Efficient Tea Leaf Classification : Bridging Supervised and Semi-Supervised Learning

Md Tahsin[1†], Maksura Binte Rabbani Nuha[2†], Lamia Haider[1†], Al Hossain[1†], Dr.Raihan Ul Islam[1*†], Dr. Mohammad Rifat Ahmmad Rashid[1†], Dr. Ahmed Wasif Reza[1†], Dr. Shamim H Ripon[1†], Mohammad Shahadat Hossain[3†]

[1]Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh.
[2]Department of Computer Science and Engineering, International Islamic University Chittagong, Chittagong, Bangladesh.
[3]Department of Computer Science and Engineering, University of Chittagong, Chittagong, Bangladesh.

*Corresponding author(s). E-mail(s): raihan.islam@ewubd.edu;
Contributing authors: tahsin30899@gmail.com; maksuranuha07@gmail.com; lamiahaider484@gmail.com; alhossaiin@gmail.com; rifat.rashid@ewubd.edu; wasif@ewubd.edu; dshr@ewubd.edu; hossain_ms@cu.ac.bd;
†These authors contributed equally to this work.

Abstract

Tea, a globally cherished beverage, comes up with a broad consumer base. Tea buds contain a rich diverse variety of flavours and aromas. Therefore, careful selection of tea leaves is essential in the process of quality control and grading of tea within the tea industry. The challenge of distinguishing between healthy and diseased tea leaves arises, resulting in a significant annual loss in tea yields. To address this issue, researchers conducted numerous studies, yet many lacked sufficient accuracy. In our study, we assessed the efficacy of both supervised and semi-supervised approaches. Remarkably, we used only 25% labeled data, demonstrating a new potential of semi-supervised models in tea leaf classification. For this instance, a custom dataset from Udalia Tea Garden, Bangladesh was used. In supervised learning, DenseNet-121 achieved the highest accuracy of 97% and least accuracy was 92% by EfficientNetB0. And Bridged the best Supervised model with fixmatch algorithm in the Semi-supervised method that obtained an identical accuracy of 96%. A notable observation in our study, similar performance was achieved. These findings indicate that although supervised learning models offered a higher accuracy, the performance of semi-supervised models demonstrates a significant potential, specifically with small size of labeled data. Additionally, for making quality tea, only the upper surface or young leaves are adopted. In this aspect, our model serves a noble purpose by ensuring Tea quality. Moreover, this model stimulates a notable improvement in quality assurance mechanisms in tea production.

Keywords: Tea leaf, Fixmatch, DenseNet-121, MobileNet-V2, Inception-V3, Classification, Semi-supervised, Supervised learning.

1 Introduction

Tea is the most popular drink globally offering a great variety of refreshments and versatile options which can satisfy our growing demand for flavorful, healthful and natural beverage options. As a result, global tea consumption has increased by 3.5 percent annually over the last decade and an estimate of 6.4 million tons by the Food and Agricultural Organization's(FAO) in 2021 [1]. China is the top producer of tea (2.46 million tonnes), followed by India(including Assam and Darjeeling regions) producer of tea (1.33 million tonnes). China and India are the only two Asian countries that produce over one millions tonnes of tea per year [2]. According to (GIA, 2011), the global market for hot beverages which includes coffee and tea is forecasted reaching US$69.77 billion in value and 10.57 million tons in volume terms by the year 2015 [3].

As a major tea-producing country, Bangladesh ranks to 12th largest tea producer in the world[4]. According to the state-run Bangladesh Tea Board(BTB), Bangladesh's tea production surged to a record high of 102.92 million kilograms (kg) in the just concluded year of 2023[5].By (ITC, 2010), globally tea is cultivated in 36, 91,938 ha with an annual production of 40,66,596 thousand Kg and also it is a major cash crop along with a significant export item of 0.81% Bangladesh accounting GDP (BTB,2002)[6].

The high demand for tea stimulated extensive studies into both tea production and leaf classification, impacting trading and tea pricing. Traditional tea leave classification is time-consuming and error-prone, whereas deep learning offers a faster, more consistent and cheaper solution although not widely used in agriculture[11]. In this research, we explored deep learning models to identify the most efficient methods for tea classification. Additionally, we introduce a semi-supervised approach that leverages minimal labeled data that obtains similar accuracy comparable to supervised model. The primary contribution of this study is demonstrating high classification accuracy with our collected dataset. We employed both supervised and semi-supervised models to showcase notable performance improvements. For a detailed perspective, we outlined the utilized dataset as:

For a detailed perspective, we outlined the utilized dataset as:

- Tea making leaves: In this dataset, we compiled images of healthy leaves chosen on a random basis, from Uaila Tea Estate(UTE), Fatikchari, Chittagong.
- Ineligible Tea making leaves: There are 2 types of tea leaves here, which are:
  - Diseased Tea Leaves: Several types such as, black rot, brown and gray blight, brown dots, blister blight and yellow spot tea leaves are used in the datasets.
  - Aged Tea leaves : Using these types of teas in production may not be harmful, but it does not taste fresh or provide the same factors as the healthy tea leaves.

In this paper, research similar to our study is presented in the Literature Review section. Data collection, preprocessing and models are shown in the Methodology section. On this basis accuracy rate, training model and comparison between the models are presented in the Result and Discussion section. At last, we conclude our paper in the Conclusions section.

2 Related Works

This section reviews relevant studies primarily focusing on two areas: classifying tea leaves and detecting leaf diseases. In 2020, Jing Ch.[7] developed several models for tea leaf disease classification. Among these, an enhanced Convolutional Neural Network (CNN) model named LeafNet achieved an accuracy of 90.23% surpassing Support Vector Machine(SVM) with accuracy of 60.91% and Multilayer Perceptron(MLP) at 70.94%. However, factors such as, comparable disease symptoms, seasonal variations and potential co-infection may impact the models accuracy. Similarly, W Wu[16] developed a TeaLeafNet model that validates a remarkable accuracy of 90.42% in identifying tea leaf diseases based on deep transfer learning. However, dataset variability may lead to models' poor performance.

In 13, Priyankara used a Support Vector Machine classifier(SVM) and Histogram oriented Gradient(HOG) for feature extraction. For tea buds with a length of 0 mm to 40mm, the accuracy was 55%. And mid-range lengths between 10mm to 30mm gave 75% of accuracy, as majority training samples fall within this mid-range. In another study, Puja Banerjeea developed Grad-CAM to increase transparency in tea leaf classification using pre-trained VGG16 with 81% and InceptionV3 model with 79.92% accuracy via transfer learning, to provide a visual explanation. Nevertheless, the study identifies its constraints like class unevenness in dataset[10].

During 2015, Bikash Chandra Karmokar[21] integrated a Neural Network Ensemble to develop a pattern recognition model Tea Leaf Disease Recognizer (TLDR), with an accuracy of 91% to minimize disease impact in Bangladesh. However, the image quality might lead to distorted feature extraction and disease recognition. Another group of researchers, in [8], using advanced models such as Faster R-CNN, AX-RetinaNet and VGG16, proposed a method for more accurate detection analysis of tea leaf blight disease, where the integrated AX-RetinaNet model acquired 93.83% accuracy.

In [11], Dejene Teka analyzed deep learning models Convolutional Neural Network (CNN), Inception v3, and EfficientNet B0 and RGB colour to classify tea leaves. In this experiment, the Convolutional Neural Network (CNN) model achieved the highest accuracy of 96.9%. However, data scarcity across locations depicts challenges for larger-scale tea production using this model. Likewise, in 2021, IR Paranavithana[14] developed an interactive software for tea bud(s) classification and proposed Convolutional Neural Network (CNN), Support Vector Machine classifier(SVM) and Inception V3 models for the experiment with a vast dataset. Here, CNN achieved the highest accuracy of 70.15%, which is comparatively lower than other studies.

An appealing study aspect provided by K Wei [12] that experimented with fluorescence images of tea leaves and deep learning for the test. Using VGG16 and ResNet-34 models with the same batch size of 32 and learning rate, reached peak accuracy of 97.5%. RGB images have highest accuracy while worst accuracy by grayscale images compared to other single-channel options. Furthermore, in 2023, an integrated YOLOv5 model was proposed by Jie Cao[9] to categorize picked tea leaves based on their quality and moisture content. It achieved an accuracy of 98.2% , with data sorting capability in just 4.7 milliseconds, per frame using SGD optimization. Nevertheless, systems inability to handle bigger and tangled leaves.

S Datta[15] proposed 2D Convolutional Neural Network(CNN) for tea leaf disease detection. Here the Tea-leaf disease classifier has two key steps, image augmentation for data preprocessing and using CNN to help with the classification of images of diseased tea leaves. Achieving overall proposed model accuracy of 96.56% which exceeded the accuracy of other related models. Although, the proposed model could not identify a larger number of tea leaf diseases. In [17], J Yang developed semi-supervised learning models to identify tendon tea leaves with achieving accuracy 92.62%. The logistic regression model with the Adam optimizer algorithm improved accuracy and convergence. Nevertheless, misclassification occurs due to similar color between tender and old leaves.

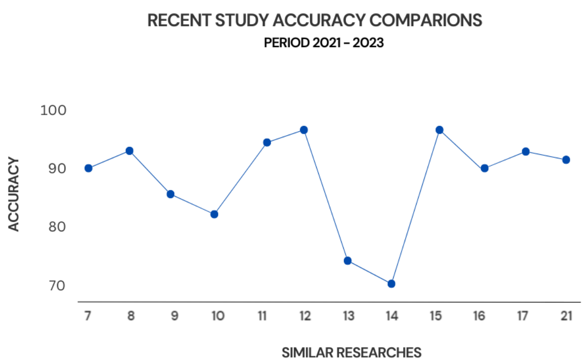

Fig. 1: Comparative analysis of Relative Study Accuracies

Figure 1 summarizes 12 recent studies from 2021 to 2023, showing varied approaches, none of which achieved accuracy levels comparable to our proposed method.

3 Methodology

This specific segment offers an in-depth summary of the techniques utilized for gathering data, along with the typical sequence of events in a machine learning study: data collection, data preprocessing, and model training. The entire procedure is succinctly represented in Figure 2.

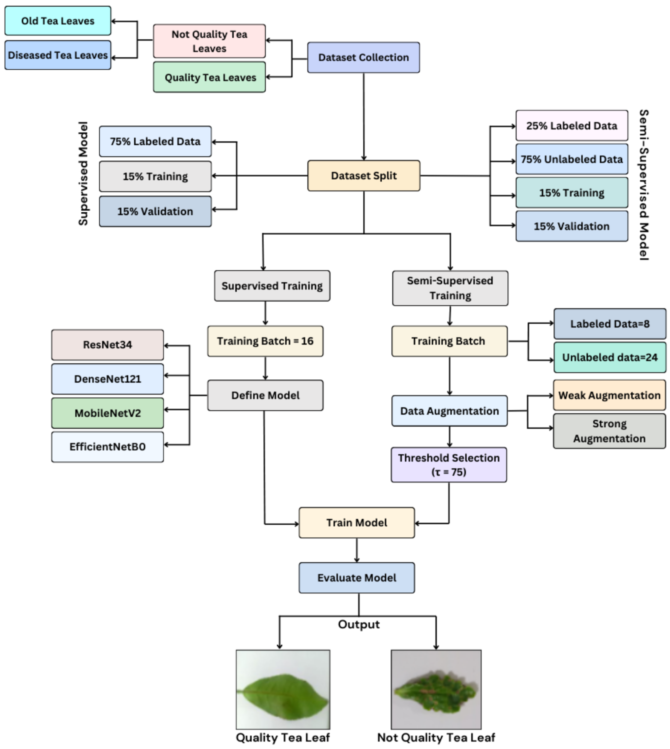

Fig. 2: Workflow diagram of Proposed models

3.1 Data Collection

Tea Leaf data was collected from Udalia Tea Garden in Chittagong shown in figure 3. We have collected approximately 800 images of tea leaves containing healthy leaves, diseased leaves assuring good image quality. In It depicts tea leaf image collection from various tea production and experimental fields in Bangladesh.

Fig. 3: Image Collection Procedure: Research Institute and Tea Garden

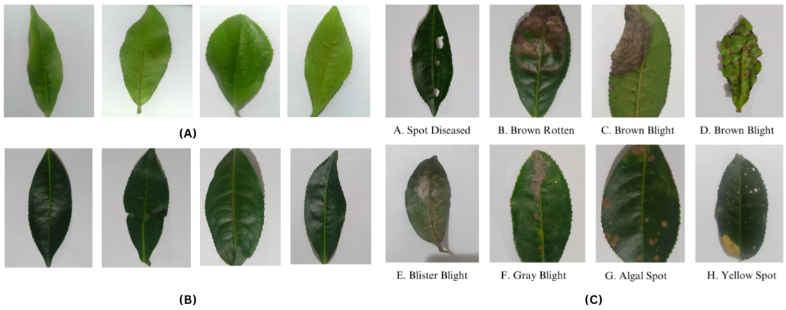

Fig. 4: From left to right (A)-(C): Tea Leaf Images - (A) Tea Making Leaves (B) Aged Tea Leaves (C) Diseased Tea Leaves

Figure 4 portraits a significant distinction between healthy young tea leaves and diseased or aged tea leaves from our sample dataset with exceptional clarity, highlighting the color differences. Figure 4(A) shows healthy young quality tea leaves. Figure 4(B) segment in the figure visualizes the Old tea leaves which have a less pliable texture with darker color. And figure 4(C), portrays the ineligible tea making leaf, that contains diseased. Here, spot diseases(A) were characterized by the distinct spots on the leaf, similarly leaves with Yellow spots(H), and small orange or reddish spots indicates algal infestation(G) and the leaf with brown patches(B), grayish patches(F) and rotten brown areas(C) indicating significant damaged tissues.

3.2 Data Processing

All images, both labeled and unlabeled, underwent normalization for training. This involved resizing the images to a common dimension of 224x224, a standard size for many pre-trained models, followed by conversion to tensors. For FixMatch specifically, a weak augmentation strategy was applied to both labeled and unlabeled data. This weak augmentation consisted of horizontal flipping and random rotations up to 10 degrees. Additionally, only unlabeled images received a stronger augmentation using the RandAugment function with parameters N=2 (number of transformations) and M=10 (magnitude of the transformations).

3.3 DenseNet121

DenseNet-121, a CNN model with 121 layers, primarily designed for enhanced image classification. For feature reuse, it employs dense blocks, where layers are interconnected to previous layers, through multiple bottleneck layers using a feedforward approach. The transition layer is used between dense blocks to reduce feature map outputs and computational costs. The function densenet121 creates a DenseNet-121 network, with two arguments: num_class and pre-trained. The num_class argument specifies the number of output classes, with a specified pre-trained model, the function loads and returns it; else, it returns a new instance [22]. The image has three colour channels(RGB). An optimizer, SGD used for minimize the loss function.

The model processes 16 images at a time. The ratio for training, testing, validation is 70% : 15% : 15%. Training consists of 20 epochs with a resized input image size of 224x224 pixels. The learning rate is 0.0001 and momentum is 0.9. Cross Entropy Loss is used as a loss function. Pretrained layers are frozen, and the dense layer 02 with units corresponding to the class uses the softmax activation function. Figure 05 shows the model's architecture, where the model normalizes the input image.

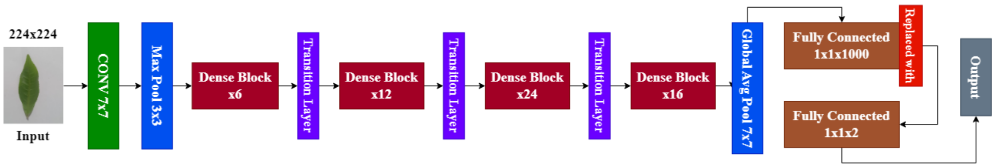

Fig. 5: DenseNet-121 Model Architecture

3.4 Fixmatch

Obtaining labeled data in the real world can be time consuming and expensive, therefore, by efficiently utilizing the unlabeled dataset using FixMatch can improve models performance, without any need for extensive labeled data which ensures robustness and cost efficiency. Fixmatch is a semi-supervised model that trains both labeled and unlabeled data to generalize better unseen samples. This algorithm combines both supervised cross-entropy loss on labeled data with weakly augmented images using pseudo-labeling. Primarily, it augments data through strong and weak methods to expand dataset variability. For simplicity training, it mostly favors stochastic gradient over Adam [19]. In Figure 06, the working procedure of fixmatch is portrayed, the labeled data trains the model and predictions are made for the unlabeled dataset and the confident predictions assign pseudo labels. The unlabeled data here uses multiple strong and weak augmentation to expand dataset variability, and creates different views.

For optimal p4 vcre-training of our model, using unlabeled data we used this approach. During training, it uses a threshold value of 75 to be selected for pseudo labeling. For batch sizes labeled is 8 and unlabeled is calculated as 8*3=24 where (mu=3) and validation conducted with a batch size of 8. The FM integrated model employs DenseNet121 architecture and optimises with a learning rate of 0.0001. For training 20 epochs have been used. The ratio for the training about 70% where 15% : 15% used for testing and validation.

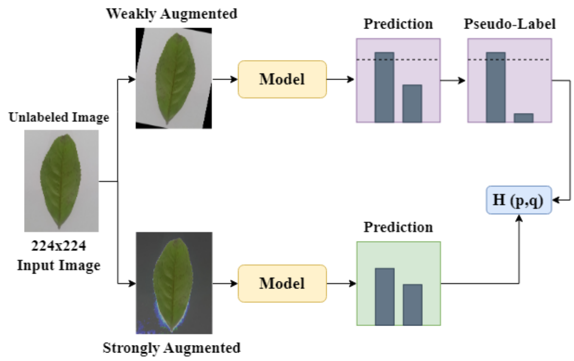

Fig. 6: Fixmatch Model Architecture

4 Result And Discussion

The performance evaluation of deep learning is a significant part of the whole process. This research calculates accuracy as the performance evaluation index. It is significant for the cases when all classes have equal significance. The accuracy is determined by the ratio of the number of correct predictions and the total number of predictions. Equations 1 to 4 represent the result for this study, which are average accuracy, precision, recall, and F1 score. Each trial's accuracy was noted, and the average is shown in the models section.

$$Accuracy = \frac{TP + TN}{FP + FN + TP + TN} \tag{1}$$

$$Precision = \frac{TP}{FP + TP} \tag{2}$$

$$Recall = \frac{TP}{FN + TP} \tag{3}$$

$$F1Score = 2 \cdot \frac{Recall \cdot Precision}{Recall + Precision} \tag{4}$$

Where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives [20]. We achieved the best results using deep learning models in our dataset, detailed in the following sections.

Table 1 summarizes the dataset distribution for training (supervised), unlabeled, labeled (semi-supervised), validation, and test datasets:

This table outlines the distribution of our datasets, with classes categorized as "Good" and "Bad". We used 70% of the data for supervised training and 25% labeled data for semi-supervised models, with 15 allocated for testing and validation.

Table 1: Number of Images For Training

Classes	Training	Unlabeled	Labeled	Validation	Test
Good	246	184	62	53	53
Bad	276	207	69	60	59
Total	522	391	131	113	112

4.1 DenseNet121 Model Evaluation

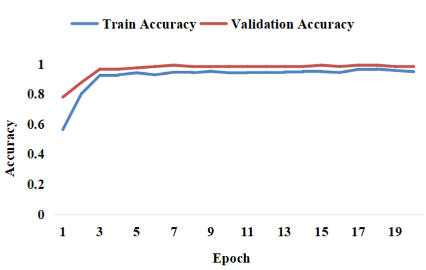

Fig. 7: Training and Validation accuracy curves for DenseNet121 model over epochs

Fig. 8: Training and validation loss curves for DenseNet121 model over epochs

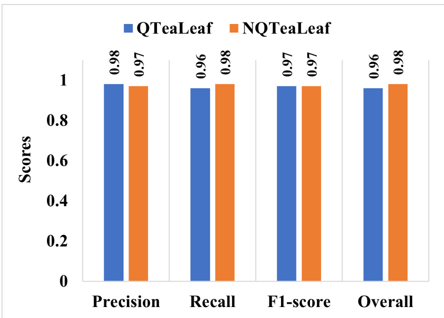

Fig. 9: Class-wise performance analysis for DenseNet121

In Figure 7, the training and validation accuracy curve and in figure 8, the training and validation loss curves for DenseNet121 model is shown, both illustrates stable performance without overfitting. Figure 9, distinguishes performance between two classes: Quality Tea Making Leaf and Not Quality Tea Making Leaf, with Precision, Recall, and F1-score values provided. Quality Tea Making Leaf achieved 96% accuracy, while Not Quality Tea Making Leaf achieved 98%.

4.2 Motivation Behind Semi Supervised Model

For regulating more efficient performance we opted a new approach with semi-supervised learning, strategically leveraging the dataset by utilizing only 25% labeled data. This decision was premised on the hypothesis that a semi-supervised model, which learns from both labeled and unlabeled data, could optimize the utility of the dataset. Notably, despite using a limited portion of data, the semi supervised model yielded a similar accuracy as the supervised model.

The advantages of the semi-supervised model are multifaceted. It mitigates the reliance on extensive lebeled dataset, which are often costly and labor intensive. In agricultural contexts, The model ensures several benefits including harvested leave quality assurance and adherence to legal and ethical policies. Thus it reduces manual labor , human error and increase efficiency. Additionally, offers a real-time quality assessment that aids farmer with quick insights and maintainence. By efficiently using a smaller labeled dataset along with a larger pool of unlabeled data, we can significantly cut down on data labeling costs and efforts. Thus, it provides a feasible and cost-effective solution with substantial quality improvement.

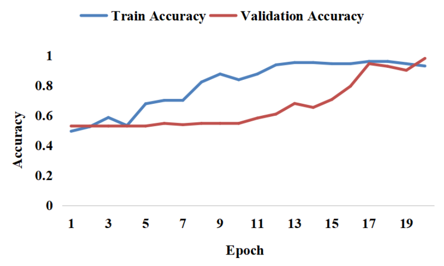

Fig. 10: Training and validation accuracy curves for FixMatch model over epochs

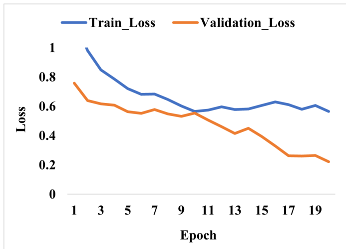

Fig. 11: Training and validation loss curves for FixMatch model over epochs

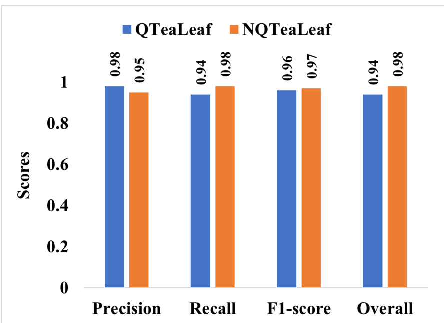

Fig. 12: Class-wise performance analysis for FixMatch

Here, Figure 10's curves illustrates training and validation accuracy curve for fixmatch model, and figure 11 demonstrates a gradual downturn in training and validation losses over epochs, indicating improved model performance with training. And in Figure 12, FixMatch's class-wise analysis showed 94% accuracy for Quality Tea Making Leaf and 98% for Not Quality Tea Making Leaf.

4.3 Overall Comparision Of Performance

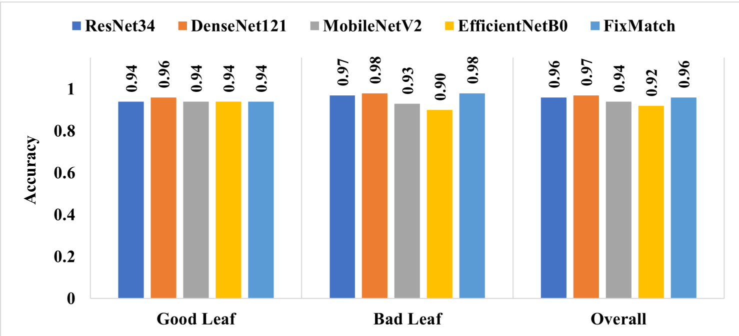

Fig. 13: Comparative analysis of ResNet34, DenseNet121, MobileNetV2, Effcient-NetB0 and Fixmatch.

Figure 13, highlights the overall comparative analysis of the used models, ResNet34, DenseNet121, MobileNetV2, EffcientNetB0 and Fixmatch, among "Good Leaf", "Bad Leaf" and an Overall accuracy. In each class, DenseNet121 seemingly outperformed all other models in Supervised model and an identical performance was demonstrated by Fixmatch.

4.4 Limitations

As for the constraints, the semi-supervised approach reduces the need for whole labeled data, but requires a substantial amount of initial labeling, which can impact scalability. Further evaluation can be made to assess efficacy for future appliances.

4.5 Suggestions for Future Research

In this section, the following suggestions aim to enhance robustness, performance and applicability of tea leaf classification systems, providing future research scope to various stakeholders in the tea industry.

5 Conclusion

In this study, we have assessed the effectiveness of diverse deep learning models for categorizing tea-making foliage. Our primary aim was to identify leaves suitable for tea production. A significant approach between, semi-supervised model and supervised model is visible. Our supervised model achieved 97% classification accuracy from DenseNet121, whereas despite using only a few sets of labeled data, the semi-supervised model yielded a similar 96% accuracy, which dictates the potential advantages to suppress the use of labeled dataset in model training. Overall, the triumph attained by these methodologies will enhance the quality, efficiency, and durability of Bangladesh's tea sector for forthcoming generations.

References

[1] Food and Agriculture Organization of the United Nations. (2022). International tea market: Market situation, prospects and emerging issues. Mark. Trade Div. Econ. Soc. Dev. Stream, CC0238EN.

[2] Ono, W. (2021). Introductory Overview of the History of Global Tea Production. International Council on Monuments and Sites 11, rue du Séminaire de Conflans 94220 Charenton-le-Pont France, 7.

[3] Basu Majumder, A., Bera, B., Rajan, A. (2010). Tea statistics: global scenario. Inc J Tea Sci, 8(1), 121-4.

[4] A. Chen, "The World's Top 10 Tea Producing Nations," WorldAtlas, Apr. 20, 2015 [Available online athttps://www.worldatlas.com/articles/the-worlds-top-10-tea-producing-nations.html]

[5] "Bangladesh's tea production in 2023 hit all-time high," nationthailand, Jan. 19, 2024. [Available online at https://www.nationthailand.com/world/asia-pacific/40034830]

[6] Ahmad, I., Hossain, M. A. (2013). Present status of Bangladesh tea sector in respect of world tea: an overview. J. Appl. Sci. Technol, 9(1), 116-128.

[7] Chen, J., Jia, J. (2020). Automatic recognition of tea diseases based on deep learning. Advances in forest management under global change, 11, 13.

[8] Hu, G., Wang, H., Zhang, Y., Wan, M. (2021). Detection and severity analysis of tea leaf blight based on deep learning. Computers Electrical Engineering, 90, 107023.

[9] Cao, J., Wu, Z., Zhang, X., Luo, K., Zhao, B., Sun, C. (2023). Sorting of Fresh Tea Leaf Using Deep Learning and Air Blowing. Applied Sciences, 13(6), 3551.

[10] Banerjeea, P., Barnwala, R. P.(2021) Visual explanation using Grad-CAM for deep-learning based classification of Tea-leaves.

[11] Dejene Teka (2024). Automatic Tea Leaf Classification Using Deep Learning Models. Department of Information Technology, Mattu University, Mattu, Ethiopia

[12] Kaihua Wei, Bojian Chen, Zejian Li, Dongmei Chen, Guangyu Liu, Hongze Lin, ORCID and Baihua Zhang (2022). Classification of Tea Leaves Based on Fluorescence Imaging and Convolutional Neural Networks. MDPI Open Access Journals.

[13] G.M.K.B. Karunasena, H.D.N.S. Priyankara (2020). Tea Bud Leaf Identification by Using Machine Learning and Image Processing Techniques. International Journal of Scientific Engineering Research Volume 11.

[14] Iromi R Paranavithana, Viraj R Kalansuriya (2021). Deep Convolutional Neural Network Model for Tea Bud(s) Classification. IAENG International Journal of Computer Science, 48:3, IJCS 48 3 17, Volume 8.

[15] Saikat Datta, Nitin Gupta (2023) .A Novel Approach For the Detection of Tea Leaf Disease Using Deep Neural Network. Dept. of Computer Science and Engineering, National Institute of Technology Hamirpur, Hamirpur(HP), 177005.

[16] Wu, W. (2022). Identification of tea leaf diseases based on deep transfer learning. Frontiers in Computing and Intelligent Systems, 2(3), 75-77.

[17] Yang, J., Chen, Y. (2022). Tender Leaf Identification for Early-Spring Green Tea Based on Semi-Supervised Learning and Image Processing. Agronomy, 12(8), 1958.

[18] J. E. van Engelen and H. H. Hoos (2019), "A survey on semi-supervised learning,". doi: https://doi.org/10.1007/s10994-019-05855-6.

[19] K. Abeywardana,(2024)"FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence: A Brief Summary," Medium. [Available online at https://medium.com/@kdwa2404/fixmatch-simplifying-semi-supervised-learning-with-consistency-and-confidence-a-brief-summary-cad2d8044ec0.]

[20] Koo Ping Shung, "Accuracy, Precision, Recall or F1?," Towards Data Science, Mar. 15, 2018. https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9

[21] Bikash Chandra Karmokar , Mohammad Samawat Ullah , Md. Kibria Siddiquee and Kazi Md. Rokibul Alam, "Tea Leaf Diseases Recognition using Neural Network Ensemble", Mar. 2015, DOI: 10.5120/20071-1993

[22] Shuvam Das, "Implementing DenseNet-121 in PyTorch: A Step-by-Step Guide" Medium,(2023). [Available online at https://medium.com/deepkapha-notes/implementing-densenet-121-in-pytorch-a-step-by-step-guide-c0c2625c2a60]