This work intends to detect the region of the image that has the anomaly. Therefore, classifications models will be used to execute this proposal. We use two supervised algorithms: k-Nearest-Neighbor (k-NN) and Support Vector Machines (SVM). Those two algorithms was chosen due to common use in the classification tasks. Additionally, as a tentative to improve the classification performance, a Convolutional Neural Network (CNN) model is also applied in this work. By the end, we should be able to compare both models and then determine which one is the best to identify anomaly in woods. The following subsection describes each one of the classification techniques.

#### 3.3.1 k-NN

The k-Nearest-Neighbor (k-NN) (Cover and Hart, 1967) is also known as lazy learn or instance-based, which means that the algorithm does not perform the model definition or generalization when training data is received, but during execution time. This analogical learning technique performs peer-to-peer comparisons to identify similarities between input data. To determine which data are closest or similar, it is necessary to describe them using distance notation, i.e., the euclidean distance.

#### 3.3.2 Support Vector Machine

The Support Vector Machines (SVM) (Suykens and Vandewalle, 1999) is a classification technique that can be applied to linearly separable or non-linearly separable data. This technique aims to determine the best hyperplane for class segregation using vectors to support this determination. In this work, we use different kernels to see the performance of each one separately.

#### 3.3.3 Convolutional Neural Network

The Convolutional Neural Network (CCN) is one of the most popular deep neural network. This network can have multiple layer, in particular: the convolution layer, ReLU, pooling layer and fully connected layer (Albawi et al., 2017). The CNN have the ability to train large datasets and considers millions of parameters, because of that, this network is commonly used in area of image recognition, object detection and computer vision.

A CNN performs automatically the step of feature extraction by using the convolution layer. So the model can learn all features in one pass instead of having the features being selected by an engineer. For those aspects, a Convolutional Neural Network model is going to be used in this work to avoid the features extraction step. In contrast, the need for a large amount of training data is an important feature of CNNs. In order to mitigate this requirement, we apply data augmentation techniques in our dataset. The Convolutional Neural Network that will be used in this work is the SDD MobileNetV2 (Sandler et al., 2018). This Convolution Neural Network is available in Tensorflow 2 (Abadi et al., 2015) which uses the object detection approach to detect bounding boxes. This model was chosen due to the speed and to the input size of tomography image available in the dataset.
4 EXPERIMENTS AND INITIAL RESULTS

This section presents the conducted experiments regarding the initial study of different classification techniques to assess the tomographic images. All experiments are performed using an Intel i5 desktop, 16GB RAM running macOS 10.15 Catalina. Our computational implementation was developed in Python 3.0 using Google Colaboratory (Bisong, 2019). We present the results and compare the performance of the techniques with each other.

As mentioned before, a data augmentation process was used to build the entire dataset. At the end of this process, the dataset had 5000 tomographic images and the annotation files related to them.

The first step for the common supervised models is to extract the vectors of features from each tomographic image. To be able to extract those feature vectors it was necessary to get the features that belong to the anomaly and the features that correspond to the healthy wood region. Those vectors generated by the feature extraction are the input for the classification models: SVM and k-NN. The dataset was split in 70% of images for training and 30% for test step.

Our first experiment concerns studying different types of SVM usage configurations. Table 2 shows the results of the SVM model using different types of Kernel. We evaluate the performance of the classification task under the different metrics described in the section 3.1. The best performance for each metric are highlighted in bold.

Table 2: Effectiveness of the SVM technique using different combinations (in bold, the highest value observed in each column).

<table>
<thead>
<tr>
<th>SVM (Kernel) - Feature</th>
<th>Accuracy</th>
<th>Recall</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM (Linear) - LBP</td>
<td>76.00</td>
<td>87.80</td>
<td>73.46</td>
</tr>
<tr>
<td>SVM (Linear) - GLCM</td>
<td>74.66</td>
<td>88.09</td>
<td>72.54</td>
</tr>
<tr>
<td>SVM (Polynomial) - LBP</td>
<td>85.33</td>
<td>92.68</td>
<td>82.60</td>
</tr>
<tr>
<td>SVM (Polynomial) - GLCM</td>
<td>68.00</td>
<td>64.28</td>
<td>75.00</td>
</tr>
<tr>
<td>SVM (RBF) - LBP</td>
<td>86.66</td>
<td>95.12</td>
<td>82.97</td>
</tr>
<tr>
<td>SVM (RBF) - GLCM</td>
<td>81.33</td>
<td>88.09</td>
<td>80.43</td>
</tr>
<tr>
<td>SVM (Sigmoid) - LBP</td>
<td>64.00</td>
<td>60.97</td>
<td>69.44</td>
</tr>
<tr>
<td>SVM (Sigmoid) - GLCM</td>
<td>64.00</td>
<td>66.66</td>
<td>68.29</td>
</tr>
</tbody>
</table>

According to Table 2, the best result of Accuracy, Recall and Precision occurred using the SVM with the Kernel RBF and the LBP feature. The second experiment analyzes the use of the k-NN classifier. Table 3 shows the results of the k-NN model. The values of k was chosen based on a experiment with a small portion of the dataset. This experiment showed limited improvements in the accuracy when k was incremented with a values less than 50 for each run. For this reason the experiment showing in the table be

low for the classifier k-NN is going to consider k=50, k=100, and k=150.

Table 3: Effectiveness of the k-NN technique using different combinations (in bold, the highest value observed in each column).

<table>
<thead>
<tr>
<th>k-NN (k=Value) - Feature</th>
<th>Accuracy</th>
<th>Recall</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>k-NN (k=50) - LBP</td>
<td>78.66</td>
<td>85.36</td>
<td>77.77</td>
</tr>
<tr>
<td>k-NN (k=50) - GLCM</td>
<td>74.66</td>
<td>76.19</td>
<td>78.04</td>
</tr>
<tr>
<td>k-NN (k=100) - LBP</td>
<td>74.66</td>
<td>80.34</td>
<td>75.04</td>
</tr>
<tr>
<td>k-NN (k=100) - GLCM</td>
<td>54.66</td>
<td>45.23</td>
<td>66.33</td>
</tr>
<tr>
<td>k-NN (k=150) - LBP</td>
<td>57.33</td>
<td>43.90</td>
<td>66.66</td>
</tr>
<tr>
<td>k-NN (k=150) - GLCM</td>
<td>49.33</td>
<td>16.66</td>
<td>70.00</td>
</tr>
</tbody>
</table>

In the case of the k-NN classifier, the best result was the one that use LBP considering k=50. Overall, the results shown in Tables 2 and 3 presents good accuracy values, especially those using LBP descriptor. The last classification method addressed in this work was the CNN. In this case, our experiment consists of comparing the best performances of SVM and k-NN, using LBP texture descriptor, with the results obtained by using the CNN SDD MobileNetV2. Table 4 shows the results.

Table 4: Performance comparison (in bold, the highest value observed in each column).

<table>
<thead>
<tr>
<th>Technique</th>
<th>Accuracy</th>
<th>Recall</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM (RBF) - LBP</td>
<td>86.66</td>
<td>95.12</td>
<td>82.97</td>
</tr>
<tr>
<td>k-NN (k=50) - LBP</td>
<td>78.66</td>
<td>85.36</td>
<td>77.77</td>
</tr>
<tr>
<td>SDDMobileNetV2</td>
<td>89.00</td>
<td>93.40</td>
<td>97.30</td>
</tr>
</tbody>
</table>

As we can observe in Table 4, the performance of the SDDMobileNetV2 classifier was better than that obtained by the other classifiers, regardless the used metric, Accuracy, Recall or Precision. Nevertheless, the three first experiments concerns to the task of classify the entire image as having or not an anomaly. That is, the results do not show the specific image region associated to the anomaly. Therefore, in order to identify the anomaly in the image, we propose a last experiment: perform the classification of image regions, obtained from Otsu algorithm (Otsu, 1979).

Figure 3 shows the results of the region classification process using the SVM plus LBP feature (a), and for the SDDMobileNetV2 classifier (b). In these cases, the presence of an internal defect in the central region of the wood is observed. This information is important from a wood usage classification point of view, for commercial purposes.

5 CONCLUSIONS

In this work we have presented an initial study about the use of data classification techniques in wood to-

579
Figure 3: Classification of the anomaly image region using two different techniques for the *Platanus* sp. wood: (a) SVM with LBP descriptor (region contour in green); (b) CNN with SDDMobileNetV2 architecture (bounding box area).

mography images. Our dataset consists of images obtained from ultrasound tomography, a non-destructive method capable of evaluating the wood log internal characteristics without causing any damage to it.

In order to identify whether or not an image has anomalies, we applied three different image classification methods: k-NN, SVM and CNN. The performance of these methods was evaluated according to Accuracy, Precision and Recall metrics, computed from a confusion matrix built based on the annotated images. We also performed an image region classification task, in order to obtain the region corresponding to the wood anomaly.

Our first experiments showed that the best results are obtained by the CNN classifier, regardless the metric. The accuracy, precision and recall values are higher than 85%. A last experiment carried out in this work was dedicated to identifying the region in the image associated with the internal defect.

Our contribution is also associated with the creation of a dataset with about 5000 images using data augmentation techniques. Now, our efforts will be towards characterizing and balancing the dataset, avoiding possible biases.

There are several possible suggestions as future works. First one, in order to improve the variability of the dataset, we need more wood tomographic images, and different species and anomalies.

The use of texture descriptors of different types, such as those provided by Fourier and Wavelet Transforms, should be the object of further studies. We also intend to combine different descriptors in order to verify the classification performance.

In addition to classifying the wood image as healthy or with an internal defect, we would like to properly identify the anomaly, its location and dimensions.

ACKNOWLEDGEMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

REFERENCES


