**CHAPTER TWO**

**LITERATURE REVIEW**

* 1. **Texture Classification**

Texture classification is an important task in image processing and computer vision, with a wide range of applications, such as computer-aided medical diagnosis (Khademi and Krishnan, 2008; Harms et al., 1986; Sutton and Hall, 1972) classification of forest species (Tou et al., 2009), classification in aerial/satellite images (Nasirzadeh et al., 2010), writer identification and verification (Bertolini et al., 2012) and music genre classification (Costa et al., 2012). Texture classification commonly follows the standard procedure for pattern recognition, as described by Bishop et al. in (Bishop et al., 2006): extract relevant features, train a model using a training dataset, and evaluate the model on a held-out test set.

Many methods have been proposed for the feature extraction phase on texture classification problems, as reviewed by Zhang et al. (Zhang and Tan, 2002) and tested on multiple datasets by Guo et al. (Guo et al., 2012). Noteworthy techniques are: Gray-Level Co-occurrence Matrices (GLCM), the Local Binary Pattern operator (LBP), Local Phase Quantization (LPQ), and Gabor filters.

The methods above rely on domain experts to build the feature extractors to be used for classification. An alternative approach is to use models that learn directly from raw data, for instance, directly from pixels in the case of images. The intuition is using such methods to learn multiple intermediate representations of the input, in layers, in order to better represent a given problem. For an example, an object recognition in an image; the inputs for the model can be the raw pixels in the image. Each layer of the model constitutes an equivalent of feature detectors, transforming the data into more abstract (and hopefully useful) representations. The initial layers can learn low-level features, such as detecting edges, and subsequent layers learn higher-level representations, such as detecting more complex local shapes, up to high-level representations, such as recognizing a particular object (Bengio, 2009). In summary, the term Deep learning refers to machine learning models that have multiple layers, and techniques for effectively training these models, commonly building Deep Neural Networks or Deep Belief Networks (Hinton et al., 2006).

Methods using deep architectures have set the state-of-the-art in many domains in recent years, as reviewed by Bengio in (Bengio, 2009) and (Bengio and Courville, 2013). Besides improving the accuracy on different pattern recognition problems, one of the fundamental goals of Deep Learning is to move machine learning towards the automatic discovery of multiple levels of representation, reducing the need for feature extractors developed by domain experts (Bengio and Courville, 2013). This is especially important, as noted by Bengio in (Bengio, 2009), for domains where the features are hard to formalize, such as for object recognition and speech recognition tasks.

**2.1.1 Texture Recognition in Object Classification**

In the task of object’s texture recognition, deep architectures have been widely used to achieve state-of-the-art results, such as in the CIFAR dataset (Krizhevsky and Hinton, 2009) where the top published results use Convolutional Neural Networks (CNN) (Ciresan et al., 2012). The tasks of object and texture classification present similarities, such as the strong correlation of pixel intensities in the 2-D space, and present some differences, such as the ability to perform the classification using only a relatively small fragment of a texture. In spite of the similarities with object classification we observe that deep learning techniques are not yet widely used for texture classification tasks. Kivinen and Williams (Kivinen and Williams, 2012) used Restricted Boltzmann Machines (RBMs) for texture synthesis, and Luo et al. (Luo et al., 2013) used spike-and-slab RBMs for texture synthesis and in painting. Both consider using image pixels as input, but they do not consider training deep models for classification among several classes. Titive et al. (Titive and Bouzerdoum, 2006) used convolutional neural networks on the Brodatz texture dataset, but considered only low resolution images, and a small number of classes.

The similarities between texture classification and object recognition, and the good results demonstrated by using deep architectures for object recognition suggest that these techniques could be successfully applied for texture classification.

**2.1.2 Motivation for Texture Classification**

There is a considerable set of potential applications for texture recognition, as briefly presented above. However, despite the reported success of classical texture classification techniques in many of these tasks, these problems are still not resolved and are subject of active research, with potential to increase recognition rates.

A second motivation is that traditional machine learning techniques often require human expertise and knowledge to hand-engineer features, for each particular domain, to be used in classification and regression tasks. It can be considered that the actual intelligence in such systems is therefore in the creation of such features, instead of the machine learning algorithm that uses them. Therefore, using techniques that do not rely on expert-defined feature extractors can make it easier to develop effective machine learning models for novel datasets, without requiring the test and selection of a large set of possible feature extractors.

**2.1.3 Challenges of Texture Classification**

Here is a set of perceived challenges in applying techniques and algorithms to execute texture classification problems:

* Image size: The majority of the image-related tasks where deep learning was successfully applied used images of small size. Examples: MNIST (28x28 pixels), STL-10 (96x96), Norb (108x108), Cifar-10 and Cifar-100 (32x32).
* Model size and Training time: One exception to dataset list above is the ImageNet dataset, which consists in high-resolution images of variable sizes. The best results on this dataset, however, require significant usage of somehow impossible to get kind of computer resources (such as 16 thousands cores running for three days (Le, 2013). The texture datasets commonly consist of higher resolution images, and therefore different techniques need to be tested, in order to classify the textures without using too much computing resources.

**2.1.4 Objectives of Texture Classification**

The main objective of texture classification is to test whether or not models, such as Convolutional Neural Networks, Support Vector Machine etc can be successfully applied for texture classification problems. More specifically, these methods are to be tested on multiple texture datasets, developing a method to cope with the high-resolution texture images without requiring models that are too large, or that require too much computing power to train. Six texture datasets were selected for testing, representing different domain problems, and containing different characteristics, such as image sizes, number of classes and number of samples per class. As part of this effort, we assess if it is possible to obtain a generic framework that brings good results for multiple texture problems.

Training the models on each dataset is of great essence, the accuracy of the deep models are compared with the state-of-the-art results achieved using the classical texture descriptors. Finally, another objective of this research is to evaluate a method of Transfer Learning for texture classification. Transfer Learning consists in using a model trained in one task to improve results on another task. This is particularly interesting when using Artificial Neural Networks or Support Vector Machine, as these models often require large datasets to be effectively trained.

**2.2 Classification of Wood**

Wood species recognition is a relatively new problem to be solved using computer vision techniques. The texture classification techniques have been proven to be useful to solve several real world problems, such as rock texture classification (Partio et al., 2002), face detection (Khalid et al., 2007), and wood species recognition (Lew, 2005; Tou et al., 2007). This is accomplished due to the property of the cross section surface of trees that has a pattern for different species. Therefore, by inspecting the patterns on the cross section surface, the species of the tree can be determined.

Recognizing forest species is an important task in many areas. In the construction industry, it is important to validate that the correct species is being used for a given construction, to ensure that the properties of the material are known. The manufacturing process of wood products, such as tables and chairs, may require a particular type of wood. In commerce, identifying the species is important for valuing a product, and for inspection to control the illegal trade of rare species, which is an issue in many countries. Considering that these tasks generally require a human expert, the development of an automated system could lower cost and make this process faster. For this reason, several systems have been proposed in the literature for forest species recognition (Tou et al., 2009; Nasirzadeh et al., 2010; Filho et al., 2010; Filho, 2012).

Wood is a useful raw material for furniture, railway sleepers, tools, shipping and construction industry, etc. Researchers from different domains such as paleontologists, archaeologists, forensic experts and art historians may be interested in the identification of wood (Wheeler & Baas 1998). Wood is broadly classified as hardwood and softwood species. Softwood species are conifers and 90.00–95.00% of their cells (called longitudinal tracheid) have simple cellular structure making it difficult to discriminate amongst themselves due to limited number of cell types. On the contrary, hardwood species (angiosperm) possess complex cellular structure and are clearly distinguishable among intra-species. Anatomical characteristics like vessels, fibers, parenchymas and rays play significant role in hardwood species identification (Hermanson and Wiedenhoeft, 2011). The unique cellular structure of each of the hardwood species varies widely among intra-species and serves as a signature for their identification (Bond, 2002).

It is necessary to identify wood as its characteristics vary widely (Wang et al 2013). Accurate recognition of wood species is essential for price fixation based on color, texture, scent, hardness, durability, availability and rational use of available resources. This would also help in avoiding deception by timber traders.

**2.3 Methods of Wood Classification**

There are two methods normally used for wood identification, viz., traditional approach and machine vision techniques. For many decades, traditional approaches have been instrumental in wood identification, like using 10× hand lenses to analyze the surface of the wood specimen in conjunction with their color, scent, hardness and weight. However for more reliable results, cross-sectional micro-structures of the wood samples are analyzed in the laboratory and their features are compared with available samples of hardwood species for identification (Baas et al 1989).

The unavailability of xylarium micro slides and literatures for comparing the microstructures of unidentified wood samples with known samples, and dearth of competent manpower in this field is the key challenge in wood identification (Bond, 2002). Further, imparting training to wood identification officers to get an expertise in identification of wood is a time-consuming process (Chang and Lin, 2011). Also, occupation as a wood certification officer is neither easy nor lucrative and possibility of unfairness and oversight cannot be denied. In the absence of systematic classification procedure for wood identification, a species has to be identified based on the combination of its microstructure features (Cavalin, 2013).

There is massive hardwood diversity in tropical countries. India alone has over 1,200 hardwood species and, thus, remembering the microstructure of all the species is next to impossible. Recognition of large volumes of wood species, employing traditional approach is prolonged, erroneous and unfeasible sometimes (Chandrashekar and Sahin, 2014). Therefore, to efficiently deal with the above mentioned issues researchers are looking into the possibility of coming up with computer assisted forest species/hardwood species identification system. Machine vision based systems outperform manual classifications when large volume of wood species is to be identified repeatedly with utmost accuracy, without getting fatigued (Breiman, 2001). Thus, machine vision based wood identification techniques have emerged as an alternative in overcoming deficiencies associated with traditional approaches. The primary motivation behind employing machine vision based wood identification/classification system is the limitations associated with traditional approaches of wood identification (Bremananth et al., 2009).

**2.4 Wood Texture Recognition and Classification Techniques**

Forest species recognition has been generally treated as a texture classification problem, due to the property that the cross section surface of trees has different patterns on different species (Tou et al., 2009). Texture classification techniques have been explored by several authors in recent years. Tou et al (Tou et al., 2009) investigated the usage of Gabor filters and co-occurrence matrices (GLCM). Khalid et al (Khalid et al, 2007) studied the usage of Local Binary Patterns (LBP) for extracting relevant features from the images, and used a K nearest-neighbour (KNN) classifier with promising results. Paula et al. investigated the usage of color-based features and GLCM in (Filho et al., 2010), and the combination of different classifiers using GLCM, LBP, CLBP and color features in (Filho et al., 2010).

Deep learning models have been receiving increased attention in recent years. These methods are frequently setting the state-of-the-art in many domains, as reviewed by Bengio in (Bengio and A. Courville, 2013). Besides improving the accuracy on different pattern recognition problems, one of the fundamental goals of Deep Learning is to move machine learning towards the automatic discovery of multiple levels of representation. The intention is to use raw data (e.g. image pixels) as input to the models, and let the models learn intermediate representations - that is, let the model learn the feature detectors (Bengio and A. Courville, 2013). This is especially important, as noted by Bengio, for domains where the features are hard to formalize, such as for object recognition and speech recognition tasks. In the task of forest species classification, several alternative feature extractors have been used (as stated above), demonstrating the difficulty of finding a good representation for the problem.

The popular groups of texture classification techniques that are used are the statistical and signal processing methods (Kivinen and Williams, 2012). Both are used in that paper, i.e. GLCM, Gabor filters, combined GLCM and Gabor filters and covariance matrix. The classifier used is the k-nearest neighbor (k-NN) and a two-stage verification based recognition process.

**2.4.1 GLCM**

This is a popular statistical texture classification technique ever since it is introduced by Haralick et al. back in 1973 (Haralick et al., 1973) because it is computationally simple yet useful for many texture classification problems. The GLCM calculates the occurrence of pixel pairs within the images according to the spatial distance between the pixel pairs and orientations provided (Petrou, and Sevilla, 2006). The computed GLCM can be used as a feature after it is down-sampled which we named as raw GLCM in that paper (Tou et al., 2009). A second-order feature can be obtained from the GLCM. There are 5 commonly used textural features, i.e. contrast, correlation, energy, entropy and homogeneity (Petrou, and Sevilla, 2006).

**2.4.2 Gabor Filters**

This is a signal processing method, therefore it processes on the frequency domain rather than the spatial domain (Khalid et al., 2007). In that paper, the Gabor filters are generated by using different three radial center frequencies and eight orientations as used in (Tou et al., 2007). The convolution is performed by applying fast Fourier transform (FFT), point-to-point multiplication and inverse fast Fourier transform (IFFT) (Khalid et al., 2007). Due to the complexity of the features produced, the Gabor filters are down-sampled and the singular value decomposition (SVD) is further used to reduce the dimensionality of the feature set (Tou et al., 2007).

**2.4.3. Combined GLCM and Gabor Filters**

Different techniques are often combined to be used and can produce better results compared to using them individually. In these papers (Tou et al., 2007; Bala, 1990; Recio et al., 2005; Umarani et al., 2007), the GLCM feature and the Gabor filters were combined by appending both of them into a single feature set. This combination produces a better result than either GLCM feature or Gabor filters when it is applied on the 32 textures from the Brodatz texture dataset (Tou et al., 2007).

**2.4.4 Covariance Matrix**

The covariance matrix is a statistical method that calculates the covariance between values. In the paper, the covariance matrix is used to calculate between images which are named as feature images. The feature images are a set of two-dimensional images or matrices generated by a feature extraction algorithm, such as the GLCM and Gabor filters. In the paper, Gabor filters are used to generate the feature images because it performs better than edge-based derivatives and GLCMs as feature images in (Tou et al., 2008).

**2.4.5 k-NN**

The k-nearest neighbor (k-NN) is used as the classifier in that paper. The k-NN will compare the feature set of the test sample against all the training samples and select the k samples with the shortest distance. The distance metric used is the Euclidean distance for all of the techniques described in Section 2.1 to 2.3 except for the covariance matrix where the Forstners and Moonen’s distance is used (Tuzel et al., 2006). This is because the covariance matrix does not lie on the Euclidean space.

**2.4.5. Verification-based Recognition**

This method is used as a two-stage classifier that first goes through a verification process before going through a recognition process. The feature extraction that is used in the paper is GLCM because of its simpler computations. The GLCMs are generated in eight directions to achieve rotational invariant as proposed in (Tou et al., 2009). The test sample will be tested against all training templates. Each training template will accept the test sample as the same species when the distance is determined to be lower than the threshold value defined. Species with the highest number of accepted templates will be selected as the winning class (Tou et al., 2009).

**2.4.6 Summary of Other Techniques**

Deep architectures have been widely used to achieve state-of-the-art in object recognition tasks, such as the CIFAR dataset (Krizhevsky and Hinton, 2009) where the top published results use Convolutional Neural Networks (CNN) (Ciresan et al., 2012). The tasks of object and texture classification present similarities, such as the strong correlation of pixel intensities in the 2-D space, and present some differences, such as the ability to perform the classification using only a relatively small fragment of a texture. In spite of the similarities with object classification and we observe that deep learning techniques are not yet widely used for texture classification tasks. Kivinen and Williams (Kivinen and Williams, 2012) used Restricted Boltzmann Machines (RBMs) for texture synthesis, and Luo et al. (Luo et al., 2013) used spike-and-slab RBMs for texture synthesis and in painting. Both consider using image pixels as input, but they do not consider training deep models for classification among several classes. Titive et al. (Titive et al., 2006) used convolutional neural networks on the Brodatz texture dataset, but considered only low resolution images, and a small number of classes.

Applying deep learning techniques to forest species recognition represent a contribution not only to this specific problem, but for texture recognition problems in general. With this approach we could learn the most appropriate feature representation for textures from data. However, one characteristic of the datasets used for this task (and also other datasets containing textures) is the high resolution of the images, in contrast with the low resolution of most object recognition databases to which deep learning has been applied. As a consequence, questions such as how to adapt the existing CNN architectures for these images and how to keep the training time acceptable are a matter of concern.

**2.5 Support Vector Machine and Artificial Neural Network for Texture Classification**

**2.5.1 Support Vector Machine**

Now a days, there are vast amount of data being stored in databases across the globe. Data mining offers promising ways to uncover hidden patterns from such amount of data. These hidden patterns can probably be used to predict future behaviour. Classification in the data mining is one of the tasks to uncover these hidden patterns. As it is known, kernel method is used in SVM for pattern analysis and detects the types of relations. The input for the classification is the training datasets, whose class labels are already known. Many approaches have been introduced to solve the classification problem. SVM is considered as one of the most robust and efficient methods among all well-known algorithms for classification. It is the powerful algorithm used for supervised learning, and is widely used in classification problems (Romero and Toppo, 2007; Maji et al., 2008). The major limitation of SVM is its low speed in the training and the test phases. To overcome this limitation, several researches have been proposed. Several researches have been done to reduce the computational cost by reducing the number of support vectors directly. But these efforts had not been able to give high accuracy. However, some other approaches have been used to cope with this problem. Neural Network is one of the approaches to solve this problem.

In data mining, classification is one of the most popular tasks with wide variety of applications. Many algorithms have been presented to produce an accurate and efficient classifier. All these algorithms are worked on single table as an input but in real world applications, data is stored on multiple tables. There have been many techniques for classification as Neural Networks and Support Vector Machines. However, they can only be applied to data in single table. The conversion of the multiple tables to single table is very difficult and expensive. So, Multi-relational classification which uses weighted voting technique can be applied to combine classifiers to get class label based on the contribution of tables (Chen and Chen, 2004; Romero and Toppo, 2007).

SVM classifier is a powerful classifier for the classification task of data mining. SVM is based on statistical learning theory and used to find the optimal separating hyper-plane between two classes. Optimal hyper-plane is the one giving maximum margin between training examples of different classes. SVM converts the original data point to dimensional space and the data point is viewed as a dimensional vector. The main aim of the SVM is to predict which class a new data point will be in. There are many hyper-planes that may classify the data. The best hyper-plane is the one which has largest separation margin between two classes. We choose this hyper-plane because the distance from the nearest point on the each side is maximized. This hyper-plane is known as maximum-margin hyper-plane or optimal hyper-plane. By this hyper-plane, the linear classifier is defined known as maximum margin classifier.

As it is known, kernel method is used in SVM for pattern analysis and detects the types of relations. Kernel SVM requires evaluating the kernel for a test vector and each of the support vectors. Intersection kernel SVM and additive kernel SVM are introduced in (Vapnik, 1999), which are independent of number of support vectors and gives more efficient classification than kernel SVM. SVM separates the class by parallel hyperplane but in some problems, the parallel hyperplane cannot be able to separate the classes. So, the concept of nonparallel hyperplane is introduced in (Gao and Zhou, 2009). Originally SVM was designed for binary classification but in the real world, the problems have multiple classes. This problem can be solved by constructing and combining several binary classifiers. (Chen and Chen, 2004) used multiple simple classifiers to approximate the SVM for classification problem.

The SVM is one of the best techniques and also gives high accuracy and efficiency compared to other techniques of classification. As the problem become complex, the computational cost of SVM increases. The major drawback of SVM is its computational cost. There are many approaches which can be used to reduce the cost. Decision tree can be applied to speed up the SVM in test phase. This approach has focused on reducing the number of test datapoints to be used in classification (Andral, 2002). Another approach is based on adaptive genetic algorithm to optimally reduce the solutions for SVM by selecting vectors from the trained dataset. These datasets consist of the support vectors which best approximates the original discriminant function (Hsu and Lin, 2002).

**2.5.2 Artificial Neural Network**

ANN can also be used to improve SVM’s efficiency. ANN consists of three layers: the input layer, the hidden layer, ant the output layer the input layer take the inputs, then forward to hidden layer (one or more) and finally to output layer to produce decision function. (Gayathri and Kumarappan, 2010). ANN can be used such as multilayer perceptron, recurrent neural network, radial bases function neural network, etc. These neural networks are applied to SVM to speed up the process of classification. The complexity of NN in the learning is generally less compared to that of SVM. When neural network is applied on the test phase of SVM, it reduces the complexity of SVM by approximating the number of support vectors. When NN is applied to training phase, pre-partitioning of the set of support vectors can be done to reduce the time complexity. (Kumar and Gopal, 2010).

**2.6 Related Works**

Nasirzadeh et al (2010) used LBP variants for feature extraction and NN as a classifier. The LBP-HF has reported 96.60% accuracy compared to 91.00% recognition accuracy obtained by the traditional LBPri.

Yusof et al (2010) used combined features obtained by Gabor filter and GLCM techniques from macroscopic images and achieved a recognition rate of 90.33% for test dataset with MLP-ANN classifier.

Further, Khairuddin et al (2011) employed BGLAM and SPPD to extract features. Before applying these features to the final classification stage, pre classification was carried out by K-means clustering followed by dimensionality reduction using LDA and KDA/GSVD. This scheme has accounted for a classification accuracy of 96.15% with K-NN classifier. The statistical features, namely, mean, SD, entropy and contrast have been obtained from wood stereogram images using a Gabor filter bank. The fusion of entropy, mean and SD features classified by NN classifier shows 94.58% recognition accuracy (Wang et al 2012).

Yusof et al (2013a, b) proposed a kernel genetic algorithm for selection of nonlinear features of macroscopic images of tropical wood species obtained by GLCM, BGLAM and SPPD techniques. This approach has brought in dimensionality reduction, and reported a classification accuracy of 98.69% with LDA classifier. Subsequently, a pre-classifier approach that uses fuzzy-logic concept was introduced by Yusof et al (2013a, b) with an aim to cluster the database. The experimental outcome reported around 93.00% accuracy compared to 88.90% accuracy obtained without incorporating fuzzy-logic pre-classifier.

The ACA was employed by Ahmad & Yusof (2013) to train and test the feature dataset obtained by BGLAM and SPPD techniques. A recognition rate of 96.75% was reported for experimental work performed using 24 clusters. Wang et al (2013) employed a MMI feature extraction technique to obtain features of wood stereogram images. Further, K-NN and SVM classifiers were investigated and a classification accuracy of 86.53% was achieved for statistical features of MMI with SVM classifier.

Martins et al (2013) presented a database of microscopic images of forest species that consist of 37 softwood and 75 hardwood species. Further, they obtained structural, GLCM and LBP features of these species, and a recognition accuracy of 98.60% and 86.00% was reported for two-class and multi-class (112) classification using SVM classifier.

Afterwards, Cavalin et al (2013) extracted GLCM, LBP, and LPQ features from the images obtained by quad-tree decomposition method. The fusion of GLCM and LPQ features reported a recognition accuracy of 93.20% using SVM classifier. The GLCM features obtained from Gabor wavelet images of hardwood species produced 92.60% classification accuracy with MLP-BP-NN classifier (Yadav et al 2013).

Subsequently, Yadav et al (2014) acquired Coiflet DWT based features of the hardwood species. A recognition accuracy of 92.20% was reported for these features using MLP and logistic classifiers with 10-fold cross validation approach.

Paula-Filho et al (2014) proposed a two level divide-and-conquer classification strategy to categorize the macroscopic images of 41 species using SVM classifier. Feature set obtained by a combination of several feature extraction techniques was classified with six number of classifiers and reported best recognition accuracy of 97.77 %.

**2.7 Deductions from the Related Works**

A comprehensive analysis of the reviewed work shows that different texture features and classification algorithms are used for wood species identification (macroscopic, microscopic and stereogram images). The review discloses the fact that the texture feature information is extracted from single scale (original grayscale) images that could not acquire all the significant information of the images helpful in the efficient classification of hardwood species. Pan and Kudo (2012) have stated that microscopic images of wood carry significant information useful in its precise identification as compared to limited information possessed by macroscopic images. Hence, most works utilized microscopic images for hardwood species identification.

Further, the classification accuracy of hardwood species can be improved either by incorporating appropriate texture feature extraction technique capable of obtaining discernible features of the image or by using suitable classifier. However, the classifiers performance highly depends on the quality of its input features. Thus, the task of the viable texture classification is the correct extraction of the significant texture features of the microscopic images of hardwood species for their efficient classification. The discrete Cosine Transform (DCT) has been proposed in this final year project due to its multi resolution capability for analyzing images at different frequencies for several levels of resolutions according to (Mallat, 1989). In the light of the above reasoning, the DCT based comparative texture feature extraction techniques (SVM and ANN) capable of acquiring significant features of hardwood species has been proposed in this project work.