Grading Tea leaves in Low Country Region of Sri Lanka using Deep Learning Techniques

Olivia George
Research Department
Amazon
olivia.george@amazon.com

Isha Choudhary
Research Department
Amazon
shahooda637@gmail.com

***Abstract*—Identifying tea leaves according to their quality is a vital cog in the tea production pipeline. This identification process includes categorizing tea leaves into three categories known as Best, Below-best, and Refused-tea. During tea production, this identification process is done manually through experts' opinions. The manual identification is done in 3 steps and involves counting of individual leaves manually. The identification by experts is not always accurate when you consider the flaws that come with their human factor such as fatigue and personal biases. The literature reveals that limited work has been carried out focusing on tea leaf classification. The majority of work has been conducted on identifying tea leaves that get infected with diseases. The literature further unveils that the research that falls into the domain of tea leaf classification has mostly been done offshore, where the outlook of a tea leaf is different from that of tea leaves grown in Sri Lanka. It is very important to conduct the research and discover the results in the area we are aiming to grow the tea garden as the difference in climatic conditions can lead to different results . In this work, a novel deep learning model is proposed in order to classify tea leaves into the three categories using deep convolutional neural networks (CNN). The proposed network is specifically designed for and experimented on grading the leaves in low country region of Sri Lanka. A dataset of 336 tea leaf images was collected for the research, and the model was trained using 216 tea leaf images after performing image augmentation on the original dataset. Advanced CNN architectures, AlexNet, VGG16 and ResNet-50 model with transfer learning were implemented for performance comparison with the proposed architecture. AlexNet and VGG16 models did not perform well on classifying tea leaves whereas the ResNet-50 model with transfer learning reached an average accuracy of 96.53%. The architecture of the proposed CNN network consists of 3 convolutional layers and 2 max-pool layers resulting in the best accuracy for the dataset. The proposed CNN model is trained with the data with 3 different architectures that have 8, 10 and 12 layers. The best accuracy is found for the model with 12 layers. The evaluation of the proposed CNN model revealed that the model was able to reach an average accuracy of 93.81% for the three classes. Smaller prediction time combined with a simple layering structure made the proposed novel CNN model much more suitable for the classification problem than the advanced CNN models. The final evaluation shows that the proposed CNN architecture with 12 layers and ResNet-50 with transfer learning show the best accuracy for classifying the given leaf image as Best, Below-best or Refused tea.**

***Index Terms*—Deep Learning, Tea Grading, Convolutional, Neural Networks, Transfer Learning**

I. INTRODUCTION

Tea leaves that are grown in Sri Lanka are grouped into three categories according to the region that they are grown in [Tea, 2020]. Low-grown tea from sea level to 600m, medium-grown tea from 600m to 1200m and high grown tea from 1200m upwards. Tea leaves grown in these three regions show vastly different characteristics mainly due to the environmental factors that change with different climates. Tea leaves grown in Low country regions of Sri Lanka (i.e., low grown tea) are subjected to long periods of sunshine. Therefore, they are grown in somewhat dry and warm conditions. Due to this, they exhibit a burgundy brown liquor and a malt, heavy note with a black leaf appearance. Tea leaves grown in upcountry regions of Sri Lanka (i.e., high grown tea) are primarily grown in wet and humid conditions, which give them an extraordinarily light feel with a greenish grassy tone.

These are why studies should be conducted separately for the low country and upcountry regions because findings from one region cannot be applied to the other.

Tea is the most significant export product in Sri Lanka [tea, 2020b]. It covers 52% of all the exports. It is an industry that provides employment to around 1 million Sri Lankans, while tea plantations cover around 4% of real estate in the country. Because of these factors, it would not be an understatement to say that the tea industry is a crucial part of Sri Lanka's economy and millions of Sri Lankans depend upon its growth as well.

The quality of tea mainly depends on the quality of the tea leaves plucked from tea plantations. There are different manufacturing processes that produce different qualities of tea. However, unlike most industries, these manufacturing processes are designed around the quality of the tea leaves, i.e., good quality tea cannot be produced with bad quality tea. Tea factories in Sri Lanka produce around 340 million kilograms of tea per annum ([tea, 2020a]). Every tea manufacturing company in Sri Lanka accepts tea leaves provided by third party suppliers to produce these quantities on a yearly basis. These third-party suppliers bring thousands of kilograms of tea leaves to tea factories all around the country. Nevertheless, the process of checking the quality of tea leaves that they bring, is done manually. It is done by experts who are experienced in identifying the quality of tea leaves by observation. The process of manual quality checking is as follows.

1.) Take tea leaf samples from different bags that carry tea leaves.
2.) Categorize the collected tea leaf samples into three classes: Best, Below-best, and Refused-tea leaves.
3.) Count the number of tea leaves in each class and calculate the good tea leaf percentage

This tea leaf quality identification process is highly prone to human errors. The lengthy and crucial task of identifying each leaf one by one is really hard and time consuming. The mistakes made by humans knowingly or unknowingly cannot be identified or cross checked as this would not be very efficient and beneficial economically.

As discussed above tea plantation is a major part of Sri Lanka's economy, it is very important to improve this identification method and make it error free, accurate, efficient, and fast. This paper proposes a deep learning model based on a Convolutional neural network to take the leaf images as input and process it to classify it either as Best, Below - best or Refused - tea category. The model proposed shows 93% accuracy on the test dataset and outperforms all the other models already present.

II. PROBLEM STATEMENT AND RESEARCH GAP

Out of the three categories, Best and Below-best classes are also known as good and average tea leaves, whereas refused-tea class represent bad tea leaves that are rejected. In the first phase of the process, Best and Below-best tea leaves are counted as one in order to calculate the good leaf percentage. Tea factories that manufacture high-quality tea will only accept the tea leaf supply if the good tea leaf percentage is higher than 65%. For standard quality tea manufacturers, 50% is the benchmark. According to the Tea Research Institute of Sri Lanka, the whole supply should be rejected if these standards are not met. However, in reality, rejection of a supply rarely happens because people who measure this good tea leaf percentage tend to accept the supply even when the percentage is lower than the recommended rate.

In order to solve this problem, this research is focused on building a tea classification model that will identify and classify tea leaves into the three categories. This model will be applicable for the classification of tea leaves grown in low country areas of Sri Lanka. According to the literature, this research marks its novelty by using deep CNNs to build a classification model for tea leaves grown in Sri Lanka. Also, this research is conducted in the hope that this will be the first step in automating the process of calculating the good leaf percentage of tea leaf shipments.

III. RELATED WORK

The problem of leaf quality/disease identification has been bothering the human kind for a long time now. Much research work has been done on this issue, trying to connect the technology with agriculture and help the people make the right decision faster and more accurately. Some of the works related to our research topic have been discussed in this section.

A study done by [10] in Columbia has found the necessity to classify tea leaves in order to speed up the tea production process. They have conducted a study to classify the tea leaves into three classes known as green shoots, green stems and non-green leaves. For this purpose, they have photographed 630 tea leaves where each leaf is placed on a white background. Out of 630 images, 210 were chosen as the final dataset after filtering out images due to various factors such as inadequate lighting conditions and separability. An Artificial Neural Network and K-means clustering algorithm was used for the classification. Under different configurations, the Neural Network achieved a success rate between 90% and 95%, whereas the K-means algorithm was able to achieve a success rate of 89.34%.

[3] have conducted their research based on identifying tea leaves. They have used a collection of 60 tea leaf images in order to build the model. For the pre-processing part, they have used various techniques to extract more reliable features when identifying tea leaves. These techniques include Fuzzy denoising using DT-DWT algorithm and boundary enhancement using a Laplacian Filter. For the classification models, they have used a Radial Basis Function Neural Network (RBFNN) and K-nearest neighbor algorithm. RBFNN produced a success rate of 86.2%, whereas the success rate is 78% for K-nearest neighbor. Also, the execution time of RBFNN was less than the K-nearest neighbor algorithm as well.

Tea bud leaf identification is another aspect where [8] focused their study. This research falls into the category of computer vision, where they have used a cascade classifier to identify tea bud leaves. For this research, [8] has used a Histogram of Oriented Gradients (HOG) in order to extract features, and a SVM has been used for the classification. The cascade classifier used one hundred fifty positive images and three hundred negative images. They contain images of tea leaves that are of four different length types. The overall accuracy of the model was 55%.

Using deep learning methods to image classification problems have gained popularity over the years mainly due to the fact that they do not require manual feature extraction ([4]). [7] presents the idea of using an improved [?] to identify diseases in tea leaves. The diseases that they wish to identify are tea leaf blight, tea bud blight and tea red scab. A set of 36 images have been photographed, which contains images from all three disease categories. The size of the dataset has been increased using data augmentation techniques such as rotation, up-down swapping, and left-right swapping. In the improved CNN, convolutional kernels in the convolutional layers have been constructed with different sizes. Standard convolution is replaced with depth wise separable convolutions, which are used to improve the calculation speed and reduce the number of parameters of the model. The proposed model performed far better, with an accuracy rate of 92.5% than traditional models such as SVM (accuracy rate - 85%), KNN (accuracy rate – 72.5%) and BP Neural Network (accuracy rate - 73%).

[4] 's research was focused on eliminating the time waste which was taken to identify tea leaf diseases using microscopic identification methods and molecular biological and spectroscopic techniques. They have used a CNN to build a classification model and compared it against models built using a SVM and Multi-Layer Perceptron algorithms (MLP). Feature extraction for the SVM and MLP algorithms were done using a DSIFT (Dense Scale Invariant Feature Transform) based BOVW (Bag of Visual Words) model. A total of 3810 images were taken. They were scaled down from 4000x3000 pixels to 256x256 pixels. As shown in figure 1, different data augmentation techniques were used, such as vertical flipping, 180-degree rotation, 90-degree right rotation, 90-degree left rotation and random cropping in order to increase the size of the dataset. The final outcome was a total of 7905 images. 90% of the data was split into an 80/20 ratio of training/test data, and the remaining 10% was used to validate the model. The model was built to classify seven different diseases. The CNN architecture used by [Chen et al., 2019] is an improved version of the AlexNet model called LeafNet. Input images were rescaled to 227x227 pixels, and all three color channels were processed directly by the network. Accuracy and Mean Class Accuracy indices were used to evaluate the performance. Accuracy and MCA of the LeafNet model were 90.23% and 90%. The same measurements for the SVM were 60.95% and 60.62%. Moreover, for the MLP algorithm, it was 70.94% and 70.77%. The LeafNet model has outperformed SVM and MLP models.

[15] have also conducted their research based on the tea leaves taken from the plants in India. The use of CNN model with 4 hidden layers and 2 fully connected layers have given promising results to [15]. The input images are classified into 8 classes such as normal leaf, Algal leaf spot, Gray blight, White spot, Brown blight, Red scab, Bud blight and Gray blight. The deep learning mode proposed in [15] achieved an accuracy of 94.45% on identifying the tea plant disease.

The semi-supervised learning and image processing technique by [16] identifies the early spring green tea leaves. The method proposed by [16] identifies the tender leaves using deep learning. They trained their model both in 2- and 3-dimensional space and all the three R, G and B components of tender leaves and their background. The results showed 93.62% accuracy with the misjudgement of 18.86%. [16] claims that the model proposed has strong versatility and adaptability.

In their research [12] also presents the idea of tea leaf disease classification. Their image acquisition method is a bit different from the other research mentioned earlier. [12] have used a conveyor belt apparatus in order to place the diseased tea leaves and photograph them one by one after every 2-second interval. Using this approach, they were able to capture 15063 images of 6 kinds of common leaf diseases. After using data enhancement methods such as rotating, flipping, and noise addition, a total of 25186 images were obtained. During the pre-processing part, they have used Gaussian image smoothing to deal with interfering elements in images. Then they have converted all the images into grayscale images in order to find the image outline. Then the sobel edge extraction method was used to extract the edges in the

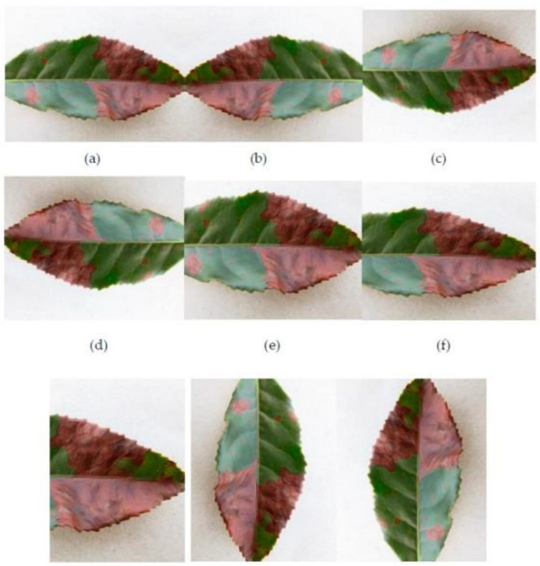

Fig. 1: Augmented Images

vertical direction. This pre-processing part and the sheer size of the dataset gives this research a significant advantage over the previously mentioned research. To avoid the overfitting problem, Max pooling algorithms has been used on the dataset. Rectified Linear Unit (ReLU) activation function has been used for both the fully connected layer and the convolutional layer of the CNN. The generalization ability of the model has been improved by introducing the Local Normalization Layer. The accuracy score of the CNN was 93.75%. The accuracy of the CNN model was compared against SVM and BP algorithms. SVM and BP algorithm produced accuracy scores of 89.36% and 87.69% respectively. [12] have further stated that adjustment of the learning rate throughout the process has consumed much time. Therefore, they have pointed out the necessity of developing a method to find the optimal learning rate.

IV. METHODOLOGY

This research focuses on investigating and classifying tea leaves found in low country areas of Sri Lanka. The classification will divide the tea leaves into three categories which will help in identifying and separating good leaves from bad leaves. In order to achieve this purpose, a constructive research approach ([Lukka, 2003]) was followed throughout this research. Design science, which is a part of constructive research, was also applied to this research design as well.

A. Data collection

As the first part of the research design, it was essential to collect a significant number of tea leaf images belonging to the classes of Best, Below-best and Refused-tea. Collecting similar amounts of images in each category was vital to building a balanced model. The images were captured using a Nikon D3500 DSLR camera. Camera configurations include an aperture level of 36, ISO level of 800 and a shutter speed of 1/50. These configurations were necessary since a LED focus light was used to illuminate the features of tea leaves. Initial experiments were done using a mobile device, but the attempt failed since the mobile device camera started capturing the flicker of the LED focus light flicker, which led to the capture of unclear images. Therefore, for the images to be clear and precise, these images were taken inside a photo booth setup. This setup also helped avoid unnecessary feature absorptions during the model's training. These unnecessary features include shadows and random variations of brightness. Each leaf was be placed on a white background inside the photo booth, and optimal lighting conditions were be applied using a LED focus light as shown in figure 2.

Fig. 2: Photobooth Setup

The tea leaves that were captured using the above setup were collected from two different tea estates in the low country region of Sri Lanka. The tea estates are namely Homadola Estate Udugama (Figure 3) and Talangaha Estate Nakiyadeniya. Both of these tea estates belong to Watawala Plantations Ltd.

Fig. 3: Homadola Estate Tea Repository

There have been cases reported in previous work where machine learning models have learned biases of the people who have collected the data ([Fuchs, 2018]). To mitigate this factor, the opinions of two experts with different levels of experience in the tea industry were used during the data labelling phase. A total of 336 images were captured during data collection. The table I shows the number of images collected in each category.

Image class	Frequency
Best	110
Below best	108
Refused tea	118

TABLE I: Image frequency according to class

B. Preprocessing

Even though the photo booth setup mitigated most of the background noise, there were still images with some issues, such as images with shadows, blurred images and images with background noise. As the initial step in the image preprocessing stage, the images that cannot be used to build the model were filtered out. Then the leaves from the three categories were separated and labelled according to their class with the help of the experts.

The images were then stored in a Python list data structure and were downsampled to fit each model's input requirement. The original image resolution was 6000x4000x3 pixels when including the RGB values. One copy of the dataset was downsampled to 224x224x3 pixels to accommodate VGG16 architecture and the proposed novel architectures. Another copy of the dataset was downsampled to 227x227x3 to accommodate the AlexNet architecture. Data augmentation techniques were needed to enhance the dataset's size and increase the model's generalizability. The data augmentation techniques used include vertical flipping, horizontal flipping, 180-degree rotation, 90-degree right rotation, random zoom and random cropping. Both copies of the dataset were then converted to two NumPy arrays and were normalized by dividing each RGB value by 255(i.e., the highest RGB value). Finally, the dataset was split as follows, 80% training data, 10% test data and 10% validation data.

C. Model Building

The literature review revealed that deep CNNs were the best approach to building a classification model because of the advantages they provide. Three novel CNN architectures were proposed and implemented to select the best approach for the tea classification problem.

1.) *Proposed CNN model 1 (M1):* The M1 architecture contains three convolutional layers where the number of filters has doubled in each layer. Thirty-two filters in the first convolutional layer, 64 filters in the second layer and 128 filters in the third layer. Max pooling was used as the pooling method and a max pooling layer follows each convolutional layer to reduce dimensionality. Two fully connected dense layers come after the pooling layers, where the second dense layer acts as the classifier. The CNN architecture M1 is shown in figure 4.

The original images were down sampled from 6000x4000x3 to 224x224x3 to accommodate the model's input. Non-linearity is introduced to the model using the ReLu activation function in the convolutional layers. Each convolution is performed with a 3x3 kernel. Max pooling operation uses a 2x2 pool size with a stride of 2. The final layer consists of 3 output neurons where each neuron represents a class of tea leaves. A SoftMax function was used in the output layers, which will transform the output of each neuron to a value between 0 and 1.

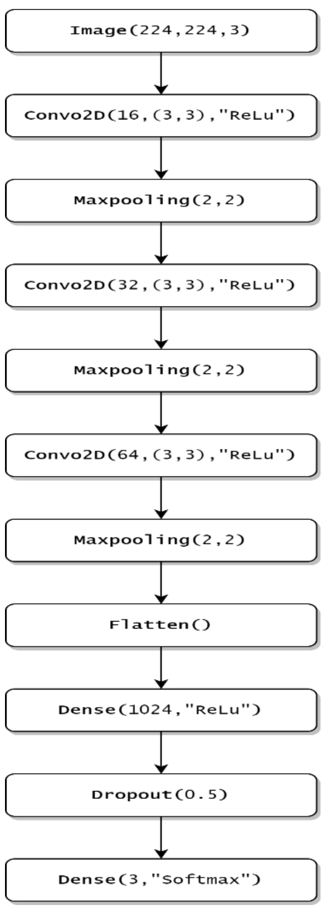

Fig. 4: Proposed Architecture M1 Flowchart

2.) *Proposed CNN model 2 (M2):* Increasing the number of layers tends to increase CNN models' accuracy levels on some occasions. This layering approach is more appropriate when the dataset is larger since more consistent features can be extracted with more layers and fewer strides. Nevertheless, these factors can change according to the classification problem at hand. The proposed CNN architecture M2 is similar to the previous M1 architecture, but with one more convolutional layer. This newly added convolutional layer contains fewer filters than the other convolutional layers. In both M1 and M2 architectures, the number of filters increases from one layer to another. The idea is to gradually identify more features as the data passes through the layers.

Non-linearity is introduced to the model using the ReLu activation function in the convolutional layers. Each convolution is performed with a 3x3 kernel. Max pooling operation uses a 2x2 pool size with a stride of 1. A dropout layer has been

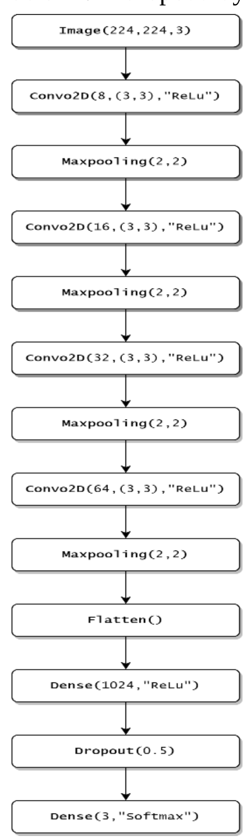

Fig. 5: Proposed Architecture M2 Flowchart

added in order to deal with overfitting issues. The proposed M2 model is illustrated in figure 5.

3.) *Proposed CNN model 3 (M3):* Max pooling layer performs dimensionality reduction to the data by using the feature map generated from the convolutional layer. It calculates the maximum value for patches of the feature map and creates a pooled feature map. It is better to have fewer max pooling layers in some cases since it reduces the complexity of the model. The proposed M3 architecture is similar to the M1 architecture, but it contains one less max pooling layer than M1. Each convolution is performed with a 3x3 kernel. Max pooling operation uses a 2x2 pool size with a stride of 2. Figure 6 depicts the proposed M3 architecture.

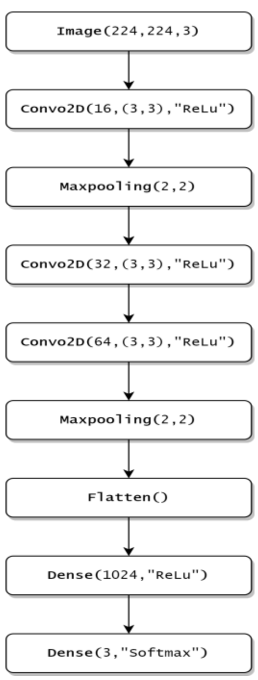

Fig. 6: Proposed Architecture M3 Flowchart

D. Existing CNN Models

To show the effectiveness of the proposed models, further experiments were conducted using two advanced CNN architectures, VGG16 and AlexNet.

1.) *AlexNet Architecture:* The AlexNet consist of eight layers with learnable parameters. Five of which are convolutional layers combined with max pooling followed by three fully connected layers. ReLu activation function was used in each of these three fully connected layers except the output layer. This was the first time the ReLu activation function was used in a CNN, and it accelerated the training process by almost six times. In order to stop the model from overfitting, they have used multiple dropout layers as well. Visualization of the AlexNet architecture is depicted in the figure 7.

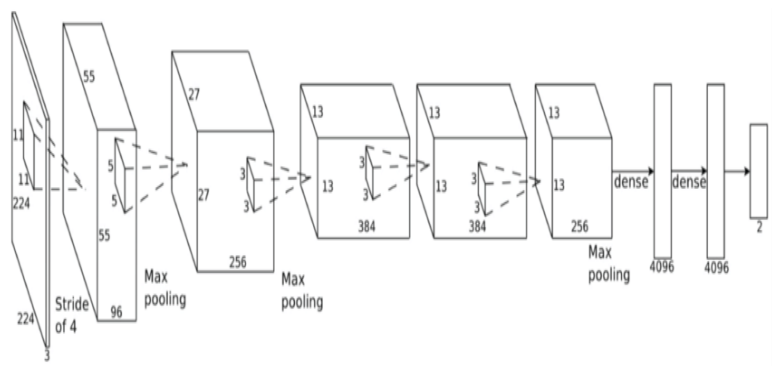

Fig. 7: AlexNet Architecture

2.) *VGG16 Architecture:* The VGG16 [Simonyan and Zisserman, 2014] architecture is a considerably extensive network that consists of approximately 138 million parameters. It was first introduced in 2014, where it won the ImageNet Challenge 2014 and is considered to be one of the best vision model architectures up to date. The architecture consists of three dense layers, five max pooling layers and thirteen convolutional layers. The VGG16 model is unique because instead of having a large number of hyperparameters, it has a set of consistent convolution layers of 3x3 filters and max pool layers of 2x2 pool size of stride 2. Also, it uses the same padding. The dense layer consists of 2 fully connected layers and a SoftMax layer for output. The VGG16 architecture is depicted in the figure 8.

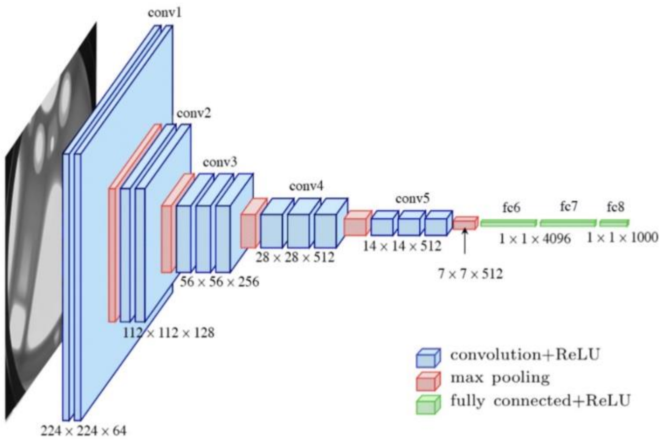

Fig. 8: VGG16 Architecture

3.) *Transfer Learning with Deep Residual Learning Network:* Transfer learning is the concept of using a pre-training model as the starting point of building a model for a new task ([Torrey and Shavlik, 2010]). Transfer learning can yield much higher performance when the dataset size is limited. Transfer learning is so frequent that training a model for image processing tasks from scratch is quite rare. Researchers and data scientists prefer to begin with a pre-trained model that knows how to identify objects and has learned general features such as edges and shapes in photos.

Deep Residual Neural Network (i.e., ResNet) is a highly advanced CNN architecture that won the ILSVRC 2015 image classification competition by showing an error of only 3.57%. Layer saturation is a problem in convolutional neural networks where the performance of the model deteriorates with the increase of layers ([He et al., 2016]). ResNet model was introduced as a solution to this issue. Residual blocks are used in deep residual learning networks to increase model accuracy. The strength of this form of the neural network is the concept of “skip connections” which is at the heart of the residual blocks. There exist different variations of ResNet, and for this research, ResNet-50, which is a 50-layer Convolutional Neural Network architecture, was used as the transfer learning model. ResNet-50 architecture is depicted in figure 9.

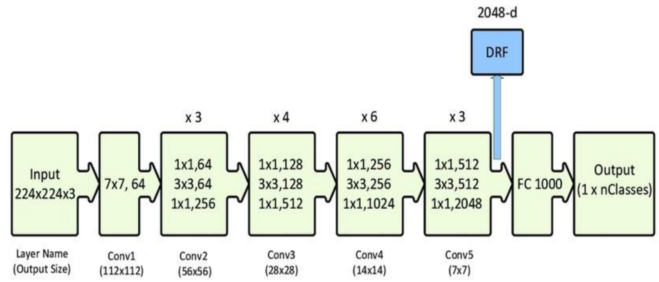

Fig. 9: ResNet-50 Architecture

IV. EVALUATION

The proposed methodology is evaluated by obtaining accuracy, precision and recall metrics using True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) values which are defined below. Since the proposed model is a multi-class classification model, we have to separate each class when calculating these performance measures.

True Positive (TP): The model correctly predicts the positive class.

True Negative (TN): The model correctly predicts the negative class.

False Positive (FP): The model incorrectly predicts the positive class.

False Negative (FN): The model incorrectly predicts the negative class.

The following equations shows how the three accuracy measures Accuracy, Precision and Recall are calculated using the above values.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{}$$

$$\text{F1score} = \frac{TP}{TP+1/2(FP+FN)}$$

The performance of the proposed CNN models M1, M2 and M3 are compared with the existing advanced CNN models AlexNet, VGG16 and ResNet-50 with transfer learning. This comparison aims to check whether there is a performance improvement in the proposed novel models concerning the already existing models.

A. Proposed CNN Architecture M1

The proposed CNN architecture M1 starts with an augmentation layer is followed by three convolutional layers with one max pooling layer after each convolutional layer. A dropout layer has been added between the two dense layers to deal with overfitting issues. The architecture of model M1 is visualized in figure 10. Convolutional layers are represented by grey, while teal represents the max pooling layers. Red, green and pink colours represent the flatten layer, dense layers and dropout layer, respectively. Accuracy-loss measure, which is common in training and evaluating CNN models, was used in the proposed M1 architecture. The model was able to achieve a training accuracy of 85.24% and a test accuracy of 92.82%. The model was able to perform considerably well even in the prediction phase by accurately classifying 375 out of 404 samples, as shown in the diagonal of the confusion matrix in figure 11. The model identified all the images from the Best tea leaf category but had difficulty differentiating between Below best and Refused-tea since there were 29 misclassifications out of 404 images. The model correctly predicted 118 images from Below-best class out of 130 and 129 images from Refused-tea class out of 142. The confusion matrix for the proposed novel. CNN model M1 is shown in figure 11

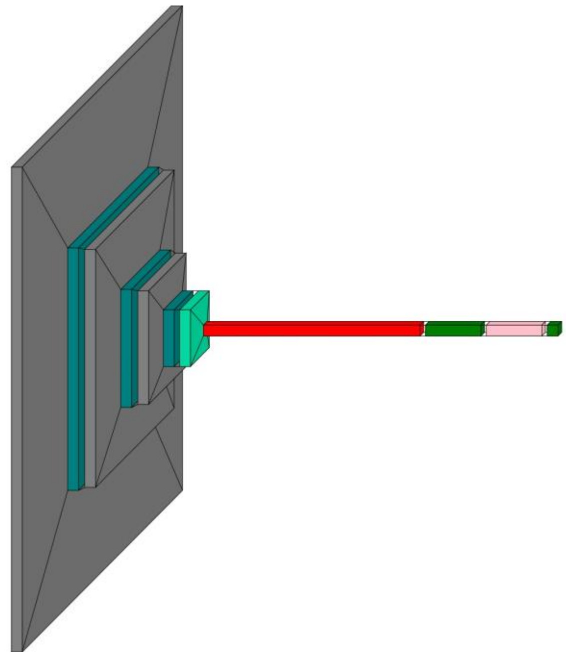

Fig. 10: Proposed CNN Architecture M

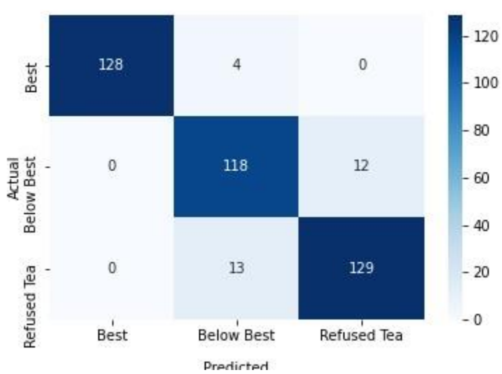

Fig. 11: Confusion matrix for the proposed CNN M1

Epoch-Accuracy plot shown in figure 12 shows clearly that training accuracy and the validation accuracy increase gradually with respect to the epoch count. The progress of accuracy of the model with respect to the number of epochs is shown in figure 12.

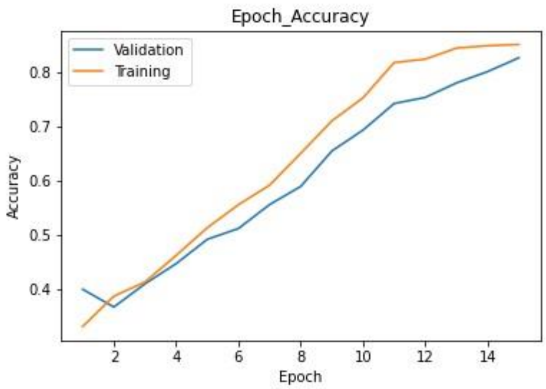

Fig. 12: Accuracy against number of epochs for proposed CNN M1

B. Proposed CNN architecture M2

The model M2 reached a training accuracy of 91.88%, and the testing accuracy was recorded at 93.81%. Both of these values were an improvement to those of the values achieved in the proposed M1 model. The confusion matrix for the model M2 is depicted in figure 13. The diagonal of the confusion matrix shows that the model was able classify 336 out of 404 images in the test dataset. It further points out that the model has made misclassifications in all three classes, albeit the

number of misclassifications was meager in the Best tea leaves class. The model has made 2, 12 and 11 wrong predictions in Best, Below-best and Refused-tea classes, respectively.

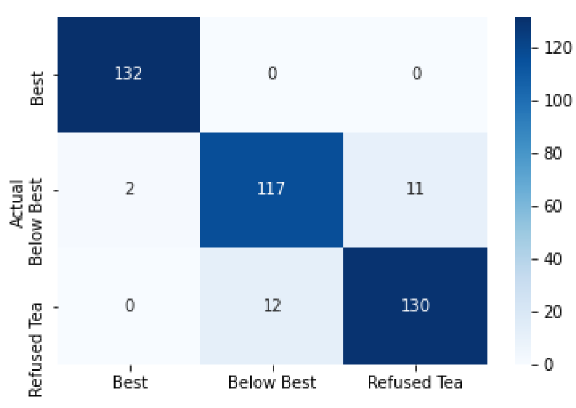

Fig. 13: Confusion matrix for the proposed CNN M2

The evolution of the training accuracy in the proposed architecture M2 is depicted in figure 14.

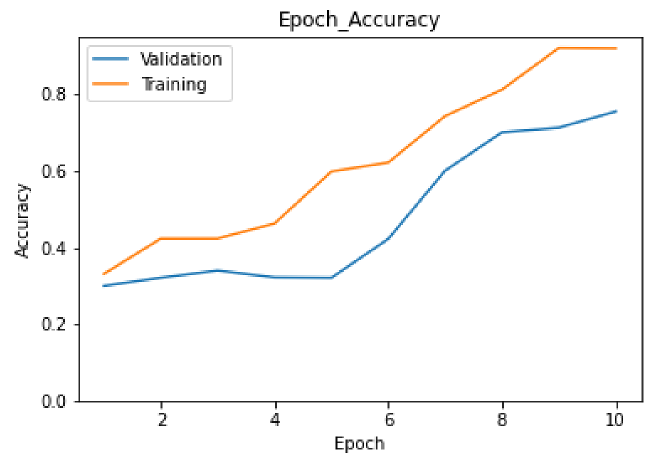

Fig. 14: Accuracy against number of epochs for proposed CNN M2

C. Proposed CNN architecture M3

The proposed CNN architecture M3 is structured with three convolutional layers and two max-pool layers. A data augmentation layer comes before the convolutional layers, and the model ends with one fully connected layer and a SoftMax layer.

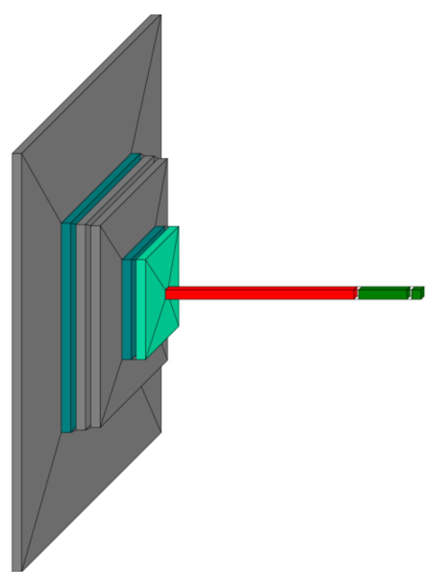

Fig. 15: Proposed CNN Architecture M3

The training accuracy of the proposed model was 75.43%, and the test accuracy was 89.60%. These values are relatively low when compared to the models M1 and M2. The confusion matrix for the proposed model is shown in figure 16. The proposed model M3 made 362 correct predictions out of 404 images in the test dataset. The model has made 4, 24 and 14 misclassifications in Best, Below-best, and Refused-tea classes, respectively.

The architectural difference is also evident from the number of parameters that exist within the models. The three novel CNN models, M1, M2 and M3, were trained using the same tea leaf image dataset that was collected for the sole purpose of this research and the experimental results obtained are used in the performance comparison.

Model	Layers	Parameters	Train-Acc	Test-Acc
Proposed CNN Architecture M1	10	47,803,427	85.24%	92.82%
Proposed CNN Architecture M2	12	11,104,211	91.88%	93.81%
Proposed CNN Architecture M3	8	198,274,083	75.43%	89.60%

TABLE II: Comparisons of Proposed CNN architectures

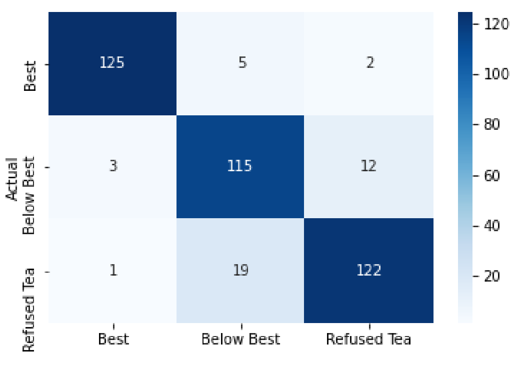

Fig. 16: Confusion matrix for the proposed CNN M3

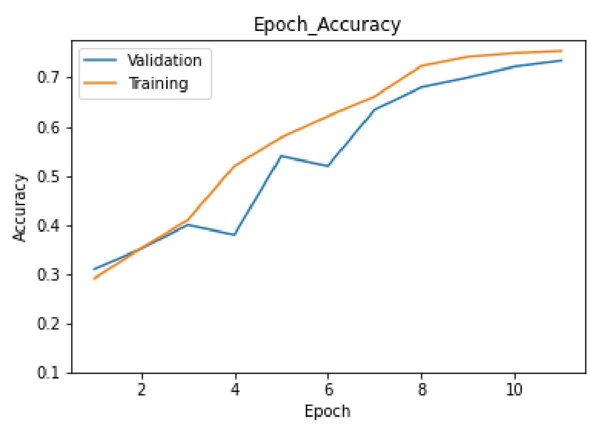

Fig. 17: Accuracy against number of epochs for proposed CNN M3

Figure 17 shows the progression of accuracy with respect to the number of epochs.

D. Comparison of proposed CNN architectures

The three novel CNN architectures were introduced to classify low country tea leaves into three categories known as Best, Below best and Refused. Model structures variate according to the number of layers and how those layers were used within each model.

Tea Leaf Class	Performance measure	Proposed CNN M1	Proposed CNN M2	Proposed CNN M3.
Best	Accuracy	0.99	0.995	0.972
	Precision	0.969	1.00	0.946
	Recall	1.0	0.985	0.968
	F1-score	0.984	0.992	0.957
Below best	Accuracy	0.928	0.938	0.903
	Precision	0.907	0.900	0.884
	Recall	0.874	0.906	0.827
	F1-score	0.890	0.903	0.855
Refused tea	Accuracy	0.938	0.943	0.915
	Precision	0.908	0.915	0.859
	Recall	0.914	0.921	0.897
	F1-score	0.911	0.918	0.877

TABLE III: Performance of proposed Novel CNN architectures

Tea Leaf Class	Performance measure	M1	M2	ResNet-Transfer Learning	AlexNet	VGG16
Best	Precision	0.969	1.00	0.984	0.00	0.44
	Recall	1.0	0.985	0.968	0.00	1.00
	F1-score	0.984	0.992	0.988	0.00	0.61
Below best	Precision	0.907	0.900	0.969	0.00	0.00
	Recall	0.874	0.906	0.940	0.00	0.00
	F1-score	0.890	0.903	0.954	0.00	0.00
Refused tea	Precision	0.908	0.915	0.943	0.35	0.00
	Recall	0.914	0.921	0.964	1.00	0.00
	F1-score	0.911	0.918	0.953	0.52	0.00

TABLE IV: Performance Evaluation of All Architectures training accuracy and testing accuracy concerning the three proposed models. Among the three models, proposed architecture M2 shows the best training accuracy by reaching 91.88%. Furthermore, the proposed CNN architectures M1 and M2 shows the best testing accuracy of 93.81% and 92.82%, respectively. The proposed architecture M3 shows the lowest accuracy levels of all the models. Accuracy, precision and F1-score were calculated as performance measures for each model, and they are summarized in table According to the results shown in the table III, it is evident that adding an extra convolutional layer and a max pooling layer could have a significant impact on the model performance. The proposed models M1 and M2 has outperformed the model M3, their performance values will be compared with the advanced CNN architectures in the following sub-sections of this chapter.

V. ALEXNET MODEL PERFORMANCE

The AlexNet model reached a training accuracy of 83.52%and a testing accuracy of 35.29%. These accuracy scores indicated that the model was overfitting the data. Due to this reason, the model could not correctly classify most tea leaf images into their respective classes.

The figure 18 shows the classification report of the AlexNet model.

```
Classification Report:
              precision    recall

           0       0.00      0.00
           1       0.00      0.00
           2       0.35      1.00

    accuracy
   macro avg       0.12      0.33
weighted avg       0.12      0.35
```

Fig. 18: Classification Report of the Alexnet Model

Figure 19 depicts the accuracy of the model with respect to the number of epochs

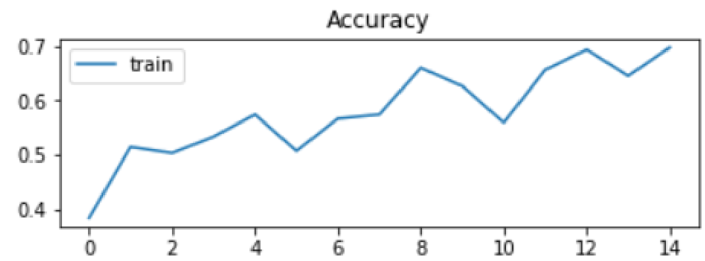

Fig. 19: Accuracy against number of epochs for AlexNet

VI. VGG 16 MODEL PERFORMANCE

The model reached a training accuracy of 29.8% and test accuracy of 44.12%. The classification report shown in figure 21 combined with the accuracy values suggests that the model was not able to classify the images into the three known classes.

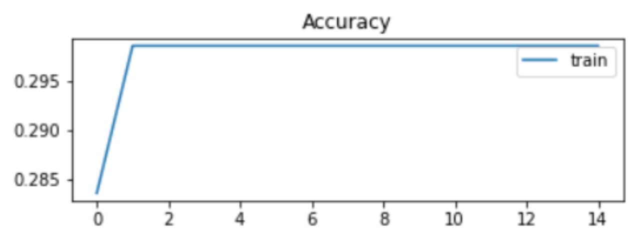

Fig. 20: Accuracy against number of epochs for VGG16

Accuracy variation with regards to the number of epochs is shown in figure 20.

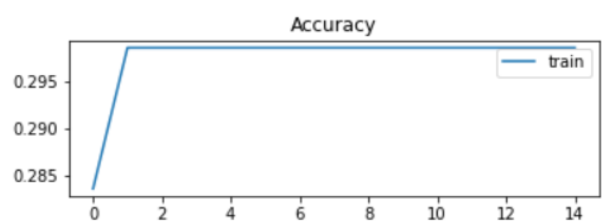

Fig. 21: Classification Report of the VGG16 Model

VII. TRANSFER LEARNING WITH RESNET-50

To compare the results with the proposed CNN architectures, the transfer learning approach was used on the model ResNet-50. The model shows a training accuracy of 95.38% and a testing accuracy of 96.53%.

Figure 22 shows the confusion matrix for the ResNet-50 transfer learning model.

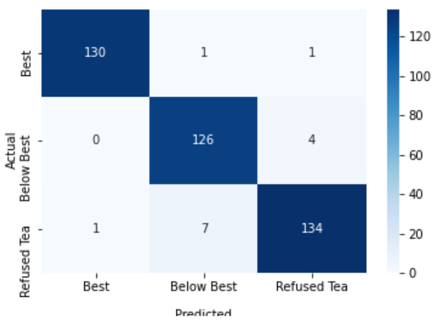

Fig. 22: Confusion matrix for the transfer learning model ResNet-50

The figure 23 shows a set of 12 predictions made by the ResNet-50 transfer learning model for the Below-best class

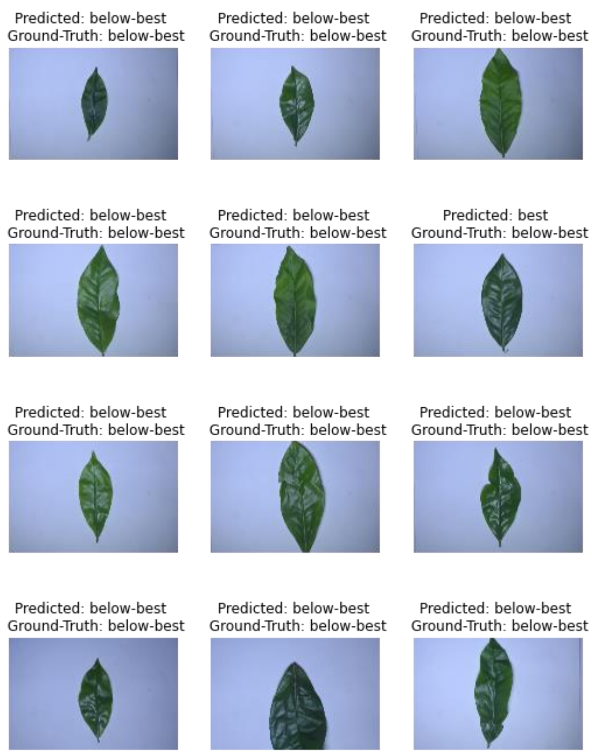

Fig. 23: Predictions of the ResNet-50 Transfer learning Classifier

VIII. FINAL EVALUATION

Precision, recall and F1-score values of the proposed CNN models M1, M2 and existing advanced CNN models AlexNet and VGG16 are shown in table IV. Proposed model M3 was omitted since the performance of the M1 and M2 models were higher than the M3 model. These values are represented with respect to the tea leaf classes Best, Below-best and Refused-tea. The proposed CNN models M2 and ResNet-50 with transfer learning show the best competitive performance among all the CNN models.

CONCLUSIONS

A. *Conclusion on research questions*

The main avenue explored through the research questions was the kind of techniques that could be used to classify and grade tea leaves in the low country region of Sri Lanka. From the analyzed literature, it became clear that deep CNNs offer significant advantages over the other techniques that could be used to pursue this goal. Automatic detection of features and reduced computational cost could be named some of these advantages over the other techniques. By following the research questions that were established, collecting, and building a dataset that consists of tea leaf images, designing and developing simple CNN architectures with significant performance, using implementations of advanced CNN architectures such as AlexNet, VGG16 and ResNet-50 with transfer learning for performance comparison could be considered as the contributions made by this research. The deep CNN models were evaluated through a data-driven approach. The pre-trained ResNet-50 model trained using the ImageNet dataset was able to reach an accuracy of 96.53%, which is a good performance. However, due to the model size and training/prediction time are both slightly increased. Experimental results show that the proposed architecture M2 is the most effective model that could be used for the tea leaf classification. The architecture M2 shows an accuracy level of 93.81% with a simple layering structure. Due to many parameters and the layering structure of the advanced CNN models AlexNet and VGG16 did not produce high levels of accuracy compared to the proposed

CNN models M1, M2 and M3. The CNN models AlexNet and VGG16 displayed more significant predictions times when compared to the proposed models as well. Using advanced concepts in Computer Science, this research has taken the first step toward automating the process of identifying and grading tea leaves before they go into production. The research has developed a novel framework for classifying tea leaf classes known as Best, Below-best and Refused-tea using Deep Learning and Computer Vision fields. This work will undoubtedly bridge the gap between technology and the tea industry of Sri Lanka and pave the way for more agricultural solutions using Computer Science.

B. Limitations and Future Work

The evaluation results show that all the models have difficulty separating Below-best tea leaves from Refused-tea due to the small margin of difference between the two types of leaves. The images were taken under proper lighting conditions and without much noise in the background. Therefore, there is a possibility that the model may not perform as well for images with a lot of background noise. For future work, these issues could be mitigated by collecting more tea leaf images under different lighting conditions with background noise and using them to improve the model. Furthermore, it is possible to develop different novel CNN architectures that will fit different image types (i.e., images taken under different conditions) as well. Since CNNs are computationally less expensive than other models, this approach will not produce any additional overhead in practical scenarios.

ACKNOWLEDGMENT

The authors would like to thank the staff members of Watawala Plantations Ltd., who dedicated their time to help during the data collection phase

REFERENCES

[1] "*Sri Lanka export development board"* , [tea, 2020a] (2020a).

[2] *"Trade Economics Website",* [tea, 2020b] (2020b).

[3] Arunpriya and Thanamani, Arunpriya, C. and Thanamani, A. S. *"An effective tea leaf recognition algorithm for plant classification using radial basis function machine".* Int. J. Modern Eng. Res. (IJMER), 4(3):35–44, 2014.

[4] Chen, J., Liu, Q., and Gao, L., "*Visual tea leaf disease recognition using a convolutional neural network model", Symmetry*, 11(3):343, 2019.

[4] Dyrmann, M., Karstoft, H., and Midtiby, H. S., *"Plant species classification using deep convolutional neural network",* Biosystems engineering, 151:72–80, 2016.

[5] Fuchs, D. J. , *"The dangers of human-like bias in machine-learning algorithms",* Missouri S&T's Peer to Peer, 2(1):1, 2018.

[6] He, K., Zhang, X., Ren, S., and Sun, J. , *"Deep residual learning for image recognition",* In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.

[7] Hu, G., Yang, X., Zhang, Y., and Wan, M. , *"Identification of tea leaf diseases by using an improved deep convolutional neural network"*. Sustainable Computing: Informatics and Systems, 24:100353, 2019.

[8] Karunasena, G. and Priyankara, H. , *"Tea bud leaf identification by using machine learning and image processing techniques"* Int. J. Sci. Eng. Res., 11(8):624–628, 2020.

[9] Lukka, K., *"The constructive research approach"*, Case study research in logistics. Publications of the Turku School of Economics and Business Administration, Series B, 1(2003):83–101, 2003.

[10] Silva, J., Palma, H. H., Nu´nez, W. N., Ruiz-Lazaro, ˜ A., and Varela, N. , "*Neural networks for tea leaf classification*". 1432(1):012075, 2020.

[11] Simonyan, K. and Zisserman, A. (2014). "*Very deep convolutional networks for large-scale image recognition*", arXiv preprint arXiv:1409.1556, 2014.

[12] Sun, X., Mu, S., Xu, Y., Cao, Z., and Su, T., "*Image recognition of tea leaf diseases based on convolutional neural network",* arXiv preprint arXiv:1901.02694, 2019.

[13] Tea (2020). *"Tea Exporters Association Sri Lanka".*

[14] Torrey, L. and Shavlik, J., "*Transfer learning. In Handbook of research on machine learning applications and trends: algorithms, methods, and techniques",* pages 242–264. IGI global, 2010.

[15] R.S. Latha, G.R. Sreekanth, R.C. Suganthe, R. Rajadevi, S. Karthikeyan, S. Kanivel, B. Inbaraj, *"Automatic Detection of Tea Leaf Diseases using Deep Convolution Neural Network",* Automatic Detection of Tea Leaf Diseases using Deep Convolution Neural Network

[16] Jie Yang, Yong Chen, *"Tender Leaf Identification for Early-Spring Green Tea Based on Semi-Supervised Learning and Image Processing",* Agronomy 2022, 12, 1958. https://doi.org/10.3390/agronomy12081958 agronomy, 2022