TEAQNET: A VISION-BASED DEEP LEARNING FRAMEWORK FOR EXPERT-LEVEL TEA REGION CLASSIFICATION USING STANDARDIZED VISUAL EVALUATION ENVIRONMENT.

By

A.E.P.P. Jayasekara

PS/2020/010

A thesis submitted in partial fulfillment of the requirements for the degree of

BACHELOR OF SCIENCE HONOURS

in

Computer Science

DEPARTMENT OF STATISTICS & COMPUTER SCIENCE

FACULTY OF SCIENCE

UNIVERSITY OF KELANIYA, SRI LANKA

2025

DECLARATION

I would like to declare that this thesis is my original work and has not been submitted previously for a degree at this or any other university/institute. To the best of my knowledge, it does not contain any material published or written by another person and I have correctly acknowledged the work of others.

Candidate: A. E. P. P. Jayasekara Signature: ....

Date: ....

As supervisor of the thesis, I certify that the above statements are true to the best of my knowledge.

Supervisor: Dr. B. M. T. Kumarika Signature: ....
Senior Lecturer,
Department of Statistics & Computer Science,
University of Kelaniya,
Sri Lanka.

Date: ....

Coordinator: Dr. B. M. T. Kumarika Signature: ....
Senior Lecturer,
Department of Statistics & Computer Science,
University of Kelaniya,
Sri Lanka.

Date: ....

Abstract

Tea is one of the most widely used drinks globally. Among the various tea producing nations, Sri Lanka is known as Ceylon which holds a prestigious place in global tea industry, and it has global stamp as Ceylon tea. Ceylon tea has good market value, quality and global branding depending on various tea regions in Sri Lanka. Traditionally, tea testers use sensory based methods to manually classify the tea regions based on their judgments vary with their experience, standard lighting and environmental conditions with limiting reproducibility and scalability. This study introduces standardized vision based deep learning framework to automate tea region classification based on tea liquor color which directly influenced by chemical composition and processing variations. The research has been constructed in a controlled vision evaluation environment to collect tea liquor image dataset. The dataset has preprocessed to remove the background reflection, and it has increased the data augmentation to improve robustness. It archives expert level classification across Sri Lankan tea regions using Convolution Neural Networks (CNN). That CNN based architecture has trained on region labeled samples and evaluated using accuracy, precision, recall, f1-score and intersection over union (IOU). The results reveal that the ShuffleNet V2 model is the most effective having a mean accuracy rate of 99.99 %. The Model demonstrates strong potential as a non-destructive, scalable tool for real-time industry deployment, particularly when integrated into mobile applications for on-site testing. To summaries, these findings highlight the importance of standardizing visual evaluation environment in improving the reproducibility of tea region classification. The future studies should explore larger datasets, cross factory domain adaptation and multimodal fusion to further advance in automated tea evaluation technologies.

Dedication

I dedicate this work to my parents whose support, love encouragement has fulfilled every movement of my academic life. They believed my abilities and their patience during movement of doubt and constant motivation gave determination to pursue this project with diligence and passion. This work is a reward for their sacrifices of every financial support given for personal growth.

I dedicate this research to all tea growers, producers and enthusiasts. Because their tireless efforts sustain the rich cultural and economic heritage of tea around the world. Their commitment has preserved the unique characteristics of each tea growing region. It has inspired the conceptualization of this framework which aims to combine traditional expertise with modern technological advancements.

Furthermore, I dedicate this research project to the broader community of researchers and innovators in Computer Vision and Artificial Intelligence. Their contribution to the field has paved the way for the development of vision-based frameworks like TEAQNET.

I hope this work contributes even in a small way. But the advancement of intelligent systems has bridged human expertise with automated, standardized evaluation methods. It also supports the future of precision agriculture and quality control in the tea industry.

Through this dedication, I acknowledge that every achievement in this project is rooted not only in my effort. It has a part for who gave me guidance, inspiration and support, nurtured me, motivated me and provided the foundation for this work.

Finally, I acknowledge the contribution of the broader AI and Computer Vision communities. Those include the developers and maintainers of libraries such as PyTorch, TorchVision, OpenCV and Sklearn whose work provided the foundational tools required to implement this research. This project has succeeded because of the collective effort, guidance and support of everyone who has committed to it. I am sincerely thankful to everyone who plays a part in making this project possible to complete.

Acknowledgments

I wish to express my deepest gratitude to all who guided, supported and inspired me to complete this research project. First and foremost, I am profoundly thankful to my research supervisors Dr. Thosini Kumarika and Dr. Dakshika. Their invaluable guidance, mentorship and constructive feedback are at every stage of this project. Their expertise in Machine learning, Computer Vision and domain-specific knowledge in agriculture and chemistry gave proper guidance throughout this entire research project.

I would also like to extend my gratitude to the Department of Computer Science and Department of Chemistry for providing the necessary lab facilities and technical support. It enabled the experimental and practical aspects of this project.

I am equally grateful to my colleagues, friends and peers who contributed through engaging discussions, brainstorming sessions and moral support throughout the research. Their constructive critiques, enthusiasm and encouragement have motivated me during the challenging phases of experimentation and model development.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

1. Chapter 01 - Introduction

Tea remains one of the most used beverages in the world, and it is of immense cultural, economic and social significance as one of the leading export crops produced in most nations, including Sri Lanka. The regions of tea determine its price in the marketplace as well as among consumers. Historically, tea regions have been identified by skilled human panelists who subjectively judge the appearance of dry tea leaves, color of liquor, aroma and taste based on sensory evaluation. These traditional methods are subjective, labor-intensive and susceptible to inter-and intra-observer variation. Therefore, it is not reproducible and scalable at a commercial scale. There are a massive number of parameters used for identifying the regions of tea, such as chemical parameters, sensory parameters and physical parameters. In physical parameters, the researchers observe the dry leaf appearance, leaf infusion, particle size, and liquor color [1].

Tea color in its dry form as well as after brewing is one of the various quality parameters among the various ones, a prominent visual parameter associated with its chemical composition and processing characteristics. These compounds are responsible for establishing tea's appearance and are also associated with flavor, aroma, and antioxidant activity. The researchers have investigated how infusion color and taste is related to chemical compounds. Those are mainly Theaflavins, Kaempferol, Astragalin and 5-p-coumaroylquinic acid [2]. These polyphenolic compounds are chemically modified during processing, especially during oxidation stages. For instance, theaflavins impart brightness and briskness, whereas Thearubigins (polymeric polyphenols) impart fullness and body to black tea liquor [3]. Since artificial intelligence and deep learning, it is rapidly moving ahead, applications of machine vision have acquired massive usages in fields of food science and agriculture as well. Convolutional Neural Networks (CNNs) showed improved object detection as well as feature extraction of intricate visual information [4]. Their ability to recognize and classify specific areas in pictures well places them for quality indication through the assessment of subtle differences in tea leaf and liquor color. However, image-based approaches rely heavily on stable and controlled imaging conditions to counterbalance the influence of extraneous sources such as lighting, background, and camera settings [5].

This proposes TEAQNET, a deep learning-based expert-specific design with CNNs within a standardized image-based testing platform to identify the tea regions. The proposed design is set to bridge human sensory classification by experts and prevailing automated systems based on the usage of visual expert-opinion-driven features and chemical composition. The target solution is expected to enhance objectivity, scalability, and consistency of identifying tea regions and thereby help ensure tea safety as well as bring more pervasive improvements to precision agriculture and smart food systems.

1.1. Background of the Tea Industry

1.1.1. Global Significance

The popularity of tea continues to grow due to its natural origin, diverse flavor profiles and well-recognized health benefits. Over the past decades, world tea production has increased at an annual rate of approximately 4% and reached 5.73 million tons in 2016, with China, India, Kenya, Sri Lanka and Indonesia. For many of these countries, tea remains not only a major source of employment. But also, it is a critical component of national export revenue, rural livelihood and socio-economic development.

Figure 1: 1.1.1 A rich black tea from the Yunnan province [6]

The global tea industry has undergone considerable diversification due to the changing lifestyles, increased urbanization and consumer expectations. The growth of the market for specialty teas, such as herbal infusion, fruit fusions, functional blends and flavored gourmet teas, has broadened the market appeal to consumers seeking novel sensory experiences and targeted health benefits.

The global market valuation has expanded due to the robust performance of the tea sector, reflected in economic terms. It was estimated at USD 46 billion in 2026 and projected to reach approximately USD 67 billion by 2023, underscoring the industry's long-term growth potential and resilience. [7].

The Asia-Pacific tea market reached USD 101.20 billion in 2024, and it shows how the region leads global consumption. Asia still dominates the market share. China generated nearly USD 51 billion in 2025, and it was five times bigger than India’s market.

1.1.2. Sri Lankan tea industry

The Sri Lankan tea industry has long been recognized as one of the most significant sectors in the national economy, contributing substantially to export earnings, rural employment, and long-term socio-economic development. Since the introduction of tea cultivation in the late nineteenth century in Sri Lanka, formerly known as Ceylon, it has developed a global reputation for producing premium-quality orthodox teas, which are widely marketed under the internationally renowned brand "Ceylon Tea." For decades, the tea industry has functioned as the country's primary generator of foreign exchange and the largest employer in the plantation sector. It is a significant source of government revenue.

Figure 2: 1.1.2 Ceylon tea leaves [8]

Sri Lanka's economic and social landscape establishes its relevance not only as a hub for agricultural commodities but also as a foundational pillar of national development. Significant changes have occurred in Sri Lanka, and current data indicate that the sector's relative economic significance has decreased gradually compared to other developing industries, such as textiles, IT, and services. Although tea remains a significant export commodity, its share in total export earnings and production volumes has declined. This decline has been attributed to multiple structural challenges that undermine industry's competitiveness in an increasingly dynamic global tea market.

Figure 3: 1.1.2 Plucking Ceylon tea [9]

The Major challenge in the Sri Lankan tea industry was the limited availability of land for tea cultivation. The total geographical area under tea has remained stagnant for several years, but some regions have been experiencing reductions due to urbanization, crop diversification, and environmental degradation. Due to its limited total production capacity, Sri Lanka finds it challenging to continue being one of the world's leading producers of tea. The low rate of replanting, which limits the replacement of ageing tea bushes with high-yielding varieties, exacerbates this problem. Low replanting rates result in reduced productivity per hectare, negatively impacting both the quality and quantity of national output. Labor shortages also present a serious concern. Usually, the tea sector has relied on labor-intensive harvesting techniques, but the industry now faces challenges in attracting and retaining workers. Younger generations are increasingly preferring employment in non-agricultural sectors, which offer higher wages and better working conditions. This labor scarcity be a reason for raising production costs and limits timely harvesting which in turn affects the quality of exports. Rising wage demands, increased input costs, and inefficiencies in plantation management has contributed to higher overall production costs compared to those of competitor nations, such as Kenya and India.

Table 1: 1.1.2 Trends in Key Tea Industry Indicators in Sri Lanka

Indicator	Current Status	Implication
Land under cultivation	Stagnant / declining	Limits production expansion
Replanting rate	Low	Lower long-term productivity
Labor availability	Declining	Difficulty meeting harvesting demands
Production cost	High	Weakens price competitiveness
Global market share	Decreasing	Loss of competitiveness to Kenya, India

To accomplish the industry's competitive position in analytical frameworks such as Porter's Five Forces Model and the Diamond Model of National Competitive Advantage gives valuable insights. The Five Forces analysis reveals the intensity of competition in the global tea market. Those are exceptionally high, driven by substitute beverages, high buyer bargaining power, and the entry of low-cost, competitive producers. Meanwhile, the Diamond Model reveals that Sri Lanka possesses both favorable and unfavorable conditions. For global brand identity, favorable elements play a crucial role and include agro-climatic diversity that produces teas of unique character, and a long-established export infrastructure. In contrast, unfavorable conditions include limited factor conditions such as high-cost structures, labor shortages, and inadequate technological advancement.

Table 2: 1.1.2 Competitiveness Assessment of the Sri Lankan Tea Industry

Diamond Model Factor	Strengths	Weaknesses
Factor Conditions	High-quality agro-climatic zones	High labor cost, aging plantations
Demand Conditions	Strong global demand for orthodox tea	Limited domestic demand for innovation
Related & Supporting Industries	Established auction system	Limited technological modernization
Firm Strategy & Rivalry	Strong global brand identity	Weak cost competitiveness vs. Kenya

These challenges require well-coordinated strategies and policy interventions, which are strengthening replanting programs and encouraging mechanization where feasible, enhancing worker welfare, and promoting value addition. It is an essential step toward revitalizing the industry.[10].

1.1.3. Importance of tea region classification

Tea region classification plays a vital role in the global tea industry in countries like Sri Lanka due to the Agro-climatic diversity directly influence the sensory, chemical and commercial attributes of tea. Classification of tea according to the geographical origin is known as terroir differentiation. It is a critical mechanism for establishing quality standards, market identity and strategic branding. In Sri Lanka there are several tea regions based the agroeconomic zones. Those are Dimbula, Nuwara Eliya, Uva, Sabaragamuwa, Kandy, Ruhuna and Udapussellawa. Each tea region has unique climatic and soil conditions that can contribute to the aroma, flavor, color and chemical composition of the tea [11].

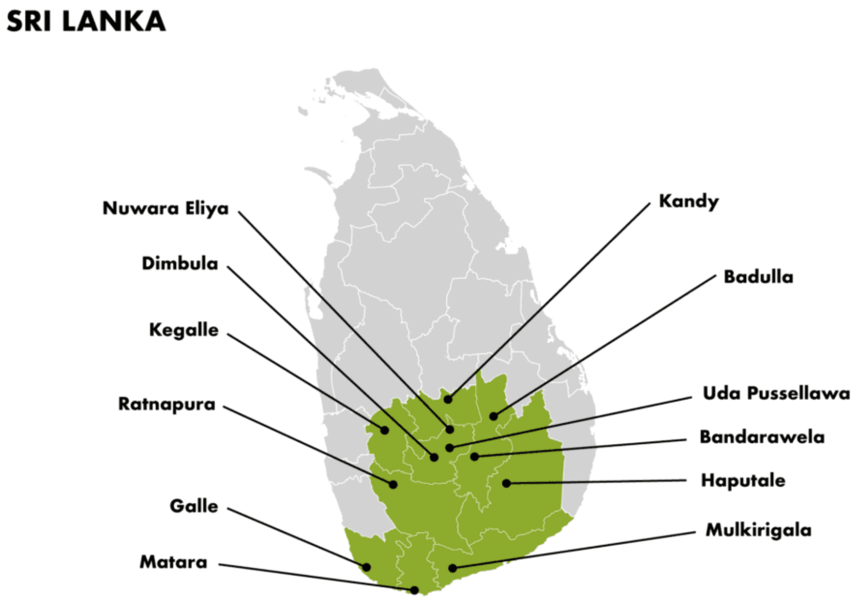

Figure 4: 1.1.3 Sri Lankan Tea Regions [12]

Tea Region classification can be considered as a function that controls the value chain differentiation and market segmentation. International buyers and consumers prefer to buy region specific teas, due to their perceived quality, traceability and cultural heritage. Therefore, classification and proper labeling of tea based on the geographical origin would be a strategic tool for strengthening the country’s global market position [13].

Tea region classification has gained importance through advance in analytical chemistry, sensor technologies and the machine learning in scientific and technological domains. Modern analytical methods such as gas chromatography-mass spectrometry (GC-MS), near-infrared spectroscopy and electric nose (E-Nose) systems have been used to identify the biochemical markers associated with geographical origins. Those are included catechins, polyphenols, amino acids and volatile compounds. Their concentration varies according to elevation, rainfall, soil PH and sunlight exposure [14].

Figure 5: 1.1.3 E-Nose [15]

The tea region classification contributes significantly to rural livelihoods and sustainable agricultural planning for economic development perspective. The region-based data allows policy makers and agronomists to implement target interventions due to the tea growing regions differing in their productivity levels.

1.2. Limitations of Traditional Methods

1.2.1. Sensory evaluation

Sensory evaluation is a scientific discipline that measures, analyzes and interprets human responses to the sensory characteristics of tea. It includes aroma, flavor, mouthfeel and after taste. In tea industry sensory evaluation is a crucial step for quality control, product grading and the market classification. Tea quality cannot be assessed solely through chemical or physical analysis. But sensory analysis integrates human perception with standardized evaluation protocols for ensuring the consistent of quality across batches and regions.

Table 3: 1.2.1 Key Sensory Attributes Evaluated in Cupping

Attribute Category	**Specific Parameters**	**Description**
Appearance	Dry leaf color, size, twist	Indicates leaf quality and processing consistency
Aroma	Floral, woody, fruity, toasty notes	Derived from volatile compounds and regional climatic differences
Flavor	Strength, briskness, sweetness, bitterness	Primary factor determining consumer preference
Liquor Color	Brightness, clarity, depth	Reflects chemical composition such as theaflavins/thearubigins
Mouthfeel	Smoothness, thickness, astringency	Influenced by tannins and polyphenols
Aftertaste	Duration, pleasantness	Important for identifying premium-quality teas

Sensory Evaluation plays a fundamental role in maintaining the global reputation of Ceylon Tea. Sri Lankan Tea Board and the International bodies like ISO 3103 have defined standard cupping methods and those are used by trained industry professionals (TEA TESTERS) who are conducting systematic evaluations. These sensory evaluations help to identify regional authenticity and verify terroir related to the attributes such as briskness in Uva teas, floral notes in Dimbula teas and full-bodied character in low-grown Ruhuna teas [16].

Table 4: 1.2.1 Comparison of Sensory Methods

Method	Description	Advantages	Limitations
Traditional Cupping (ISO 3103)	Standard hot infusion tasting by trained experts	Highly reliable; captures holistic perception	Require extensive training
Descriptive Sensory Analysis	Uses trained panels to quantify specific attributes	Detailed profiling, repeatable	Time-consuming and expensive

Electronic Nose / E-Nose	Sensor array that detects volatile compounds	Objective, fast screening	Limited compared to human perception
Machine Learning Prediction	AI analyzes chemical/sensory data to classify quality	High accuracy, scalable	Require large datasets

Sensory Evaluation also connects human perception with biochemical properties. Those sensory characteristics are influenced by tea quality attributes such as polyphenols, catechins, caffeine, theaflavins and volatile aromatic compounds. Researchers indicate that expert tester can detect subtle differences in chemical compositions which correspond to aroma compounds like linalool and geraniol [17].

Furthermore, sensory evaluation contributes to product development, market segmentation and consumer preference research. There is a high demand for specialized products such as green tea, herbal blends, flavored teas and ready to drink beverages. It guides us to understand how consumers perceive sweetness, bitterness, aroma intensity and mouthfeel which are optimizing formulations.

Recent advancements integrate sensory science with machine learning and digital tools. Artificial Intelligence (AI) models like convolutional neural network (CNN), support vector machine (SVM) and Electronic Nose (E-Nose) systems can analyze aroma compounds, leaf texture and liquor color to predict sensory quality attributes. These enhance accuracy and provide scalable solutions for quality assurance.

1.2.2. Subjectivity

Subjectivity plays a pivotal role in the evaluation, grading and classification of tea. Traditionally, tea region classification relies expert human assessment where tea tasters evaluate visual, olfactory and gustatory attributes to determine the tea regions. But the experienced tasters may classify the same tea samples differently because the sensory differences between regions are subtle.

Environmental subjectivity also plays a major role. The factors such as room lighting, ambient temperature, cup color and surrounding odors also can influence the perception during visual and aroma evaluation. Further water temperature, brewing time and the leaf quantity also affect the liquor appearance and flavor intensity [18].

Furthermore, linguistic ambiguity also contributes to classification subjectivity. In industry, tea tasters use the qualitative descriptors such as “brisk”, “soft”, “sparkling”, “floral” or “mellow”. But those varies across evaluators. Researchers have found that when attributes overlap across different regions, the professionals define sensory terms differently.

Table 5: 1.2.2 Sources of Subjectivity in Tea Region Classification

Source	Description	Impact on Classification
Physiological variability	Differences in sensory sensitivity and taste receptor genetics	Inconsistent perception of regional flavor notes
Cognitive bias	Preconceived expectations of regional profiles	Expectation-driven misclassification
Sensory fatigue	Reduced sensitivity due to repeated tasting	Lower ability to detect subtle regional differences
Environmental factors	Lighting, noise, odor, cup color, steeping conditions	Altered perception of leaf color and liquor attributes
Vocabulary differences	Varied interpretation of sensory descriptors	Communication errors and inconsistent scoring
Sequence effects	Previous sample influencing evaluation of the next	Contrast or masking effects

1.2.3. Environmental variation

Environmental variation plays a central role in distinguishing the growing of tea in Sri Lanka. Each tea region has unique agroecological conditions that are shaped by elevation, temperature, rainfall, soil type, humidity and wind patterns. These factors influence the biochemical composition of tea leaves that produce distinct sensory and chemical signatures for regional classification. These are essential for improving accuracy and supporting quality assurance, market valuation and origin authentication.

Elevation is one of the most significant determinants of environmental variation. The High grown regions in Sri Lanka such as Nuwara Eliya and Udapussellawa has experienced low temperatures, high UV exposure and lower leaf growth. It is leading to lighter liquors and elevated levels of aromatic volatiles. The low grown tea regions such as Ruhuna and Sabaragamuwa are categorized by warm, humid climates and rapid leaf growth. It results in darker liquor, stronger flavor profiles and higher polyphenolic content.

Climatic variation is another significant determinant of environmental variation. Uva region has significant withering characteristics due to reduced humidity and dry winds. In other hand Nuwara Eliya region has floral softness and Dimbula region has brisk brightness due to the weather conditions. Reduce rainfall and increased sunlight enhance the photosynthesis which resulting in higher catechin to theaflavin and it directly affect tot the regional flavor.

Figure 6: 1.2.3 Climate Changes affect tea [19]

Soil composition is another significant determinant of environmental variation. In Sri Lanka high grown regions typically contain well-drained, sandy loam soils with low nutrient density. It encourages slow nutrient uptake and complex flavor development. In other hand low grown areas often contain iron rich alluvial soils and promoting vigorous growth and stronger liquors. Amino acid and polyphenol biosynthesis are influenced by soil PH, mineral balance and organic content [20].

Humidity and rainfall patterns also introduce significant environmental variation. In Sri Lanka, Sabaragamuwa region has higher humidity compared to other regions. Therefore, it has dark leaf color and greater tannin concentration. Uva and Nuwara Eliya regions have lighter liquors and higher aroma intensity.

Table 6: 1.2.3 Key Environmental Variables Influencing Sri Lankan Tea Regions

Environmental Factor	**Description**	**Influence on Tea Properties**
Elevation	High, mid, and low-grown regions	Aroma intensity, polyphenol content, liquor color
Temperature	Regional and seasonal variations	Aroma development, bitterness, astringency
Rainfall	Monsoonal vs. inter-monsoon patterns	Leaf growth rate, flavor concentration
Humidity	Varies significantly by region	Withering characteristics, aroma compounds
Soil composition	Mineral levels, pH, organic matter	Biochemical composition of leaves
Wind patterns	Uva dry winds, highland breezes	Volatile compound formation
Solar radiation	High in mountain regions	Photosynthesis rate and flavor precursor formation

Table 7: 1.2.3 Environmental Profiles of the Seven Tea Regions

Region	**Elevation**	**Climate Attributes**	**Environmental Signature**
Nuwara Eliya	>1800 m	Cool, misty, low humidity	High aroma volatiles, light liquor
Uva	900–1500 m	Dry monsoon winds, moderate rainfall	Strong aroma, unique “Uva character”
Dimbula	1100–1700 m	Cool, seasonal variation	Bright liquor, brisk flavor
Kandy	600–1200 m	Warm, moderate rainfall	Stronger flavor, mid-grown character
Udapussellawa	1200–1600 m	Cool, cloudy climate	Rosy notes, medium body

Ruhuna	0–600 m	Hot, humid, high rainfall	Strong liquor, high polyphenols
Sabaragamuwa	0–900 m	Wet, humid, tropical	Darker liquor, soft texture

1.2.4. Inter- and intra-observer variation

Inter- and intra-observer variation represents a major source of inconsistency in traditional tea region classification practices. Usually, tea regions are evaluated using human sensory judgement. The accuracy of that classification depends on human perception of aroma, liquor color, briskness, mouthfeel and astringency. Both inter-observer variation (differences between evaluators) and intra-observer variation affected the classification reliability and reproducibility.

Inter-observer variation occurs due to several factors. Those are prior experience, training quality, cognitive bias, cultural background and physiological sensitivity. For example, tasters who are exposed to high grown teas can be easily identified the floral and aromatic markers of Nuwara Eliya teas compare to the less experienced evaluators. Similarly, tasters who are exposed to strong, low-grown teas are consumed may unconsciously favor characteristics such as liquor thickness and color depth, leading to misclassification of subtle, lighter high-grown teas.

Intra-observer variation occurs due to several factors. Those are inconsistent judgments at different times, fatigue, adaptation, mood, distraction or temporary physiological changes. The research shows that repeated exposure to samples during tasting sessions causes olfactory adaptation, which diminishes the perception of volatiles associated with regional identity.

Environmental factors further amplify observer variation. Those factors are tasting bowls, lighting quality, ambient odors and water temperature. When following the ISO 3103, the minor derivation can be reduced in brewing conditions which affect the sample appearance and flavor.

1.3. Motivation for Automation

1.3.1. AI in agriculture

Artificial Intelligence (AI) has crucial role in modern agriculture, and it is reshaping the way crops are cultivated, monitored and managed. Since the global population is expected to reach 9.7 billion by 2050, traditional agriculture methods need to be updated for enhancing productivity, reducing environmental impact and optimizing resource utilization.

In Precision agriculture, AI is used for site specific crop management, and it enhances accuracy and scalability. Using Machine learning (ML) algorithms can be analyzed in large datasets and it is easy for remote sensing with using drones, IOT (Internet of things) and satellite to monitor crop health, detect nutrient deficiencies and predict growth patterns.

Figure 7: 1.3.1 A tea picking robot [21]

In Sri Lanka, traditional classification relies on sensory evaluation, where trained tea tasters assess liquor color, aroma, flavor, astringency and briskness. For Automation AI systems can analyze thousands of samples quickly and consistently. The objective of this is to eliminate human bias in sensory perception. AI can identify micro-level differences which are not visible to human evaluators and can decrease the dependence on large, highly trained sensory panels.

1.3.2. Need for standardized color-based evaluation

Tea region classification is a multidimensional process which relies on the sensory assessment of aroma, flavor, liquor color and infusion characteristics. Among these parameters, color plays a major role for determining the market value and the geographic identify of tea. In Sri Lanka there are 7 regions of tea, and each region produce teas with distinct liquor hues. Therefore, standardizing color-based evaluation is essential for accurate tea region classification.

The tea liquor color is associated with polyphenol concentration, catechin oxidation, amino acid profiles and climate interactions. As an example, Nuwara Eliya teas have typically light golden yellow hue, Uva teas have bright reddish liquor with distinct briskness, Dimbula teas show an amber to golden appearance due to monsoon winds and cool climates and Ruhuna teas show dark, full-bodied liquor influence by low-elevation conditions.

Human perception for tea region classification is subjective. Because tea tastes rely on experience, memory and perception. Those are influenced by fatigue and sensory adaptation, personal preferences, Environmental lightning conditions and psychological biases. The tea also can be influenced by lightning conditions such as artificial light, time of delay and ambient color environment. This makes it very difficult to ensure uniformity across different tasting panels and factories.

1.3.3. Real-time classification challenges

The increasing adaptation to new technologies like Artificial Intelligence (AI), Internet of things (IOT) and computer vision have opened many avenues for automatic tea region classification. In Sri Lanka, there are seven geographical area which are famous for teas and each region have their own characteristics. It can significantly enhance quality assurance, traceability and market competitiveness.

Sri Lanka's tea industry has faced more pressure from different areas. Those are shifting global market preferences, rising production costs, labour shortages, climate variability and market competition from Kenya, India and emerging producers. The key challenges in real time tea region classification are maintaining data quality and heterogeneity. Since AI models heavily rely on large, high quality and well-labeled datasets, it has faced multiple challenges. Those are variation in lightning conditions, differences in imaging devices and limited annotated data.

1.4. Problem Statement

1.4.1. Lack of reproducible classification

Reproducibility is a critical requirement in any scientific and operational classification system. It has applicable in agricultural quality assessment in stakeholders' decisions which are pricing, market positioning and geographical indications. Sri Lankan tea region classification has significant reproducibility challenge due to variations in sensory evaluation, environmental factors, data acquisition procedures and absence of standardized computational frameworks.

The subjective nature of sensory evaluation is a key driver of non-reproducibility which includes color assessment, aroma profiling and liquor-quality scoring. Because trained tasters also exhibit inter-observer and intra-observer variance which influence by experience level, fatigue, environmental conditions and sample presentation formats.

The environmental variability also contributes to poor reproducibility which specially differences in altitude, soil chemistry, humidity and microclimatic influences across tea growing regions. These environmental attributes influence leaf morphology and chemical composition. But tea region classification is not fully captured in traditional evaluation frameworks.

The lack of digitally standardized datasets poses a major limitation for computational classification. Many tea institutes are not maintained large-scale, high-resolution classification or spectral datasets with controlled acquisition settings. Without proper standardized datasets with computer vision and machine learning applications are not appropriate to trained, validated or benchmarked.

1.4.2. No deep learning-based liquor-color system

Color of a brewed tea liquor is a fundamental indicator which is used by trained tasters for tea region classification. It can differentiate the growing regions, assess quality and determine the market value. The tea liquor color has reflected the chemical composition of tea which are including polyphenols, theaflavins, thearubigins and mineral content.

Currently color-based region classification performed through manual sensory evaluation and industry lacks deep learning-based liquor region classification system which resulting a limited technological advancement compared to other agricultural sectors. Deep learning models have demonstrated high accuracy in numerous agricultural classification tasks which are fruit ripeness detection, leaf disease identification and soil quality prediction due to their ability to extract hierarchical and non-linear features from images. Sri Lankan tea industry has not adapted to that kind of tea region classification system based on tea liquor color due to absence of controlled imaging protocols and lack of integrated frameworks. Without such automated systems, tea region classification continues depending on traditional sensory evaluation which also creates inconsistencies across brokers and auction centers.

The absence of a deep learning–based liquor-color system also limits progress toward real-time, automated classification pipelines. It requires standardized image acquisition setup which includes controlled lighting, camera image acquisition setup, including controlled lighting, camera sensors, image angles and sample preparation based on ISO 3103 brewing standards.

Deep learning architecture such as Convolutional Neural Networks (CNNs) have potentially higher accuracy which can identify the subtle variations in liquor color that are difficult for human eye to consistently distinguish. These models can integrate color histograms, texture features, chemical correlations and region-specific metadata to create highly accurate classification systems. But that kind of systems require extensive labeled datasets that capture intra-region variability and seasonal changes across Sri Lankan tea growing regions.

1.4.3. Need for standardized image environment

The accurate tea region classification depends on dry leaf appearance, infused leaf characteristics and liquor color which heavily depends on consistency of image acquisition conditions. Sri Lankan tea industry lacks a standardized image capturing environment. This leads to variations that significantly weaken the reliability and reproducibility of computer vision and deep learning system used for tea region classification.

A standardized image environment is an environment that controlled and repeatable conditions under which tea samples are photographed, including illumination, imaging angles, sensor calibration, background color, distance to subject, and environmental temperature and humidity. Inconsistent lighting is the most critical source of error and liquor exhibits high sensitivity to color distortion. The need of standardized imaging environment becomes more important for deep learning-based classification approaches. If a sample image varies in lighting, shadow or color temperature, the resulting datasets become noisier.

1.5. Aims & Objectives

1.5.1. Main objective

The primary objective of the research is to develop a robust, reproducible framework for tea region classification in Sri Lanka with using standard imaging environments and deep learning methodologies. Sri Lankan tea industry heavily relies on regional differentiation, and it is key indicator of quality, market identity and price determination.

The research aims to integrate computer vision, deep learning architecture and standardized color-based evaluation protocols. It produces automated classification models that capable of identifying the seven major tea growing regions in Sri Lanka. By leveraging high-resolution imaging under controlled environmental conditions use to generate a scalable and reproducible dataset that captures region-specific characteristics such as liquor color features. It provides actionable insights that support the Sri Lankan tea industry's transition from manual machine-learning outputs with domain knowledge from tea tasters and existing region classification standards.

Figure 8: 1.5.1 Tea Regions of Sri Lanka [12]

1.5.2. Sub-objectives

A one of sub-objective in the research is building a large, labeled, multi modal dataset across seven tea regions of Sri Lanka. That dataset included before brew, after brew and waste images of teas. The target was to collect 700 images from all regions and 100 images per region.

Another sub-objective of the research is to create web applications for centralized classification, management and visualization. This application secures the user authentication based on their role and provides batch classification, visualization of model, confidence and provenance, data management such as download .csv, download .png and download .pdf files. Classification latency is less than two seconds per image and there are many options for region classification. Those are single image classification, multi-image classification and raw image and preprocessed image wise classification and get the cropped preprocessed images from the raw images.

Another sub objective of the research is to create android app that can get real time liquor images and make the region classification of teas. It makes very robust and efficient way to classify the tea regions and expert can also use it. It takes very little time to classify the tea samples, and anyone can easily use that app with smart phone.

1.6. Scope and Limitations

The scope of this research is to develop and evaluate deep learning-based tea region classification systems using digitized images of tea liquor samples. The research emphasizes the extraction of color and texture-based features from brewed liquor from different regions of Sri Lanka. The system also supports real time classification and gives the confidence of a sample being in a particular region.

The primary scope of the research is to design a standardized imaging protocol for capturing tea liquor samples. The study is also bound by several limitations. A major limitation concerns light sensitivity. Those image base models show performance degradation under inconsistent illumination conditions. Derivation from controlled lighting environments may affect the output predictions.

The study also limited to tea liquor samples only, excluding assessments of tea leaves, infused leaves or dry tea particles. The liquor characteristics vary depending on season, altitude, soil conditions and processing variations. Then the model accuracy is limited. The device that is used to collect images is also a limitation. Because when changing the device, it also affects the quality of images.

1.7. Significance of Research

The classification of tea by regional origin is a critical factor which influences the market value. In Sri Lanka, there are seven geographical tea regions such as Nuwara Eliya, Kandy, Ruhuna, Dimbula, Udapussellawa, Sabaragamuwa and Uva. Those are heavily relying on expert graders who evaluate tea liquor color and flavor profiles to determine the region authenticity and the quality. This process is highly subjective and depends on expert experience and environment.

This research supports the industrial need for scalable and cost-efficient quality control systems. AI based liquor color analysis offers efficient and robust ways to reduce the time of tea tasters and they can verify their classification is correct or not. The ability to integrate the classification model into web applications and Android app broadens adaptation, enabling estates, factories and quality assurance labs to perform rapid evaluations without specialized tools and equipment. This improves decision making, efficiency and reduces the operational costs.

This study improves the traceability and the authenticity verification in the supply chain. Because mislabeling of tea region origin affects both export value and consumer trust. The liquor color-based classification models offer verification capabilities, ensure the batches meet regional standards before packaging. This effect to the national branding initiatives such as 'Ceylon Tea' and their origin integrity is vital for maintaining premium market positioning.

1.8. Novelty of Research

This research introduces several innovative contributions to the domain of tea region classification exclusively based on tea liquor color using deep learning techniques. Prior studies have explored chemical profiling, sensory evaluation or leaf color-based classification. Traditionally, liquor color-based classification has done with trained tasters. This study fills the gap by integrating Convolutional Neural network (CNN) based tea region classification framework using tea liquor color images.

The tea liquor color is a critical indicator of polyphenol content, fermentation levels and environmental factors linked to specific tea growing regions. By training CNN models with tea liquor color images can be extracted the deep spectral patterns that are not easily identifiable through manual observation. This study introduces a controlled imaging environment and device normalization approach that enhances the robust of deep learning-based classification. The color calibration techniques that can enable the CNN models to generalize across different capture environments.

Another novel contribution is the integration of Convolutional Neural Network-based prediction into deployable applications. In this approach includes a web platform and android mobile application for cost-effective region classification purposes. This deployment pipeline allows tea factories, tea estates and quality control laboratories to perform real time region classification. This can verify the sensory evaluation of tea tasters and reduce their time and human error when classification of tea regions based on tea liquor color. Overall, this study lies in its comprehensive integration of standardized brewing, deep learning, controlled imaging, and deployable classification systems.

1.9. Chapter Summary

The introduction chapter presented a comprehensive foundation for developing a tea region classification framework based on tea liquor color. It began by outlining the global tea industry, its economic value and necessity of objective classification system which surpass subjective human sensory evaluation. The key issues identified in current industry practices rely on lack of reproducible classification methods which absence of deep learning-based color system and non-standardized imaging environment.

The chapter further established the main objective to design the tea region classification system. Several sub objectives also have defined which include development of image dataset, feature extraction mechanism, web base interface and android app for the real time classification.

2. Chapter 02 - Literature Review

2.1. Chapter Introduction

Tea region classification has emerged as a critical research area due to the increasing demand for objective, rapid, and reproducible tea quality assessment methods. Traditional sensory evaluation has widely practiced in tea industry which often lacks consistency, suffers from evaluator bias and fails to scale with modern industrial requirements. The tea characteristics have been explored in this chapter. Those are liquor color, aroma, taste and chemical composition vary across geographical regions, numerous analytical, machine learning and deep learning approaches.

This chapter provides a comprehensive review of existing literature which is relevant to tea region classification. It begins with scientific foundational background of tea. It also covers chemistry, processing methods and sensory attributes. This chapter addresses analytical techniques such as electronic noses, electric tongues, chromatographic techniques and hyperspectral imaging. The chapter examined machine learning and deep learning. Then the chapter examines the tea liquor color-based evaluation and the previous attempts regarding the tea region classification. It has used classical chemometrics and modern neural networks.
Finally, significant research gaps are identified in standard visual evaluation environment. There is limited use of deep learning for liquor color-based tea region classification and challenges in dataset reproducibility which collectively motivate the present study.

2.2. Tea Science & Chemistry

Tea (Camellia sinensis) is a chemically complex beverage that is composed of thousands of bioactive chemical compounds. It has influenced flavor, aroma, liquor color and overall quality. The factors such as agro-climatic conditions, soil properties, altitude, plucking standard and post-harvest processing have shaped the chemical composition of tea. Its biochemical variations have resulted from environmental factors impart distinct color characteristics for the teas from different tea regions.

2.2.1. Major Chemical Groups in Tea

Tea contains major categories of chemical compounds. Those are polyphenols, alkaloids, volatile aroma compounds, amino acids, pigments and minerals. By those categories, polyphenols play a central role in tea region classification. The research has found during the fermentation level, catechins undergo enzymatic reactions forming theaflavins (TFs) and thearubigins (TRs) [22].

Table 8: 2.2.1 Major Chemical Components in Tea and Their Functional Roles

Chemical Group	**Key Compounds**	**Functional Role**	**Influence on Liquor Color**
Polyphenols	Catechins, EGCG	Antioxidants, precursor for color compounds	Forms TFs & TRs during oxidation
Pigments	Chlorophylls, carotenoids	Leaf color, aroma precursors	Limited; broken down during fermentation
Oxidation Products	Theaflavins, thearubigins	Taste (briskness/body), color intensity	Directly responsible for orange–red color
Amino Acids	L-theanine	Umami, sweetness	Minimal color effect
Caffeine & Alkaloids	Caffeine, theobromine	Stimulant effects	No color effect
Aromatic Compounds	Linalool, geraniol, aldehydes	Tea aroma profile	Minimal color effect
Minerals	Ca, Mg, K	Water hardness interactions	Can alter infusion clarity

2.2.2. Chemistry of Black Tea Formation

The Black tea production involves several factors. Those are withering, oxidation and drying. Oxidation is the most critical chemical transformation. In oxidation, catechins react with polyphenol oxidase and they form theaflavins, thearubigins and theabrownins. These chemical factors affect the tea liquor color variations. As an explanation, theaflavins contribute to brightness and golden yellow hues, thearubigins produce reddish-brown liquor and theabrownins contribute for higher molecular weight compounds associated with darker liquor [23].

Table 9: 2.2.2 Influence of Chemical Profiles on Regional Liquor Color

Tea Region (Sri Lanka)	**Typical TF:TR Ratio**	**Expected Liquor Color**	**Chemical Interpretation**
Nuwara Eliya	High TF, lower TR	Light golden	Slow leaf metabolism due to high altitude
Uva	Moderately high TF	Bright coppery	Windy, cool climate enhances TF formation
Kandy	Balanced TF/TR	Amber	Mid-altitude oxidation balance
Dimbula	Moderate TF, rising TR	Bright red	Seasonal climate variability
Ruhuna	Low TF, high TR	Dark red brown	High temperature rapid catechin oxidation
Sabaragamuwa	Lower TF, higher TR	Deep reddish	Low altitude, humid conditions
Udapussellawa	Higher TF	Pale golden	Cool climate slows TR formation

2.3. Origin, types, and processing of tea

Tea is one of world’s oldest and most widely consumed beverages. It has long history which trace back more than 4000 years to ancient China. It was initially used for medical purposes before becoming an everyday consumed drink. According to the historical evidence tea has spread from China to Japan through Buddhist monks. After that it that spread to Europe via Portuguese and Dutch traders. Finally, it has spread to the Asia through British colonial agriculture expansion. Sri Lanka emerged as one of the most globally recognized tea producers in 19th century.

The tea industry is categorized by a diverse range of tea types. It has produced specific processing techniques which influence chemical composition, sensory qualities and liquor color. Each type of tea originated from camellia sinensis. But they differ primarily by the oxidation and nature of post-harvest processing. It is directly affected by the biochemical pathway of polyphenols which specially are on catechins.

2.4. Tea processing steps

Tea processing methods significantly influence the tea liquor color, aroma and tea region specific characteristics. These steps vary depending on tea type. But those generally follow a sequence of withering, rolling, oxidation/fermentation, drying and grading.

2.4.1. Withering

This step is conducted for the removing of moisture and soften of tea leaf's structure. This has affected to the enzymatic activity which influencing the catechin concentration and the subsequent formation of theaflavins and thearubigins. This stage is impacted by the climatic factors and cause it for regional variations in chemical profiles.

Figure 9: 2.4.1 Withering of tea leaves [24]

2.4.2. Rolling or Maceration

In this step mechanical rollers rupture the leaf cell structures. It enabled oxidative enzymes to interact with catechins. The intensity of rolling directly affects the liquor body and strength and increases the substrate availability for oxidation.

Figure 10: 2.4.2 Rolling of tea leaves [25]

2.4.3. Fermentation

Fermentation is also known as oxidation. It is the most critical stage for black tea. Polyphenol oxidase catalyzes catechin oxidation into theaflavins and thearubigins. It determines the liquor's color, brightness and depth. The oxidation depends on various conditions. Those are temperature, humidity, and oxidation duration across tea regions.

Figure 11: 2.4.3 Fermentation of tea leaves [25]

2.4.4. Drying

The tea is put through the drying process after the fermentation process. It takes 20 minutes to complete this process. It halts oxidation and stabilizes the chemical compounds. It has ensured the aroma development and the moisture reduction for shelf stability through high temperature drying. The tea has dried in hot air at 80-90 Celsius which gives the dark brown, black color of tea particles.

Figure 12: 2.4.4 Drying of tea [25]

2.5. Sensory components of tea

The sensory attributed of tea plays a critical role for determining tea quality, origin and consumer acceptance. These attributes are categorized into appearance, aroma, flavor, liquor color and mouthfeel. Traditionally sensory evaluation of tea has been done by highly trained expert tea tasters. In the present, due to the improvement and performance of computational intelligence these sensory attributes can be cached through the liquor color of the tea.

2.5.1. Aroma of tea

Aroma is one of the most influential sensory attributes of tea. It is a critical chemical composition for processing the quality and the geographical region. Recent studies have found that aroma formation mechanisms are closely linked to biochemical pathways. It also indirectly influences the liquor color and enables improving tea region classification.

Tea aroma is primarily shaped by the nature of plants. It influences biochemical formation and releases volatile compounds through multi environmental factors. The research identifies that six major tea type. Those are green, white, yellow, oolong, black and dark tea. Those are produced from same cultivar under controlled conditions which exhibit distinct volatile profiles due to the different processing methods. Techniques like gas chromatography-mass spectrometry (GC/MS) analysis have revealed 168 volatile compounds. The black tea has shown highest concentration of 710 µg/g, while green tea has the lowest concentration of 20 µg/g [26].

The study further establishes that aroma molecules in tea have formed from four precursor groups. Those are caratenoids, fatty acids, glycosides and amino acids. These biochemical precursors differ across regions due to the terroir, soil nutrition, climate, altitude and cultivator genetics. Tea liquor color also inherently associated with the same precursor pathways which are responsible for regional aroma profiles. These chemical relationships can be improved classification robustness especially in automated AI based systems.

2.5.2. Taste of tea

Tea taste is one of the most influential sensory attributes of tea. It can be distinguished between tea types, origins and processing styles. For the tea region classification through tea liquor color, taste profiles provide complementary information for the color-based evaluation. The taste attributes are determined by chemical composition which turn into shaped by regional agro-climatic factors, cultivator variations and manufacturing process.

Recent research has emphasized that taste of green tea is influenced by hundreds of non-volatile compounds which include catechins, polyphenols, amino acids, sugars and organic acids. These factors collectively define the bitterness, astringency, sweetness, umami and smoothness in liquor.

Recent research has integrated the unsupervised and supervised machine learning methods to categorize and predict tea tastes more correctly. The study analyzed 88 Chinese green tea infusions and applied Quantitative Descriptive Analysis (QDA) with using hierarchical K-mean clustering approach [27].

The research also identified the taste types with higher TP or catechins concentrations tended to display higher bitterness and astringency. These findings are supported to integrate chemical taste attributes in computational models for tea region classification. Due to the sensory challenges in tea region classification in taste profiles, tea liquor color is the best way to capture the feature of region-specific tea liquors more correctly.

Table 10: 2.5.2 Primary Chemical Contributors to Tea Taste and Their Sensory Functions

Chemical Group	**Representative Compounds**	**Taste Attributes**	**Impact on Liquor Color**
Tea Polyphenols (TP)	EGCG, EGC	Bitterness, Astringency	Darker yellow to green due to oxidation reactions
Catechins	ECG, EC	Bitterness, Astringency	Correlated with deeper green hues in green tea liquor
Amino Acids	L-Theanine, Glutamic acid	Umami, Sweetness	Brighter and clearer liquor appearance
Alkaloids	Caffeine	Bitterness	Minor contribution; increases brightness
Soluble Sugars	Sucrose, Glucose	Sweetness	Lightens liquor tone
Organic Acids	Malic, Citric acids	Sourness, Freshness	Enhance clarity and transparency

2.5.3. Color of tea

Color is one of most influential organoleptic attributes of tea. It plays a critical role in consumer acceptance, quality grading and region classification. It also provides essential cues regarding the chemical composition, processing intensity and leaf maturity. In China both herbal teas and green teas are widely consumed.

Tea contains high levels of catechins, flavonoids, pigments and oxidation which influence liquor color under standardized brewing conditions. Among the varies teas, green tea undergoes minimal enzymatic oxidation. It gives yellow green to bright green hues due to the high catechin content and lower level of theaflavins and thearubigins [28].

Matcha is a green tea type. A study has reported that rising global interest in matcha is linked to the consumers growing which focuses on health and wellness. It is recognized for its antioxidant composition. The researchers specially examined how the levels of shading (0%, 50%, 70%, 90%) influenced the physical, chemical and sensory qualities of matcha. These findings refer that increasing the intensity produced matcha with darker color values, reducing the moisture content (approximately 9.26% to 8.64%) and solubility (from 17.10% to 15.37% in 0% versus 90% shading) [29]. This gives an idea that color and shading can make a significance change across tea regions. Then tea region classification is applicable to do with tea liquor color.

2.6. Chemical composition of tea

The chemical composition of tea is critical determinant of its sensory profile, functional properties. It is potential for tea region classification. Tea leaves contain a complex mixture of biochemical compounds. Those are polyphenols, amino acids, alkaloids, volatile compounds, vitamins and minerals. Those significantly vary with cultivar type, agro-ecological conditions and post-harvest processing. Under those polyphenols catechin has the highest contribution for bitterness, astringency and antioxidant capacity. Caffeine is another important compound that directly influences the stimulant properties of tea. It forms a biochemical signature often used in tea region classification.

The chemical composition of tea is highly sensitive to environmental variables. Those are altitude, soil chemistry, shading practices, rainfall and temperature. These factors have influenced the biosynthesis of phenolics, secondary metabolites and pigments. For example, teas from high elevation regions have higher concentration of catechin and flavor precursors due to slower leaf growth. It has also increased exposure to stress-related metabolic pathways.

Among different tea varieties in China, Oolong tea (OT) represents a semi-fermented category. Research has demonstrated that OT undergoes distinctive biochemical transformation during the processing. It has resulted the unique chemical profile. The processing workflow of OT includes withering, bruising, partial fermentation, fixation, rolling and drying. These stages have influenced the concentration of catechin, theaflavins, thearubigins, flavanols and volatile aromatic compounds. The research indicates that the phytochemistry of OT has dominated by polyphenols, especially catechin, such as epigallocatechin-3-gallate (EGCG), flavonoid glycosides, alkaloids, amino acids and various pigments from oxidative polymerization. Overall, polyphenols and pigments play a central role in determining liquor color. It has marked the chemical composition analysis as a vital foundation for the deep learning-based tea region classification systems [30].

2.7. Chemical reactions in fermentation

Fermentation is a critical biochemical stage in orthodox and semi-fermented tea production systems. It has directly influenced the pigment profile, aroma compounds and liquor color which discriminate features for tea region classification. It is primarily an enzyme-driven oxidation process. Research has found that microorganisms from Protista, Fungi and Bacteria kingdoms participate in general fermentation process. It has produced enzymes, pigments and macromolecules which indirectly affect biochemical transformations.

2.7.1. Metabolic Principles Relevant to Tea Fermentation

Research shows that microbial metabolism consists of anabolic and catabolic pathways. Catabolic-like oxidation generates theaflavins (TFs) which is a primary pigment of determining the liquor color of tea. Anabolic-like enzyme actions facilitate the transformation of catechin into complex polymeric structures. The scientific structures have shown that temperature and PH directly impact to the metabolic activity in microorganism. It has given the optimal activity around 20-30 Celsius which consists of mesophilic biological systems.

2.7.2. Fermentation Kinetics and Tea Color Development

The microbial growth in fermentation follows lag, log, stationary and dead phases. The research has used these kinetics models to analogize the rate of biological oxidation during tea processing. These controlled kinetics are essential for developing reproducible liquor color for critical AI based region classification [31] .

Table 11: 2.7.2 Fermentation Kinetics

Kinetic Phase (Biological Model)	**Parallel in Tea Fermentation (Research Mapping)**
Lag Phase	Leaf enzymes equilibrate with oxygen; moisture redistributes.
Log Phase	Rapid oxidation of catechins → formation of TFs and TRs; major color development.
Stationary Phase	Oxidation slows; pigment stabilization occurs.
Death Phase	Over-fermentation leads to pigment degradation, darkening, and quality loss.

2.8. Traditional Tea Evaluation

Traditional tea evaluation is also known as organoleptic or sensory evaluation. It is done by highly experienced tea tasters based on their human sensory judgements. Those sensory parameters are liquor color, aroma, taste, briskness, strength, appearance and mouthfeel. But those methods suffer from subjectivity, environmental variability and assessor depending on inconsistency. Then, there is a need for standardized, objective and automated classification approaches for tea region classification. There are three types of sensory evaluation. Those are Dry leaf evaluation, Liquor taste evaluation and liquor color evaluation.

2.8.1. Dry tea leaf-based evaluation

Dry tea leaf evaluation is a fundamental stage of tea quality control process. It plays a significant role in understanding the biochemical potentials of tea samples before the processing. Research has found that the chemical composition of postharvest tea leaves directly influences the subsequent formation of liquor color, aroma volatiles and polyphenolic oxidation pathways which are essential to determine tea region classification. Recent studies have found chemical markers like tea polyphenols, amino acids and the polyphenol to amino acid ratio is a reliable predictor compared to the human evaluation alone [32]. It has been found that dry leaf composition varies significantly across cultivars, plucking standards and agro-ecological regions. Therefore, dry tea leaves chemical composition significantly varies among tea regions. Then the deep learning-based color classification system seeks to capture this variation, and it gives better accuracy without considering the chemical composition.

2.8.2. Liquor taste and aroma-based evaluation

Tea taste is a fundamental dimension of tea quality control. It also checks along with aroma and mouthfeel. Research has found that both aroma compounds and non-volatile taste compounds can be used for tea flavor evaluation [33]. It summarizes 184 known odorants which contribute to the aroma and off flavor in teas. It has highlighted those modern analytical methods like extraction/enrichment, GC-MS, sensory panels and chemometric approaches widely used in flavor evaluation.

Another recent study has identified several varieties of black tea with different aroma characteristics like floral and sweet. It has used the combination of GC-E-Nose, GC-IMS and Odor Activity Value (OAV) analysis. The study has identified 38 volatile compounds. Out of

38, 15 odorants had OAV>1 which means they are likely to significantly contribute to perceived aroma. Among those compounds, linalool, nonanal, benzaldehyde, octanal were deemed important.

2.8.3. Liquor color-based evaluation

Liquor color of tea is one of most critical attributes which use in tea quality control process. It has provided rapid, non-destructive and objective information about the biochemical composition of tea. Because tea liquor color is strongly influenced by the terroir, processing techniques and chemical constituents. Those are theaflavins (TFs), thearubigins (TRs) and polyphenol oxidate activity (PPO).

Research has found that black tea undergoes substantial chemical transformation during fermentation. It is primarily driven by polyphenol oxidase and peroxidase which are the enzymes that catalyze the oxidation of polyphenols [34]. It also reports that more than 70% of the chromatic characteristics of black tea liquor have resulted from the balance between pigment fractions of theaflavins (TFs), Thearubigins and theabrownins. Considering those pigments, TFs impact for golden yellow brightness, TRs contribute for crimson to ruby-red hue and TBs contribute for the dark brown tones. These pigments have directly influenced both color and taste attributes such as brightness, strength, sweetness, astringency and freshness.

2.8.4. Limitations of human sensory testing

Human sensory testing has been done by trained tea tasters in industry. They used sensory attributes such as flavor, aroma, appearance, liquor color and mouthfeel for their classification and quality control of the tea. But the sensory analysis is inherently subjective, variable and limited to support large scale classifications.

Human sensory perceptions have been influenced by numerous physiological, psychological and environmental factors. Those affect to the result inconsistencies of both within the same evaluator overtime and between different evaluators. It also depends on the environmental factors such as lighting conditions, angle of observation, surface color of the teacup and ambient environmental reflections.

Another limitation of human sensory testing is evaluation lacks quantifiable precision. As an example, liquor color terms such as ‘bright’, ‘coppery’, reddish’, ‘brisk’, ‘golden ring presence’ do not correspond to the measurable values. The modern studies emphasize the need of digital colorimetry which use the computer vision and spectral imaging technologies to replace qualitative descriptors with quantifiable parameters. Those are RGB values, CIELAB, Hue, saturation, chroma and absorbance.

Another limitation of human sensory testing is limited scalability for large scale industry operations. It is time-consuming and labour intensive, which makes it not suitable for real time or large-scale region classification. The tea auctions and processing centers must evaluate thousands of samples daily. Therefore, manual sensory testing-based evaluation is impractical. Then there exists a clear need for the automated computer vision, deep learning-based region classification system.

2.9. Analytical Techniques

Analytical techniques play a crucial role in modern tea science. It is essential for tea region classification, where chemical, sensory and image-based parameters must be accurately quantified to differentiate regional origins. Traditional sensory evaluation cannot ensure subjective, reproducible and high-resolution classification. Therefore, scientific techniques like chromatographic techniques, spectrometric techniques and computer vision have widely adapted.

2.9.1. Chromatographic Techniques

Chromatography is one of most commonly used approaches to the categorize chemical composition of tea. A High-Performance Liquid Chromatography (HPLC) has applied to quantify catechins, theaflavins, thearubigins, caffeine, amino acids and volatile compounds. A recent study has shown that specific chemical markers like epigallocatechin gallate (EGCG) concentration in green and oolong teas differ across regions [35]. A technique such as GC-MS is effective for profiling volatile aroma compounds. Those are terpenoids, aldehydes and phenolics.

Table 12: 2.9.1 Chromatographic Techniques Used in Tea Classification

Technique	Target Compounds	Application in Region Classification
HPLC	Catechins, caffeine, TFs, TRs	Identifies regional chemical signatures
GC–MS	Aromatic volatiles	Aroma-based discrimination
LC–MS/MS	Polyphenols & flavonoids	High-precision metabolic fingerprinting
GC–FID	Hydrocarbons & aldehydes	Aroma pattern grouping

2.9.2. Spectroscopic Techniques

Spectroscopic techniques have been widely used for enabling rapid, non-destructive and high-throughput evaluation. It can be used to quantify liquor color parameters because TFs, TRs and TBs correspond to characteristic wavelength.

Techniques like Near-Infrared Spectroscopy (NIR) and Fourier Transform Infrared Spectroscopy (FTIR) have allowed analysis of organic functional groups. Those have emerged as accurate tools for discriminating tea based on their biochemical compositions. There are Machine learning models which trained on NIR/FTIR spectra. Those have achieved accuracy levels exceeding 90% in regional differentiation.

2.9.3. Electronic Nose (E-Nose) and Electronic Tongue (E-Tongue)

Research has found that Electric Nose (E-Nose) systems provide a non-destructive and highly sensitive method for evaluating tea aroma profiles. It primarily consists of two integrated subsystems. Those are sensing systems and automated pattern-recognition systems. The sensing unit often includes an array of gas sensors. It responded to the complex VOC mixtures by producing distinct electrical signatures [36]. In tea research the most extensively utilized sensors are metal oxide semiconductor gas sensors. It is very stability, low cost, highly sensitive and commercially available. It demonstrates that E-Nose technology has potential for fermentation monitoring, quality assessment and aroma profiling.

Electronic tongue (E-Tongue) systems have emerged as powerful analytical instruments for assessment of tea quality. It has overcome the key limitation of conventional human sensory evaluation. Recent research has shown that electronic tongue systems are coupled with Artificial Neural Network. In the study there were four grades of green tea which analyzed using custom E-Tongue system with multi sensor array. It has used modeling techniques like KNN and ANN using cross validated parameter optimization. The ANN model has achieved 100% identification accuracy for both training and validation datasets. It has demonstrated the exceptional capacity in tea grade level classification [37].

2.10. Computer Vision in Tea Science

2.10.1. RGB, HSV color models

Color is one of the most informative visual cues used in tea quality control. Digital imaging-based tea classification systems frequently adopt RGB and HSV color models due to simplicity, efficiency and strong correlation with chemical variation of tea. The RGB color model represents the images using three additive primary colors which are RED (R), Green (G) and Blue (B). In tea studies, R-values typically indicate stronger oxidation in low grown black tea and G-values associated with high grown teas that remain more catechin. But RGB values are sensitive to camera sensor variability. It can be reproduced across experiments. Therefore, standardized lightning environment is essential for tea region classification.

Researchers have converted the images into HSV (Hue, Saturation, Value) color space to overcome the issues of RGB color space. It is closer to human color interpretation and separate chromatic information. Hue is particularly important tea evaluation due to it accurately captures chromatic shifts associated with regional characteristics. Saturation reflects the purity of liquor color, and it has been linked with flavonoid concentration and fermentation intensity. Value represents brightness, and it is influenced by transparency and particle content in tea infusion. Research has found for tea bud identification technology based on HIS/HSV color transformation. It has been reported that RGB color model is widely used in general computer vision tasks. It has discriminated for isolating young tea buds due to complex background of tea plantation [38].

2.10.2. Hyperspectral imaging

Hyperspectral imaging (HSI) has integrated two-dimensional spatial imaging with high resolution spectral analysis. It allows simultaneous capture of both spatial and chemical composition information for each pixel in a sample. HIS captures the reflection and absorbance across many narrow wavelength bands. It allows detection of chemical compounds like polyphenols, pigments, moisture and amino acids. It has influenced the color, taste and liquor chemistry [39]. But these techniques require very complex datasets and need people with domain knowledge.

2.10.3. Image correction & calibration techniques

Accurate color-based tea evaluation is very important for tea region classification. The raw images are often affected by various illuminations, camera sensor characteristics, geometric distortion and environmental conditions. These techniques ensure the extraction of color features like RGB, HSV or spectral reflectance values [40].

Variations of ambient lighting has introduced the large uncertainties into tea color analysis. It has directly impacted on the performance of machine learning models in tea region classification.

Typically, correction techniques include:

1. Gray-world normalization - which assumes that the average reflectance of a scene is gray and adjusts RGB gains accordingly.
2. White balance calibration - where a reference white surface is used to compensate for color temperature variation.
3. Shading correction or homomorphic filtering - which separates illumination from reflectance components to reduce shadows and glare.

Camera sensors vary in spectral sensitivity. Because RGB values unreliable across different cameras or acquisition environments. The calibration techniques ensure the consistency between datasets collected across different sessions.

1. Color reference charts (e.g., X-Rite Color Checker) – It is used to map camera-dependent RGB values to device-independent color spaces such as CIE-Lab or sRGB.
2. Polynomial regression color calibration - It is used reference patches to fit a transformation matrix from raw RGB to calibrated color values.
3. Spectral camera calibration – It is used in multispectral or hyperspectral imaging to correct for sensor non-linearity and wavelength drift.

2.10.4. Standardized imaging environments (ISO standards)

Standardized imaging environments are critical for ensuring the reproducibility and comparability in liquor color-based tea region classification. The tea samples have introduced various inconsistencies in illumination, background, vessel geometry and sample preparation.

The International Organization for Standardization (ISO) has provided a controlled framework for preparing tea for testing purposes.

One of the most relevant standards for preparing tea is ISO 3103. It has specified the procedures for brewing tea (quantity of leaf, water temperature, infusion time, pot geometry) to ensure consistency in tea liquor appearance. That standard was originally designed for human sensory testing. The standard minimizes the variability caused by brewing conditions. It has enabled the computer vision system to attribute observed color differences to genuine regional chemical differences rather than preparation inconsistencies. The standard specifies the parameters such as:

1. Tea-to-water ratio: 2.0 g of tea per 100 mL of freshly boiling water.
2. Water temperature: 100°C using freshly boiled, non-distilled water.
3. Steeping duration: 6 minutes precisely.
4. Brewing vessel: White porcelain pot with tight-fitting lid (150 mL or 300 mL).
5. Liquor separation: After brewing, the infusion must be fully drained into a matching white porcelain bowl for evaluation.

It has reduced the color variation caused by inconsistent brewing and provides a uniform reference background for vision algorithms. It also ensures that differences in measured RGB/HSV/L* a* b* values correspond to regional tea variations which not for environmental artifacts [41].

2.11. Machine Learning in the Tea Industry

2.11.1. Feature Engineering

Feature engineering plays a critical role in the accuracy and robustness of tea region classification which relying on liquor color as primary discriminative parameter. It is influenced by biochemical compositions such as catechins, theaflavins, thearubigins, mineral content and polyphenol oxidase activity. Therefore, transforming raw image data into meaningful discriminative features is essential for building reliable machine learning and deep learning model approaches. Traditional color models like RGB, HSV/HIS and CIELAB have enabled the quantification of hue, chroma and brightness. It is sensitive to subtle variations in oxidation and fusion chemistry. Commonly extracted features include:

1. Mean pixel intensity of each color channel

2. Color histograms (global or region-specific)
3. Chromaticity coordinates (e.g., a*, b* in CIELAB)
4. Saturation and lightness distributions

Table 13: 2.11.1 Common Feature Types Used in Tea Liquor Classification

Feature Type	Examples	Purpose
Color Features	RGB means, HSV histograms, CIELAB a*, b*	Capture chromatic differences among regions
Texture Features	GLCM, LBP, Gabor filters	Encode spatial patterns and liquor clarity
Statistical Features	Variance, entropy, skewness	Quantify distribution-based variability
Deep Features	CNN embeddings, feature maps	Learn high-level representations automatically
Fusion Features	Concatenated vectors, PCA-compressed features	Improve accuracy & robustness

2.11.2. Classical ML models

Machine Learning (ML) has played a crucial role in automating the tea region classification. Those models are Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors and Logistic Regression.

SVM has been one of the most frequently applied methods in agricultural imaging tasks. SVM has demonstrated high accuracy when distinguishing sublet variations in tea leaf morphology and tea liquor color. Researchers have combined SVM with hand-crafted feature descriptors such as histogram of oriented gradients, gray level Co-occurrence matrix and color moment features.

Another classical model used in this domain is K-Nearest Neighbors (KNN). It is very useful when regional variations gradually increase rather than distinct. It is also sensitive to noise and feature scaling. It has performed competitively when combined with robust preprocessing and normalization methods.

Another classical model used in this domain is Random Forest. It has widely applied due to their ability to handle heterogeneous features and provide meaningful feature importance rankings. It has mitigated the overfitting and improved predictive stability and making it suitable for datasets to collect under varying environmental conditions. Those are differences in lightning, leaf maturity or camera hardware. It has been used to integrate multi-modal features such as color indices, texture metrics and handcrafted morphological attributes which leading to improvements in classification robustness [42].

2.12. Deep Learning in the Tea Industry

Deep learning has transformed the global tea industry with providing advanced capabilities for automated region classification. Deep Learning network automatically learns hierarchical feature representations from raw data which enables the superior robustness against environmental variability, noise and non-linear interactions among chemical and visual tea attributes. DL has become a promising tool for precise, scalable and real time tea assessment. The recent studies demonstrate that Convolutional Neural Network (CNN) can effectively model tea leaf morphology, aroma related cues for tea quality monitoring [43].

2.12.1. Tea quality analysis

Research has found that quality of tea leaves directly affects the quality of finished tea. It plays vital role in automating liquor color analysis particularly for tasks involving color consistency, quality grading and correlation with biochemical parameters.

Models like CNN can learn fine-grained spatial chromatic patterns with the infusion of hue, saturation, brightness and turbidity. It has strong performance in predicting chemical constituents such as polyphenols, theaflavins and thearubigins which directly from RGB or HSV images of tea leaves [44].

2.12.2. Tea region classification attempts

Recent studies have demonstrated the trace-metal profiling of tea. It combined with multivariate pattern-recognition techniques which provide a powerful basis for distinguishing the geographical origin of tea. The study has used the inductively coupled plasma-atomic emission spectrometry (ICP-AES) and inductively coupled plasma-mass spectrometry (ICP-MS). Those techniques have shown that metal content in tea leaves, strongly influenced by soil composition, geological background and environmental conditions. These factors differ substantially across Asian and African tea producing regions. It has enabled reliable classification when appropriate chemometric models are applied.

Research has found that 17 trace elements which include Al, Ba, Cd, Co, Cr, Cu, Cs, Mg, Mn, Ni, Pd, Rb, Sr, Ti, V and Zn form a sufficiently discriminative chemical signature for analyzing tea provenance. In study there were 85 tea samples (36 Asian, 18 African, 24 Commercial blends, 7 unknown Origin). The techniques like CP-AES and ICP-MS were used to obtain comprehensive elemental fingerprints. These multi-element datasets were subsequently analyzed using advanced pattern recognition algorithms [45].

2.13. Summary of Literature Gaps

A comprehensive review of existing literature on tea chemistry, sensory evaluation, image-based analysis and machine learning driven region classification reveals several persistent research gaps which limit the development of robust, scalable and generalizable tea region classification systems.

Although ISO 3103 provides a foundation for standardized brewing, many studies have failed to implement it. In this research, controlled lightning environments, calibrated imaging systems and standardized container geometry and background conditions have applied. Color features are highly sensitive to illumination and imaging setup. The lack of unified standards introduces noise measurement which reduces the classification reliability and inhibits cross study comparability.

Another significant gap is limited use of Deep Learning for region classification. Deep learning has been applied to leaf quantity and disease detection. But tea region classification has primarily been used for classical ML models (RF, SVM, LR, KNN). There is lack of CNN-based region classification studies, transformer based spectral models and domain adaptation and transfer learning studies for cross region generalization. The deep learning models remain underutilized due to the absence of large, annotated datasets.

Table 14: 2.13 Comparison of approaches

Common Goals	Chemical and sensory based approach	Vision based Deep Learning Approach
Improve accuracy and consistency in tea quality evaluation.	Require lab equipment (LC-MS, GC-MS, HPLC, NIRS, etc.).	Image-based CNN model analyzing color (RGB, LAB) to assess quality
Identify key chemical or sensory markers (like catechins, caffeine, aroma).	Involve manual sensory evaluation (panels tasting and scoring tea).	Fully automated, non-destructive, and suitable for real-time use
Focus on effects of processing methods, especially fermentation and steeping.	Emphasize biochemical compound quantification.	Doesn't require expensive lab equipment or trained tasters

Table 15: 2.13 Suitability of the TEAQNET

Aspect	Other Approaches	TEAQNET
Equipment	Requires costly lab tools.	Just a camera & trained model.
Speed	Slow, lab dependent.	Real-time prediction.
Expert involvement	Needs trained panels.	Fully automated.
Scalability	Limited by lab & human effort.	Highly scalable.
Sensitivity	Limited sensitivity in aroma and E-nose based approaches.	Highly sensitive through wide range of color.
Focus	Chemical & sensory features.	Visual features (color changes due to chemical reactions).
Practicality	Visual features (color changes due to chemical reactions).	Visual features (color changes due to chemical reactions).

2.14. Chapter Summary

The Literature Review chapter synthesized the findings from multidisciplinary research related to tea chemistry. Those are sensory evaluation, analytical instruments, imaging methods, machine learning and deep learning. It began with overview of tea industry, including chemical composition, aroma and flavor compounds and regional biochemical variation.

Traditional sensory based evaluation methods were critically examined. It has revealed the inherent limitations such as sensory fatigue, variation among tasters and lack of reproducibility. It has included analytical techniques like electronic nose, electronic tongue, HPLC, GC-MS which demonstrated their utility in profiling tea components. These methods remain time-intensive and impractical for industry-scale rapid classification.

The review has explored the imaging-based approaches which presenting RGB, HSV and hyperspectral imaging standards (ex-: ISO 3103 for liquor preparation). These studies have highlighted progress in objective measurement. But also expose the absence of standardized color-based region classification workflows.

The chapter has addressed the machine learning and deep learning applications in the tea sector for leaf grading, quality detection, visual inspection and early region classification attempts using chemometrics. These works have demonstrated promising accuracy, gaps which remain regarding liquor color modeling, standardized imaging environments and deep learning-based classification at scale.

3. Chapter 03 - Research Methodology

3.1. Chapter Introduction

The research adopts systematic experimental design to develop robust tea region classification model based on tea liquor color. The methodology has followed a structured pipeline which consisting of image acquisition, preprocessing, feature learning using deep convolutional neural networks (CNN) and model evaluation.

3.2. Research Design Overview

3.2.1. Conceptual framework

The conceptual framework is grounded in tea liquor color containing geographically discriminative visual signatures. It is influenced by factors such as soil chemistry, climatic conditions and regional processing variations. The framework integrates four major components which are standard image capture, image preprocessing and calibration, deep feature extraction and classification and performance evaluation.

Table 16: 3.2.1 Conceptual framework

Component	Description
Standardized Image Capture	Tea liquor images are collected under controlled lighting and background conditions to ensure minimal variability.
Image Preprocessing & Calibration	Background elements, reflections, and noise are removed. Images are normalized to reduce illumination-based discrepancies and standardized to the input size required for CNN models.
Deep Feature Extraction (CNN-Based)	A convolutional neural network automatically learns hierarchical color-texture representations from the liquor samples, removing the need for manual feature engineering.
Classification & Performance Evaluation	The trained model predicts the regional origin of tea samples and performance is

	assessed using accuracy, F1-score, IOU precision–recall analysis, and confusion matrices.

This framework has ensured a reproducible, data driven approach that is capable of identifying subtle color differences around different tea regions which is difficult perceived by human evaluators.

3.2.2. Processing Pipeline

The processing pipeline of tea region classification system has been designed to ensure data quality, enhance feature consistency and enable robust model learning. Tea liquor color is highly sensitive to environmental variations. Those are lightning intensity, background interference, imaging hardware and sample preparation. The pipeline incorporates multiple stages of preprocessing and transformation before any deep learning model is trained. This section has presented a comprehensive overview of each stage which includes image acquisition, annotation, preprocessing, normalization, augmentation, feature extraction through Convolutional Neural Networks (CNN), model training and evaluation.

The accuracy of tea region classification based on tea liquor color fundamentally relies on collecting high-quality image datasets. The image collection stage follows standardized procedures to ensure reproducibility and reduce noise. The tea samples are prepared following ISO 3103 brewing guidelines to minimize the variability in infusion strength and steeping conditions. Each tea sample is poured into standardized white porcelain cupping bowls to maintain consistent background reflectance to isolate liquor color as primary discriminative parameter.

Figure 13: 3.2.2 Professional tea cupping [46]

Images have captured using controlled using a controlled imaging box which equipped with LED lighting (5000-5500 K) to reduce shadows, glare and chromatic biases. The images have captured at a fixed distance and angle using a consumer grade smartphone camera. This has ensured the method is accessible and reproducible. The optical parameters which include exposure, white balance, ISO level and shutter speed are fixed manually. It prevents automatic camera adjustments from distorting liquor color representation. Each sample is photographed multiple times to minimize the sample specific noise and to expand dataset diversity.

After collection of images of tea liquor color, those images are labeled automatically based on the tea region, tea type and the sample number. There are seven tea regions such as Dimbula, Nuwara Eliya, Sabaragamuwa, Ruhuna, Udapussellawa, Kandy and Uva which have belonged to tea types mainly like BOP, BOPF, OP in Sri Lanka. As a example for a particular sample which belongs to Dimbula region also of the type BOPF and sample id of 001 that is labeled as DI_BOPF_001.

The raw tea images often contain unwanted artifacts such as bowl edges, table surface, shadows or specular highlights which are caused by reflective surfaces. These artifacts have degraded model performance by introducing irrelevant cues that the CNN misinterpret as discriminative features. To address these issues, a structured background removal step has been implemented.

In the first step, images are converted into appropriate color space (RGB), and it enables the identification of high contrast boundaries between liquor surface and the background.

The dataset undergoes normalization to reduce the inter-image variation caused by the different imaging devices, lightning conditions or sensor responses. Therefore, tea images should be normalized to reduce these variations.

3.3. Dataset Development

3.3.1. Tea sample acquisition

The main stage of this research is Tea sample acquisition. For this study, a systematic and standardized sampling strategy was designed to collect representative black tea samples across Sri Lanka agro-climatic tea regions. These regions are intentionally recognized for producing teas with distinct organoleptic and physicochemical characteristics due to variations in elevation, soil type, rainfall patterns and microclimatic conditions. The acquisition process has aimed to achieve followings:

1. Obtain regionally authentic black-tea samples reflecting standard production conditions.
2. Include multiple tea grades - Broken Orange Pekoe Fannings (BOPF), Orange Pekoe (OP), and Broken Orange Pekoe (BOP) to ensure grade-level representation and reduce classification bias.
3. Maintain traceability, standardization and controlled conditions to support reproducibility for imaging and liquor-color analysis tasks.

Tea samples were produced directly from licensed factories and Sri Lankan Tea Board. Each estate has provided documentation that confirms the precise geographic origin, production batch details and tea type. For each of the seven tea regions mainly three tea grades (BOP, BOPF, OP) were collected. There were collected seven hundred images from all tea regions such as hundred images per tea region.

3.3.2. Brewing standard (ISO 3103)

ISO 3103 is internationally recognized for preparing tea infusions for analytical and sensory evaluation. It was established by International Organization for Standardization. It provides a controlled, repeatable and unbiased method for brewing tea samples. It has ensured that variations in tea liquor color, aroma and flavor arise from tea itself rather than inconsistencies in preparation. Tea liquor color is key predictive feature that adherence to ISO 3103. It is essential to guarantee data quality and comparability.

The primary objective of ISO 3103 is to establish a uniform brewing protocol for eliminating human induced variation in brewing water temperature, leaf water ratio, infusion time, vessel material & geometry and filtration and presentation of liquor.

Table 17: 3.3.2 ISO 3103 standards for tea infusion

Parameter	**ISO 3103 Specification**	**Relevance to Liquor Color Classification**
Water Temperature	100 °C (boiling)	Ensures consistent extraction of pigments (theaflavins, thearubigins).
Leaf Mass	2.0 g	Controls concentration, ensuring uniform chromatic intensity.
Water Volume	100 ml	Standardizes dilution, critical for color modeling.
Brewing Time	6 min	Ensures consistent pigment release across samples.
Brewing Vessel	White porcelain pot, 150 mL capacity	Minimizes color contamination and enhances optical contrast.
Filtration	Decantation through the pot lid	Avoids particulate variability in imaging.

Tea liquor color serves as a discriminative cue which reflects regional agro-climatic variations, cultivar differences, chemical composition shifts and manufacture variation for computer vision-based classification. Experimental noise is increased for these subtle differences without standard brewing conditions. ISO 3103 reduces variance introduce by over-steeping/under-steeping, inconsistent leaf quality, variable water chemistry and non-uniform lightning in imaging environments. This ensures that machine learning and deep learning models capture true region dependent color signatures without relying on brewing anomalies [47].

3.3.3. Liquor color preparation methodology

The preparation of tea liquor samples is very difficult step for ensuring that color based analytical outcomes. It remained scientifically valid, reproducible and comparable across diverse tea growing regions. This study follows a rigorously standardized methodology grounded in ISO guidelines and it has established sensory evaluation practices.

For this research, tea samples were collected from Sri Lankan seven tea regions focusing on three widely produces black teas. Those are BOPF, BOP and OP. A 2.0g subsample of each tea sample was weighted using laboratory calibrated digital balance. The standardization controls variability in infusion strength which leads to unreliable characteristics.

The fresh boiled water was used for all infusions. A fixed 100ml volume of water has dispensed for each sample. The water temperature at the moment of infusion was maintained at (98 ± 2) °C in line with standard black tea extraction protocol.

The weighted tea samples were infused in a porcelain vessel which ensures neutral color inference. The steeping time is set to 6 minutes which consists of ISO 3103 guidance for tea sensory evaluation. It has allowed sufficient extraction of polyphenolic compounds and pigment molecules (theaflavins, thearubigins) affect the liquor color. During the steeping period, the vessel has covered to minimize thermal loss. After that liquor was filtered using stainless steel or nylon mesh which does not impart color distortions.

3.3.4. Imaging setup

The dedicated imaging lightbox was constructed to isolate samples from ambient light variations. The lightning box constructed from regiform as a rectangle shape. It has a small hole in the top corner for capturing the liquor images. The lightbox minimized the specular reflections and scattering effects which are common when imaging liquids.
The camera of the phone was mounted on top of the lightbox to maintain geometric consistency across samples. The distance between camera lens and teacup surface was standardized at 35-40 cm. It has determined calibration experiments to optimize focus sharpness while capturing spatial detail.

Exposure settings were manually configured to eliminate fluctuations caused by adjustment algorithms. The parameters such as shutter speed, ISO sensitivity and aperture were kept constant for all tea liquor samples.

3.3.5. Dataset size

The reliability and generalizability of machine learning and deep learning models for tea region classification are strongly influenced by size, diversity and representativeness of the dataset. For this study dataset was collected using tea liquor samples prepared from three commonly produced black tea grades in Sri Lanka. To ensure statistical robustness a minimum of 100 liquor image samples were collected per region. It makes total of 700 tea liquor image samples.

3.3.6. Annotation

The annotation of tea liquor image samples is a critical phase in developing supervised learning systems. In this study, annotation was conducted at sample level which ensuring precise labeling of each liquor image with respect to its origin and characteristics. Each annotation record has included region of origin, tea grade and the sample id. (ex-: DI_BOPF_001)

3.3.7. Data labeling protocol

To ensure consistency and reproducibility, a standardized data labeling protocol was implemented. The protocol defines a set of rules and guidelines to ensure that each tea liquor is enabled accurately and uniformly. This naming convention encodes region and tea grade explicitly, enable efficient dataset loading and parsing and reduce risk of mislabeling.

Table 18: 3.3.7 Tea sample labeling strategy for sample id 001

Region	Type	Sample Id	Label
Dimbula	BOPF	001	DI_BOPF_001
	OP		DI_OP_001
	BOP		DI_BOP_001
Nuwara Eliya	BOPF	001	NU_BOPF_001
	OP		NU_OP_001
	BOP		NU_BOP_001
Kandy	BOPF	001	KA_BOPF_001
	OP		KA_OP_001
	BOP		KA_BOP_001

Ruhuna	BOPF	001	RU_BOPF_001
	OP		RU_OP_001
	BOP		RU_BOP_001
Sabaragamuwa	BOPF	001	SB_BOPF_001
	OP		SB_OP_001
	BOP		SB_BOP_001
Udapussellawa	BOPF	001	UP_BOPF_001
	OP		UP_OP_001
	BOP		UP_BOP_001
Uva	BOPF	001	UV_BOPF_001
	OP		UV_OP_001
	BOP		UV_BOP_001

3.4. Image Preprocessing

Image preprocessing is a critical stage in developing a robust and generalizable tea liquor color-based tea region classification framework. These variations have been introduced by imaging devices, illumination changes, brewing inconsistencies and environmental factors which lead to significant intra-class variability. The preprocessing of images ensures that input images fed into the classification model which are standardized, noise free and color accurate.

Figure 15: 3.4 Raw tea liquor Image

Figure 14: 3.4 Preprocessed tea liquor image

3.4.1. Resizing and Background Removing

Images have been captured from the standardized imaging setup naturally vary in resolution depending on camera hardware and lens specifications. Deep learning models such as CNN architecture require fixed dimensions to maintain consistent convolution operations. In this study all liquor images are resized to 224 × 224 pixels to align with widely used architectures such as ResNet, MobileNetV2, and EfficientNet. Resizing also reduces computational load, enabling faster training while retaining essential chromatic cues relevant to regional differentiation. The background of the raw tea liquor image has been removed using OpenCV technologies.

3.4.2. Normalization

Normalization has been employed to stabilize the training process and keep pixel intensities with a consistent numeric range. Normalization helps to minimize undesired discrepancies and enhances model generalization across different tea regions due to the tea liquor images inherently exhibiting variations in brightness, illumination intensity, sensor response, and exposure settings.

Tea liquor color is a sensitive and highly discriminative parameter for differentiating Sri Lankan tea regions. But the environmental inconsistencies such as slight changes in lightbox illumination, camera aperture fluctuation or reflections. It caused the pixel value distribution to vary across images. Normalization addresses these issues by standardizing pixel intensity. It focuses on true chromatic differences rather than noise introduced during image capture.

1. It has reduced inter-sample variability from imaging conditions
2. It has improved feature stability for chromatic attributes such as brightness, saturation, and hue
3. It has enhanced training convergence due to well-bounded input distributions
4. It has used prevention of dominance by high-intensity pixels in gradient computation

3.4.3. Color calibration

Color calibration is a critical preprocessing step in imaging-based tea region classification. Because of the slight variations in illumination, sensor characteristics and environmental conditions are introduced inconsistencies in recorded tea liquor color. A machine learning or deep learning model are created for getting better accuracy. Color calibration corrects systematic errors in color acquisition by mapping image pixel values to a standardized color reference space.

Tea liquor color varies across Sri Lanka’s seven tea regions due to several factors. Those are climatic conditions, soil composition, elevation, leaf chemical constituents and manufacturing style. These variations are overshadowed by imaging artifacts. Those are uneven illumination inside the lightbox, Variability in camera sensor spectral response, Incorrect white balance, Shadows or glare on the liquor surface, Differences in lens vignetting and Aging of LED light sources.

3.4.4. Noise reduction

Noise reduction is a critical preprocessing step in image-based tea region classification. It distorted the chromatic and textural properties of tea liquor and consequently degrade model performance. Images are captured under controlled environment such as standard lightboxes. But there was various noise due to sensor limitations, lightning inconsistencies, thermal fluctuations and reflection from the liquor surface. Tea liquor images have several noise artifacts which include:

1. Sensor noise (Gaussian noise): It has been caused by electronic fluctuations in the camera sensor, and particularly under low-light or high-gain conditions.
2. Specular highlights and micro-reflections: It has occurred due to glossy surfaces of tea liquor or glassware which create bright spots that distort color uniformity.
3. Salt-and-pepper noise: It has been caused by pixel errors, dust particles and environmental interference during liquor color image acquisition.
4. Photon Noise: It has resulted from variation in lighting intensity and particularly when using non-uniform LED sources.
5. Compression artifacts: It has been introduced when images are stored or transferred in lossy formats.

3.4.5. ROI extraction

Region of Interest (ROI) extraction is a critical preprocessing step in tea liquor color–based classification. It is isolated only the informative area of the image. When capturing a raw image of tea liquor color, there exist irrelevant elements such as cup edges, shadows, background textures and reflections. Significant noise into the feature space is introduced by the incorrect or inconsistent ROI extraction. ROI extraction protocol ensures reproducible and comparable results across samples, sessions, and imaging setups.

ROI extraction focuses on segmenting the central liquor area. It is reliable to represent the color and optical characteristics of brewed tea. It preserved chromatic features while minimizing artifacts caused by light reflections, uneven illumination and cup geometry. For the image preprocessing perspective ROI extraction has involved three main stages. Those are cup boundary detection, liquor region segmentation and ROI cropping or masking. It used algorithms such as the Hough Circle Transform. Detection of the cup rim has enabled the system to estimate the inner circular region containing the liquor.

Liquor area is segmented by generating circular masks corresponding to inner region. This has ensured that the extracted ROI represents only the uniform color region of liquor. The ROI is cropped or extracted as a standalone image patch which is resized to the standardized model input size (224×224 pixels).

3.5. Model Architecture

3.5.1. CNN model

Convolutional Neural Networks have demonstrated exceptional performance in image-based classifications due to their ability to learn hierarchical visual features automatically. CNN can extract subtle spatial and chromatic patterns such as hue distributions, saturation gradients and micro texture variation of liquor surface. CNN model follows a standard deep learning architecture which is composed of four fundamental components. Those are convolution layers, pooling layers, normalization operations and non-linear activation functions. These components have enabled robust features learning from the preprocessed liquor images which resize to 224 ×224 pixels.

Convolutional layers represent the core of CNN architecture. It has applied to learnable kernels (filters) the input image to compute local feature maps. Each kernel slide has over the spatial dimension of the image. It performed on element wise multiplication and summation to generate activation responses. In this study, early convolutional layers focus on extracting low level chromatic cues associated with liquor color which include brightness gradients, color edge transitions and local tone variations and textural differences across tea types.

Pooling layers serve to progressively reduce the spatial dimensions of the feature maps. This operation has supported translation invariance, lower computational burden and prevent overfitting by suppressing noise and redundant information. Max pooling refers to the select maximum value within a local neighborhood. It is very effective for images with subtle color transitions.

Normalization stabilized the learning process by adjusting the distribution of activations within the network. There are two approaches for normalization. Those are Batch Normalization and Layer Normalization. Batch normalization is used after convolutional operations immediately. BN Normalizes intermediate activations such as:

$$\hat{x} = \frac{x - \mu_B}{\sigma_B}$$

where $\mu_B$and $\sigma_B$represent the batch mean and variance, respectively.

Activation functions give non-linearity to CNN. It has enabled the network to model complex relationships between liquor color features and tea regions. The Rectified Linear Unit (ReLU) has been used due to its computational efficiency and ability to mitigate.

$$f(x)=max(0,x)$$

ReLU allows the CNN to learn strong discriminative features by highlighting positive activation. In liquor color analysis, ReLU supports the extraction of intricate chromatic boundaries. Those are transitions between reddish, golden or dark amber liquor tones which are important indicators of cultivator, region and processing variations [48].

3.5.2. State-of-the-art architecture

CNN architecture has strong performance in image classification tasks. Among CNN models ResNet-18, EfficientNet, MobileNet, ShuffleNet and SqueezeNet have representational capacity, parameter efficiency and computational complexity.

ResNet-18 follows the concept of residual learning. It mitigates gradient vanishing issues when training deeper networks. It has skip connection and making model to learn identity mapping more effectively. In liquor color analysis, subtle chromatic variations across regions must be captured. It provides stable model architecture which is capable of learning discriminative color and texture features [49].

EfficientNet follows a compound scaling method that uniformly scales network depth, width and resolution. This balanced scaling method is used to improve accuracy compared to conventional architecture. EfficientNet model has an ability to preserve fine-grained features while maintaining low conceptual overhead. It is suitable for high resolution tea color images. Because high resolution images have minor differences in color saturation and brightness [50].

ShuffleNet uses pointwise group convolutions and channel shuffle operations. These operations have enhanced accuracy, reduced computational costs and preserved feature diversity across channels. ShuffleNet is very suitable for the scenarios where tea region classification models must be trained on low resource devices [51].

SqueezeNet uses parameter reduction through its fire modules. It compresses and expands feature representations using 1×1 and 3×3 filters. It has small model size compared to traditional CNN models. It enables fast inference with minimal storage requirements. These architectures illustrate a spectrum of options which balancing accuracy, computational efficiency and deployment requirements [52].

3.5.3. Transfer learning

Transfer learning is a strategy for improving model performance when there is limited data to be trained. Transfer learning initialized model with pre-trained weights derived from large scale datasets such as ImageNet. These learned weights got generic visual features such as edges, shapes and color gradients. In this research has adapted the transfer learning strategics when training pre-trained models like ResNet-18, EfficientNet, MobileNet, ShuffleNet and SqueezeNet.

3.5.4. Model optimization

Model optimization plays a crucial role in achieving optimal performance for tea region classification tasks. In this research Adam optimizer is adopted due to its learning rate mechanism. It computes individual learning rates for each parameter using first and second moment estimators. It has rapid coverage with stability and then makes it suitable for color-based image classification.

3.5.5. Proposed TEAQNET architecture

The TEAQNET architecture is designed for tea region classification using tea liquor color analysis. It combined efficient feature extraction modules, lightweight convolutional blocks and color sensitive operations to capture fine-grained chromatic variations across tea varieties. Input layer of this architecture accepts RGB tea liquor images which have resized to 224 ×224 pixels. It has applied channel-wise normalization and color sensitive convolution filters to emphasize minor hue differences. It uses global average pooling for reducing spatial dimensions.

TEAQNET generates progressively abstract feature maps. Its early layers capture the basic color gradients, middle layers extract the region-specific chromatic signatures, and deeper layers encode high level differentiating cues. There are many hyperparameters used for training this.

Table 19: 3.5.5 Hyperparameters of TEAQNET

Hyperparameter	Description & Values
Optimizer	Adam
Initial Learning Rate	0.001
Learning rate Schedular	Factor =0.1, patience =5
Batch Size	32
Epochs	50
Weight Initialization	Normal
Regularization	L2 weight decay (1e-4), Dropout (0.4)
Data Augmentation	Random Rotation, Brightness Shift and Horizontal flipping

3.5.6. Loss function

The model is optimized using Cross-Entropy Loss (Loss Function). It is widely used in multi-class image classification tasks. It measures the divergence between the predicted probability distribution and ground-truth distribution. According to the given image, the model resulting the output class probabilities corresponding to different tea regions. The loss function represents incorrect predictions by increasing the loss value. Mathematically, cross-entropy loss is defined as:

$$\mathcal{L} = -\sum_{i=1}^{C} 1\, y_i \log(\widehat{y_i})$$

where C represents the number of tea regions, y_i denotes the true class label, and $\widehat{y_i}$ denotes the predicted probability for class i. Cross – entropy is very useful to stable coverage, and it performs well with SoftMax-based classifiers [53].

3.5.7. Regularization

Regularization techniques are used to enhance the model's generalization ability. It also reduces the risk of overfitting when dataset size is limited. L2 regularization penalizes use large weight parameters by adding a proportional term to the loss function. This factor encourages the model to maintain smaller and more stable weights.

3.5.8. Training setup

Training setup is designed to ensure stability, reproducibility and optimal performance of the TEAQNET. The training is conducted on GPU-enabled environment to accelerate computation for convolutional operations. It has been designed with NVDIA GPU and CUDA. The dataset is divided into training, validation and testing. In dataset 10% of images are allocated for testing phase. Out of other images, 60% of images are allocated for the training phase and remaining portion has been allocated for the validation phase. A batch size of 32 is used to balance memory usage and gradient stability with 50 epochs. It has an Adam optimizer and learning rate strategy. In these criteria early stopping has employed. It prevents unnecessary training cycles and avoids overfitting.

3.6. Evaluation Metrics

The evaluation metrices use to assess the performance of proposed TEAQNET architecture. It has multiple evaluation architecture. Since tea region classification is a multiclass classification problem, accuracy does not fully cover the model effectiveness. Therefore, precision, recall, F1-score and Intersection Over Union (IOU) are considered [54].

3.6.1. Accuracy

Accuracy measures the portion of tea liquor samples among the total number of samples that are correctly classified. It presents the classification capability of the model.

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

where TP represents true positives, TN represents true negatives, FP represents false positives, and FN represents false negatives.

3.6.2. Precision

Precision used to represent the correctness of positive predictions made by the model. In this research high precision indicates that images classified as a particular region actually belong to that region.

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

where TP represents true positives and FP represents false positives.

3.6.3. Recall

Recall also known as sensitivity. It measures the model's ability to correctly identify all images of specific tea region. If recall is high, it represents most of the tea liquor images from given region are correctly detected.

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

where TP represents true positives and FN represents false negatives.

3.6.4. F1-Score

F1-score combines both precision and precision for providing a balance measure. F1-score provides more reliable indicators of model performance by specifying both false positive and false negative.

$$\text{F1} - \text{score} = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

3.6.5. IOU

Intersection over union (IOU) also known as Jaccard index. It measures the overlap between predicted class region and the ground-truth region. IOU provides insight into the consistency of predictions across classes when analyzing confusion between closely related tea regions.

$$\text{IOU} = \frac{TP}{TP + FP + FN}$$

where TP represents true positives, FP represents false positives, and FN represents false negatives.

3.7. Experimental Setup

The experimental setup defines the computational environment, software framework and dataset allocating strategy which is used to develop the TEAQNET. It ensured the reproducibility, consistency and fairness of performance evaluation for tea region classification based on liquor color analysis.

3.7.1. Hardware

All experiments were conducted on a Dell VPro workstation which is equipped with a Core i9 processor. It has high computational capability for deep learning workloads. It has a system which is supported by a dedicated GPU with CUDA acceleration. It enabled efficient training of Convolutional Neural networks and reduced the overall training time. This hardware configuration has provided a stable high-performance environment for executing computationally intensive tasks like feature extraction, model optimization and hyperparameter tunning.

3.7.2. Software

The proposed system was created using PyTorch. It is an open-source deep learning framework. It has been widely adopted for its dynamic computation graph, flexibility and efficient GPU utilization. It is very useful for rapid model prototyping, fine-grained control over training loops and seamless integration of TEAQNET.

OpenCV is another software library which is used for the TEAQNET. It has been used for image preprocessing, resizing, color normalization, noise reduction and region of interest extraction from tea liquor images. Both software libraries provide robust software pipelines which are capable of handling both deep learning computation and classical image processing tasks which require accurate tea region classification [55].

3.7.3. Training/validation/testing split

The dataset was partitioned into training, validation and testing subsets. There were 700 tea liquor images. Out of those images 10% of images have allocated to testing subset. From remaining 630 images, 60% of images have allocated to training subset and remaining to validation subset. These allocations are done by using random split techniques.

3.7.4. Comparative models

TEAQNET architecture is compared against several traditional machine learning models to provide comprehensive performance evaluation. For fair comparison, all models were trained and evaluated using same dataset partitions and preprocessing pipeline. Initially features of liquor color images are converted into numerical format, which can be applicable for training machine learning models. In this research machine learning models such as Support Vector Machine (SVM), Random Forest, K- Nearest Neighbors (KNN) and Logistic Regression have been used.

3.8. Risk Management

Risk management strategies address potential challenges and ensure the robust performance in tea region classification. It is essential that avoid the data imbalance, overfitting and environmental noise. Overfitting occurs when a model learns noise or dataset-specific patterns rather than generalizable features. In this research has used dropout, L2 weight decay, early stopping and validation-based rate scheduling for preventing overfitting.

3.9. Chapter Summary

This chapter presented methodology for tea region classification based on tea liquor color analysis. It has outlined the systematic design, implementation and evaluation of the proposed approach. This chapter began by describing data preprocessing techniques which applied to enhance image quality and normalization techniques.

The model architecture has detailed that covering CNN models. It has introduced domain specific solutions which emphasize its lightweight design and hierarchical feature extraction. It has explained about transfer learning strategies and optimization techniques which including Adam Optimizer and Learning rate schedular.

This chapter also discussed loss function, regularization methods and training setup which are used to ensure robust and stable model learning. For the evaluations, the metrices such as accuracy, precision, recall, F1-score and intersection over union (IOU) have been used. Additionally, for comparative analyses machine learning classifiers like SVM, Random Forest, KNN and Logistic Regression models are used. Finally, this chapter has discussed potential risks which include data imbalance, overfitting and environmental noise.

4. Chapter 04 - Results & Analysis

4.1. Chapter Introduction

This chapter represents detailed analysis of experiment results which have been obtained from the proposed TEAQNET architecture. The objective of this chapter is to evaluate the effectiveness, robustness and generalization of proposed architecture through quantitative and qualitative analysis. The results are analyzed using learning curves, confusion matrixes, ROC curves, principal component analysis and standard evaluation metrics such as accuracy, precision, recall, F1-score and Intersection Over Union (IOU).

4.2. Learning Curve Analysis

Learning curves represent the training and validation performance of the model over successive epochs. It provided insight into convergence behavior, learning stability and potential overfitting and underfitting.

Table 20: 4.2 Learning Curve Interpretation

Scenarios	Curve Behavior	Interpretation
Overfitting	Training accuracy is very high, and validation accuracy stagnates or low	Model memorized training data, poor generalization
Underfitting	Training & Validation accuracy, both low	Model is too simple or not trained enough, it cannot capture patterns.
Good Fit	Training & Validation accuracy close and high	Model generalizes well

The CNN models such as RestNet-18, ShuffleNetV2, MobileNetV2, EfficientNetb0 and SqueezeNet are used for model comparisons among deep learning model architectures.

4.2.1. RestNet-18

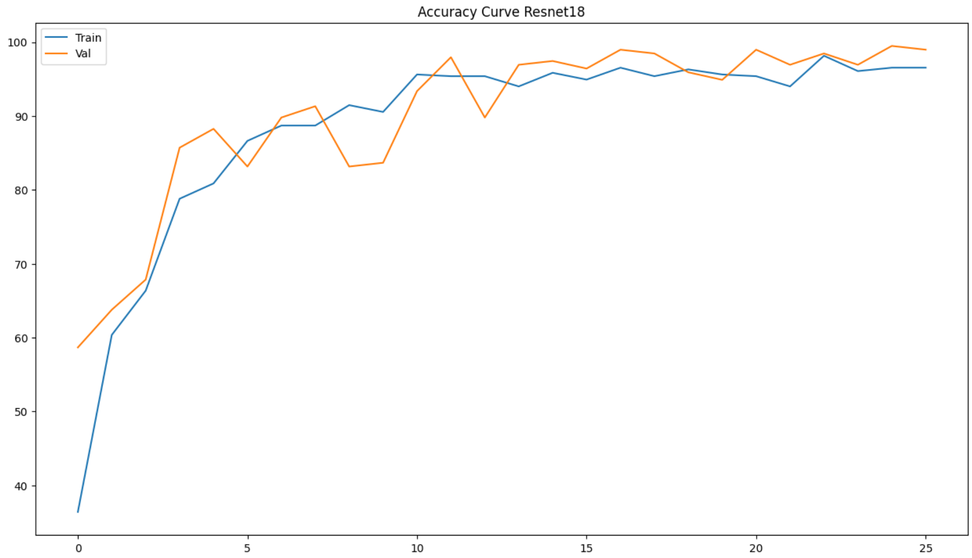

Figure 16: 4.2.1 Accuracy Curve - ResNet-18

Figure 16: 4.2.1 Accuracy Curve - ResNet-18 illustrates that training and validation accuracy of RestNet-18 model which trained for tea region classification. Both training and validation accuracy have increased from epoch 0 to 25. Initially the model was supposed to be trained on 50 epochs. In epoch 25, early stopping has triggered. Therefore, training procedure has terminated for saving the best model.

For a best fit model, validation accuracy should be higher than the training accuracy and there should be a small gap between those two learning curves. Figure 16: 4.2.1 Accuracy Curve - ResNet-18 illustrates that validation accuracy is more than the training accuracy and there is small gap between those two lines in final epoch 25. Therefore, RestNet-18 model is good fit and generalizes well.

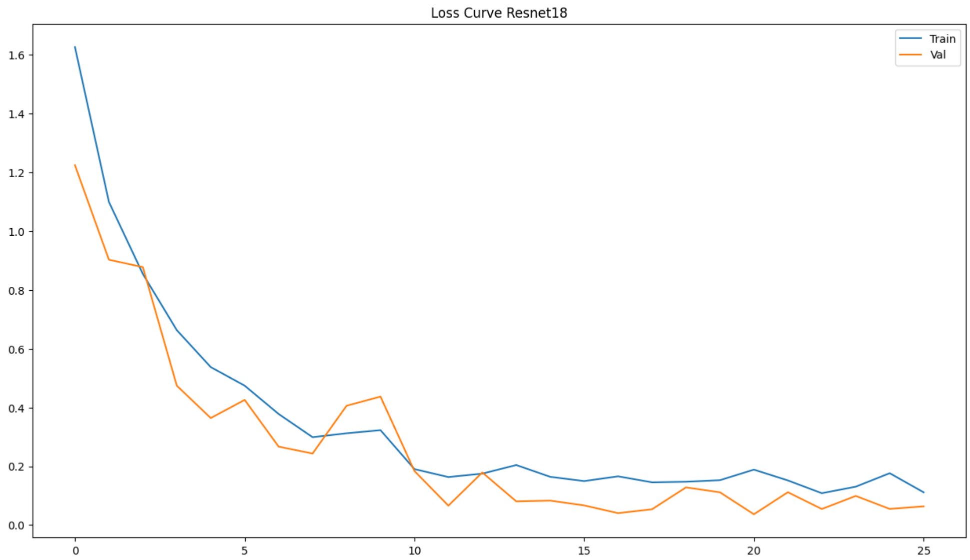

Figure 17: 4.2.1 Loss Curve - RestNet-18

Figure 17: 4.2.1 Loss Curve - RestNet-18 illustrates that training and validation loss of RestNet-18 model which trained for tea region classification. Both training and validation loss have decreased from epoch 0 to 25. Initially the model was supposed to be trained on 50 epochs. In epoch 25, early stopping has triggered. Therefore, training procedure has terminated for saving the best model.

For a best fit model, validation loss should be lower than the training loss and there should be a small gap between those two learning curves. Figure 17: 4.2.1 Loss Curve - RestNet-18 illustrates that validation loss is lower than the training loss and there is small gap between those two lines in final epoch 25. Therefore, RestNet-18 model is good fit and generalizes well. The model has got 98.98% accuracy in training approach.

4.2.2. ShuffleNetV2

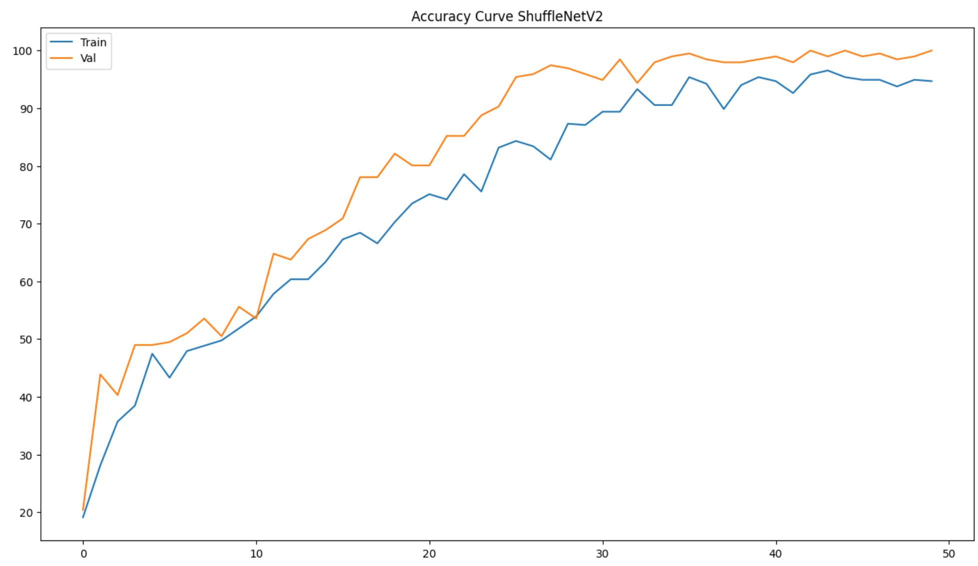

Figure 18: 4.2.2 Accuracy Curve- ShuffleNetV2

Figure 18: 4.2.2 Accuracy Curve- ShuffleNetV2 illustrates that training and validation accuracy of ShuffleNetV2 model which trained for tea region classification. Both training and validation accuracy have increased from epoch 0 to 50. Initially the model was supposed to be trained on 50 epochs. The model was trained for all its 50 epochs.

For a best fit model, validation accuracy should be higher than the training accuracy and there should be a small gap between those two learning curves. Figure 18: 4.2.2 Accuracy Curve- ShuffleNetV2 illustrates that validation accuracy is more than the training accuracy and there is small gap between those two learning curves in final epoch 50. Therefore, ShuffleNetV2 model is good fit and generalizes well.

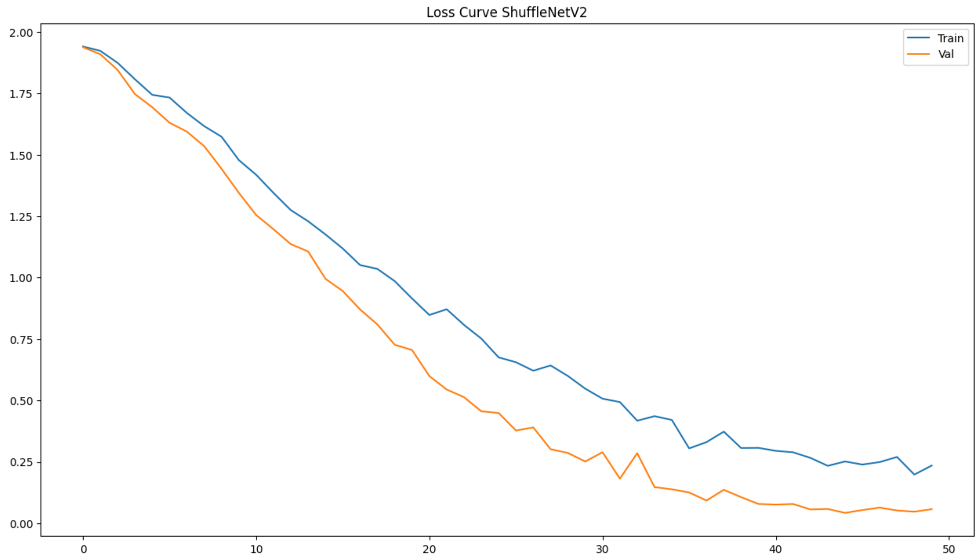

Figure 19: 4.2.2 Loss Curve - ShuffleNetV2

Figure 19: 4.2.2 Loss Curve - ShuffleNetV2 illustrates that training and validation loss of ShuffleNetV2 model which trained for tea region classification. Both training and validation loss have decreased from Epoch 0 to 50. Initially the model was supposed to be trained on 50 epochs. The model was trained for all its 50 epochs.

For a best fit model, validation loss should be lower than the training loss and there should be a small gap between those two learning curves. Figure 19: 4.2.2 Loss Curve - ShuffleNetV2 illustrates that validation loss is lower than the training loss and there is small gap between those two lines in final epoch 50. Therefore, ShuffleNetV2 model is good fit and generalizes well. The model has got 99.99% accuracy in training approach.

4.2.3. MobileNetV2

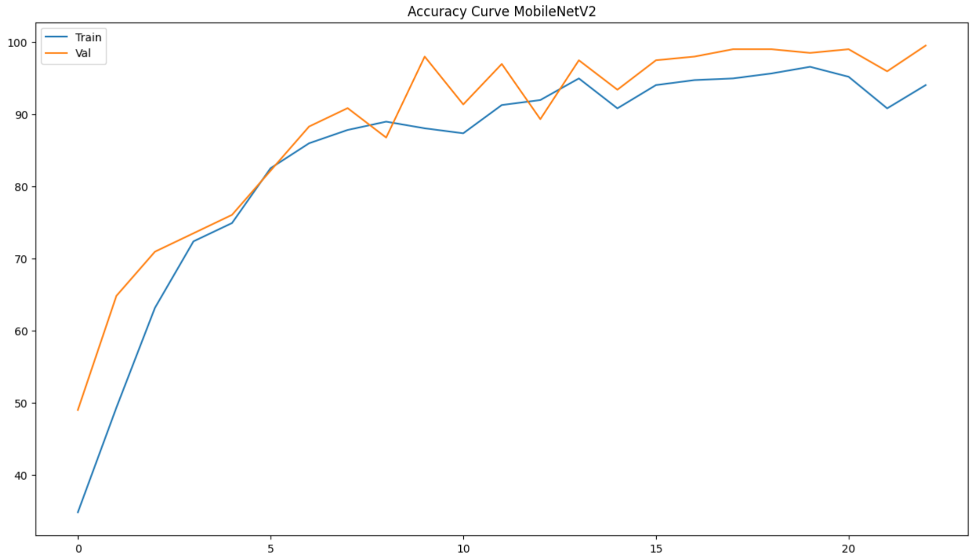

Figure 20: 4.2.3 Accuracy Curve - MobileNetV2

Figure 20: 4.2.3 Accuracy Curve - MobileNetV2 illustrates the training and validation accuracy of MobileNetV2 model which trained for tea region classification. Both training and validation accuracy have increased from epoch 0 to 23. Initially the model was supposed to be trained on 50 epochs. In epoch 23, early stopping has triggered. Therefore, training procedure has terminated for saving the best model.

For a best fit model, validation accuracy should be higher than the training accuracy and there should be a small gap between those two learning curves. Figure 20: 4.2.3 Accuracy Curve - MobileNetV2 illustrates that validation accuracy is more than the training accuracy and there is average gap between those two lines in final epoch 23. Therefore, MobileNetV2 model is moderately good fit and generalizes well averagely.

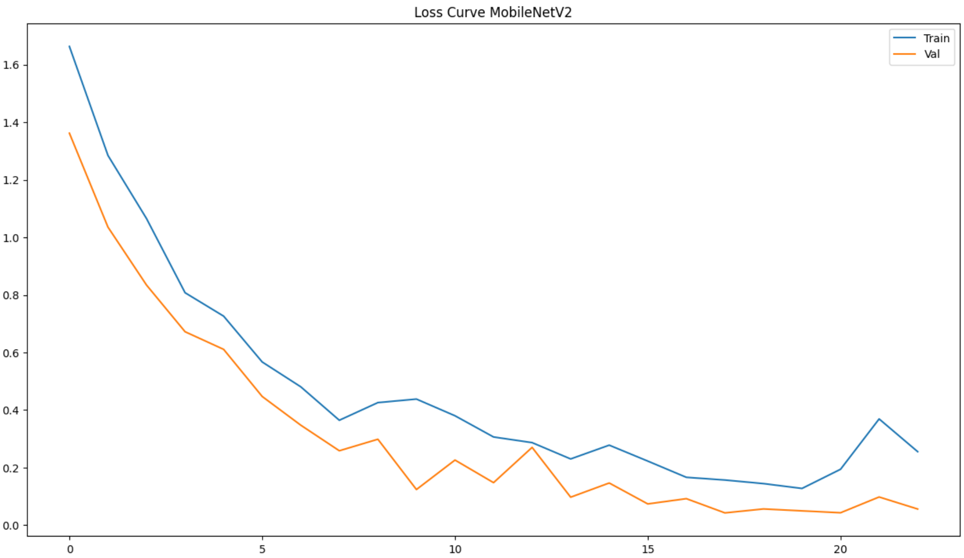

Figure 21: 4.2.3 Loss Curve - MobileNetV2

Figure 21: 4.2.3 Loss Curve - MobileNetV2 illustrates that training and validation loss of MobileNetV2 model which trained for tea region classification. Both training and validation loss have decreased from epoch 0 to 23. Initially the model was supposed to be trained on 50 epochs. In epoch 23, early stopping has triggered. Therefore, training procedure has terminated for saving the best model.

For a best fit model, validation loss should be lower than the training loss and there should be a small gap between those two learning curves. Figure 21: 4.2.3 Loss Curve - MobileNetV2 illustrates that validation loss is lower than the training loss and there is average gap between those two lines in final epoch 25. Therefore, MobileNetV2 model is moderately good fit and generalizes well averagely. The model has got 99.49% accuracy in training approach.

4.2.4. EfficientNetb0

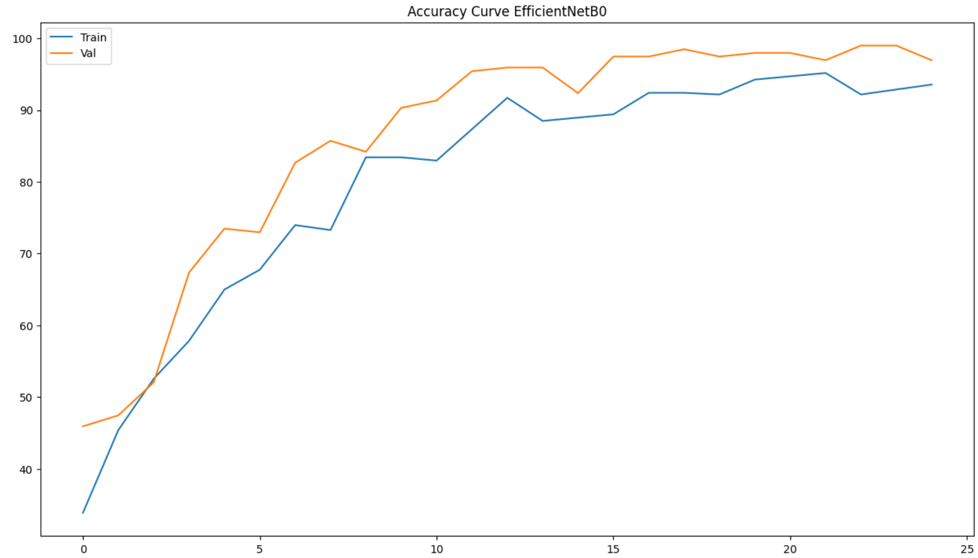

Figure 22: 4.2.4 Accuracy Curve - EfficientNetb0

Figure 22: 4.2.4 Accuracy Curve - EfficientNetb0 illustrates that training and validation accuracy of EfficientNetb0 model which trained for tea region classification. Both training and validation accuracy have increased from epoch 0 to 25. Initially the model was supposed to be trained on 50 epochs. In epoch 25, early stopping has triggered. Therefore, training procedure has terminated for saving the best model.

For a best fit model, validation accuracy should be higher than the training accuracy and there should be a small gap between those two learning curves. Figure 22: 4.2.4 Accuracy Curve - EfficientNetb0 illustrates that validation accuracy is more than the training accuracy and there is small gap between those two lines in final epoch 25. Therefore, EfficientNetb0 model is good fit and generalizes well.

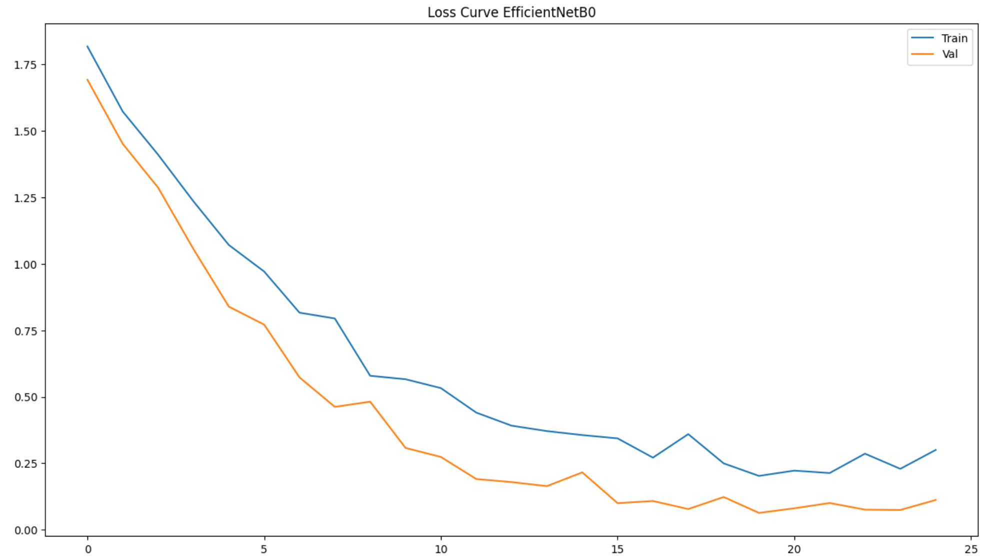

Figure 23: 4.2.4 Loss Curve - EfficientNetb0

Figure 23: 4.2.4 Loss Curve - EfficientNetb0 illustrates that training and validation loss of EfficientNetb0 model which trained for tea region classification. Both training and validation loss have decreased from epoch 0 to 25. Initially the model was supposed to be trained on 50 epochs. In epoch 25, early stopping has triggered. Therefore, training procedure has terminated for saving the best model.

For a best fit model, validation loss should be lower than the training loss and there should be a small gap between those two learning curves. Figure 23: 4.2.4 Loss Curve - EfficientNetb0 illustrates that validation loss is lower than the training loss and there is small gap between those two lines in final epoch 25. Therefore, EfficientNetb0 model is good fit and generalizes well. The model has got 97.00% accuracy in training approach.

4.2.5. SqueezeNet

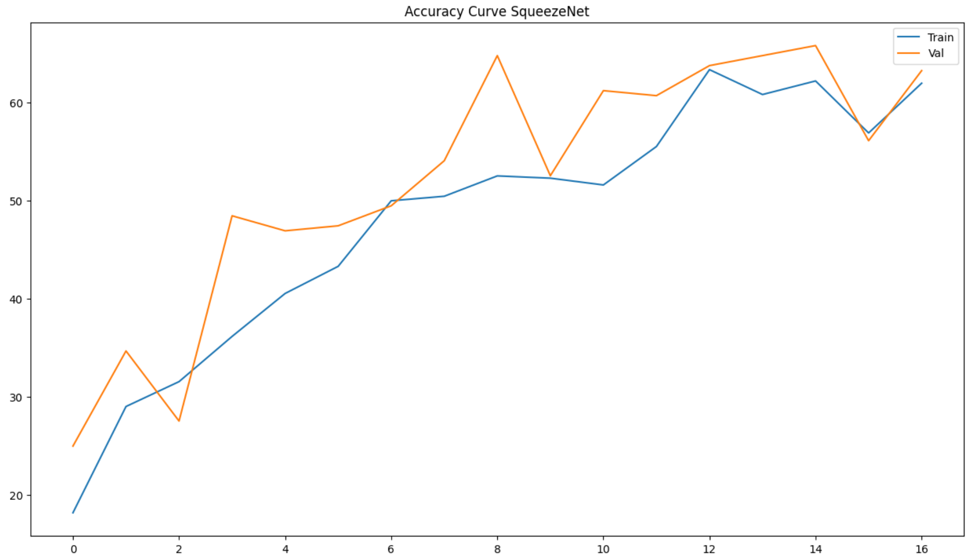

Figure 24: 4.2.5 Accuracy Curve – SqueezeNet

Figure 24: 4.2.5 Accuracy Curve – SqueezeNet illustrates that training and validation accuracy of SqueezeNet model which trained for tea region classification. Both training and validation accuracy have increased from epoch 0 to 16. But it didn’t increase linearly. Initially the model was supposed to be trained on 50 epochs. In epoch 16, early stopping has triggered. Therefore, training procedure has terminated for saving the best model.

For a best fit model, validation accuracy should be higher than the training accuracy and there should be a small gap between those two learning curves. Figure 24: 4.2.5 Accuracy Curve – SqueezeNet illustrates that validation accuracy is more than the training accuracy and there is small gap between those two lines in final epoch 16. Therefore, SqueezeNet model is good fit and generalizes well.

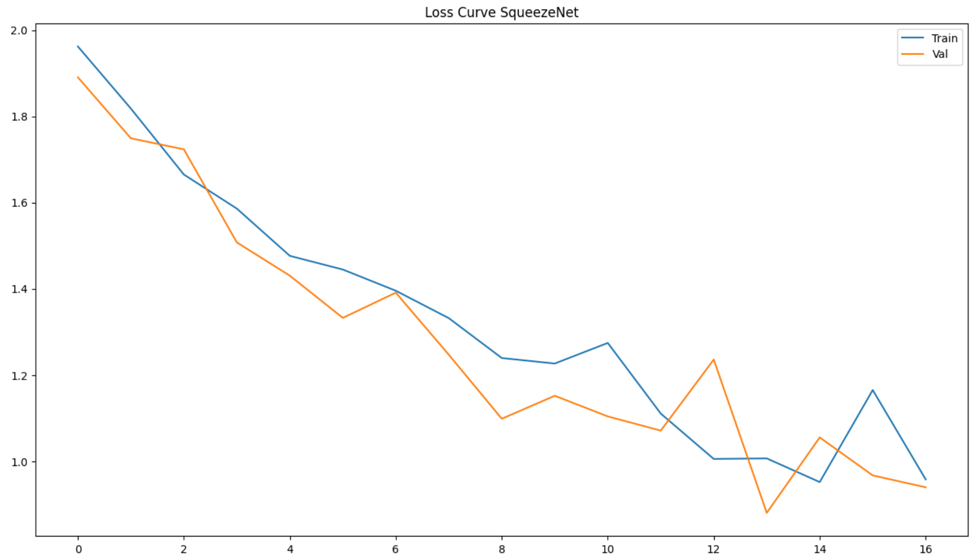

Figure 25: 4.2.5 Loss Curve – SqueezeNet

Figure 25: 4.2.5 Loss Curve – SqueezeNet illustrates that training and validation loss of SqueezeNet model which trained for tea region classification. Both training and validation loss have decreased from epoch 0 to 16. But it didn't decrease linearly. Initially the model was supposed to be trained on 50 epochs. In epoch 16, early stopping has triggered. Therefore, training procedure has terminated for saving the best model.

For a best fit model, validation loss should be lower than the training loss and there should be a small gap between those two learning curves. Figure 25: 4.2.5 Loss Curve – SqueezeNet illustrates that validation loss is lower than the training loss and there is small gap between those two lines in final epoch 16. Therefore, SqueezeNet model is good fit and generalizes well. The model has got 63.27% accuracy in training approach.

4.3. Confusion Matrix Analysis

The confusion matrix provides a class-wise breakdown of prediction performance. It has highlighted both correct and misclassified instances across tea regions. It represents the actual class while each column corresponds to the predicted class. The misclassifications are observed between visually similar regions. The CNN models such as RestNet-18, ShuffleNetV2, MobileNetV2, EfficientNetb0 and SqueezeNet are used for model comparisons among deep learning model architectures.

4.3.1. RestNet-18

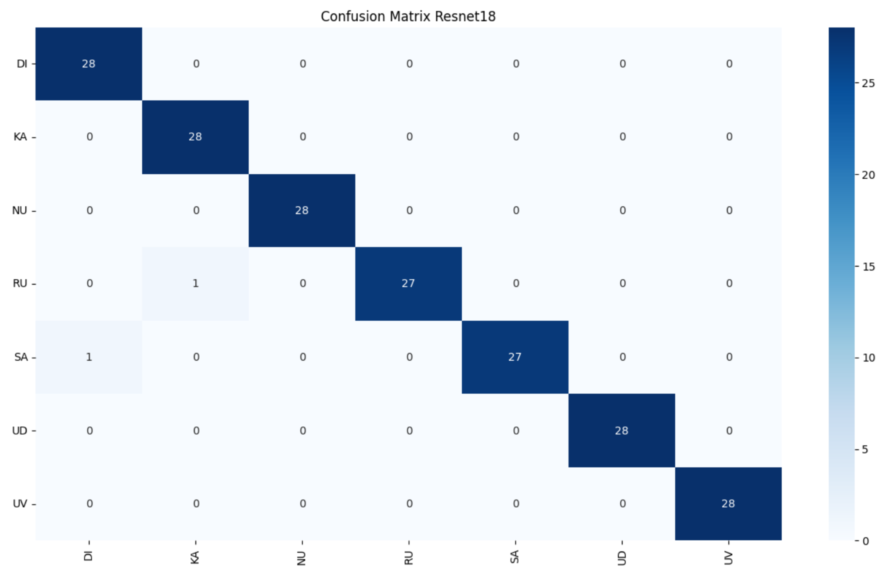

Figure 26: 4.3.1 Confusion Matrix - RestNet-18

Figure 26: 4.3.1 Confusion Matrix - RestNet-18 illustrates that RestNet-18 model for classifying tea samples into seven tea growing regions in Sri Lanka. The confusion matrix represents strongly dominant diagonal structure which indicates high number of samples are classified correctly.

Out of 196 tea liquor samples, 194 samples were correctly classified across all regions. It has demonstrated the effectiveness and robustness of the proposed deep learning approach. Among seven tea regions Dimbula , Kandy, Uva, Udapussellawa and Nuwara Eliya have achieved perfect classification. There were minor misclassifications observed in Ruhuna and Sabaragamuwa regions where one sample from Ruhuna, classified as Kandy and one of sample from Sabaragamuwa, classified as Dimbula.

The minimal off-diagonal values represent the negligible inter-class confusion. It has confirmed that liquor color serves as a highly discriminative feature for tea region classification. In conclusion, the confusion matrix illustrates that proposed RestNet-18 based model has achieved perfect regional discrimination using liquor color attributes.

4.3.2. ShuffleNetV2

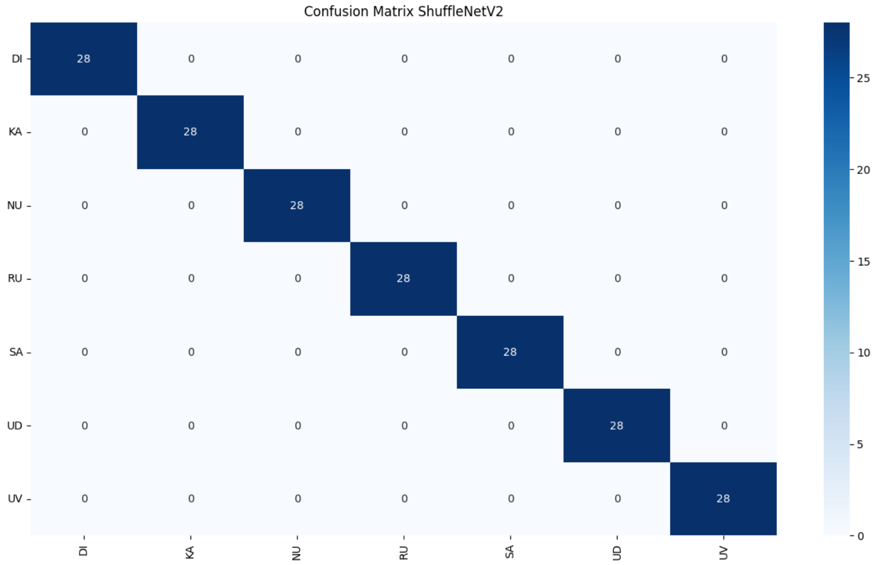

Figure 27: 4.3.2 Confusion Matrix - ShuffleNetV2

Figure 27: 4.3.2 Confusion Matrix - ShuffleNetV2 illustrates that ShuffleNetV2 model for classifying tea samples into seven tea growing regions in Sri Lanka. The confusion matrix represents strongly dominant diagonal structure which indicates high number of samples are classified correctly.

Out of 196 tea liquor samples, all samples that were correctly classified across all regions. It has demonstrated the effectiveness and robustness of the proposed deep learning approach. Seven tea regions of Sri Lanka such as Dimbula , Kandy, Uva, Udapussellawa , Ruhuna, Sabaragamuwa and Nuwara Eliya have achieved perfect classification. There were not minor misclassifications observed in any region.

The minimal off-diagonal values represent the negligible inter-class confusion. It has confirmed that liquor color serves as a highly discriminative feature for tea region classification. In conclusion, the confusion matrix illustrates that proposed ShuffleNetV2 based model has achieved perfect regional discrimination using liquor color attributes.

4.3.3. MobileNetV2

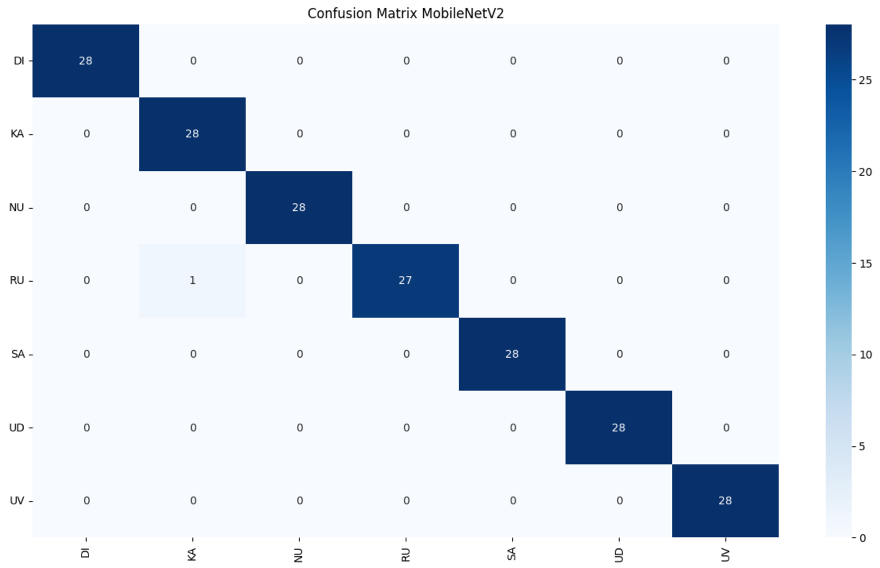

Figure 28: 4.3.3 Confusion Matrix - MobileNetV2

Figure 28: 4.3.3 Confusion Matrix - MobileNetV2 illustrates that MobileNetV2 model for classifying tea samples into seven tea growing regions in Sri Lanka. The confusion matrix represents strongly dominant diagonal structure which indicates high number of samples are classified correctly.

Out of 196 tea liquor samples, there were 195 samples that were correctly classified across all regions. It has demonstrated the effectiveness and robustness of the proposed deep learning approach. Among seven tea regions Dimbula , Sabaragamuwa, Kandy, Uva, Udapussellawa and Nuwara Eliya have achieved perfect classification. There were minor misclassifications observed in Ruhuna region where one of sample from Ruhuna classified as Kandy.

The minimal off-diagonal values represent the negligible inter-class confusion. It has confirmed that liquor color serves as a highly discriminative feature for tea region classification. In conclusion, the confusion matrix illustrates that proposed MobileNetV2 based model has achieved perfect regional discrimination using liquor color attributes.

4.3.4. EfficientNetb0

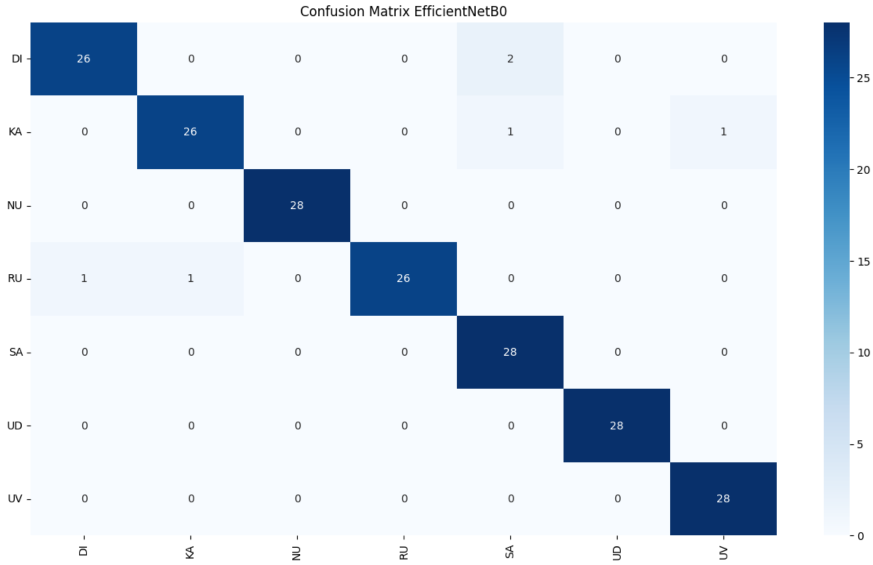

Figure 29: 4.3.4 Confusion Matrix - EfficientNetb0

Figure 29: 4.3.4 Confusion Matrix - EfficientNetb0 illustrates that EfficientNetb0 model for classifying tea samples into seven tea growing regions in Sri Lanka. The confusion matrix represents strongly dominant diagonal structure which indicates high number of samples are classified correctly.

Out of 196 tea liquor samples, there were 190 samples that were correctly classified across all regions. It has demonstrated the effectiveness and robustness of the proposed deep learning approach. Among seven tea regions Uva, Udapussellawa, Sabaragamuwa and Nuwara Eliya have achieved perfect classification. There were minor misclassifications observed in Ruhuna Kandy and Dimbula regions where some samples are classified as incorrect.

The minimal off-diagonal values represent the negligible inter-class confusion. It has confirmed that liquor color serves as a highly discriminative feature for tea region classification. In conclusion, the confusion matrix illustrates that proposed EfficientNetb0 based model has achieved perfect regional discrimination using liquor color attributes.

4.3.5. SqueezeNet

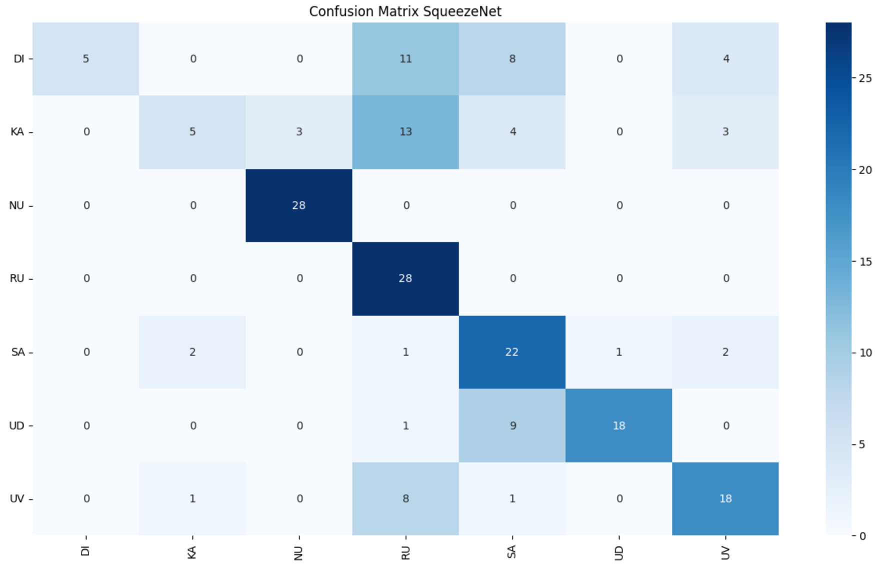

Figure 30: 4.3.5 Confusion Matrix - SqueezeNet

Figure 30: 4.3.5 Confusion Matrix – SqueezeNet illustrates that SqueezeNet model for classifying tea samples into seven tea growing regions in Sri Lanka. The confusion matrix represents average dominant diagonal structure which indicates high number of samples are classified correctly.

Out of 196 tea liquor samples, there were 194 samples that were correctly classified across all regions. It has demonstrated the effectiveness and robustness of the proposed deep learning approach. Among seven tea regions Nuwara Eliya and Ruhuna have achieved perfect classification. There were huge misclassifications observed in other regions.

The minimal off-diagonal values represent the negligible inter-class confusion. It has confirmed that liquor color serves as a highly discriminative feature for tea region classification. In conclusion, the confusion matrix illustrates that proposed SqueezeNet based model has achieved lower regional discrimination using liquor color attributes.

4.3.6. SVM

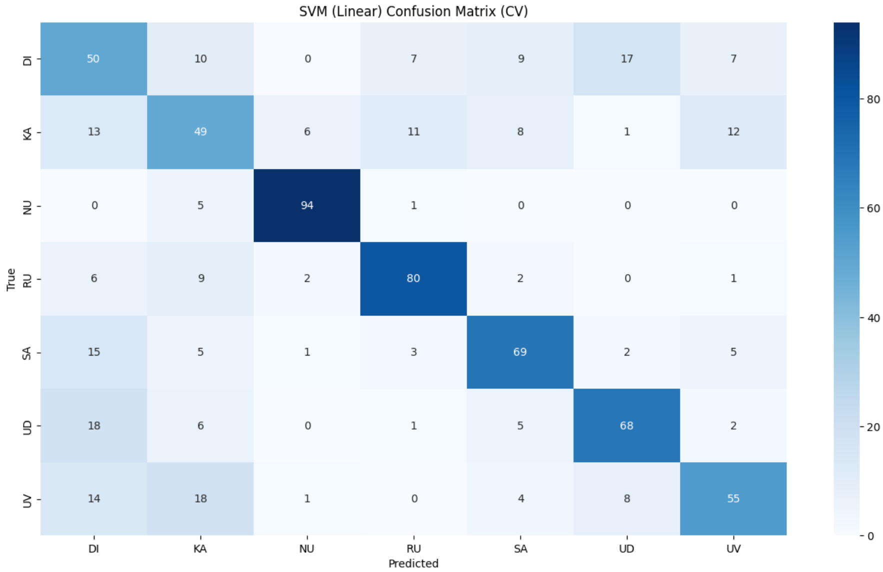

Figure 31: 4.3.6 Confusion Matrix - SVM

Figure 31: 4.3.6 Confusion Matrix – SVM illustrates that SVM model for classifying tea samples into seven tea growing regions in Sri Lanka. The confusion matrix represents week dominant diagonal structure which indicates lower number of samples are classified correctly.

Out of 700 tea liquor samples, there were 465 samples that were correctly classified across all regions. It has demonstrated the effectiveness and robustness of the proposed machine learning approach. Among seven tea regions there were no regions that classified all the samples correctly.

The minimal off-diagonal values represent the negligible inter-class confusion. It has confirmed that liquor color serves as a highly discriminative feature for tea region classification. In conclusion, the confusion matrix illustrates that proposed SVM based model has achieved lower discrimination using liquor color attributes.

4.3.7. Random Forest

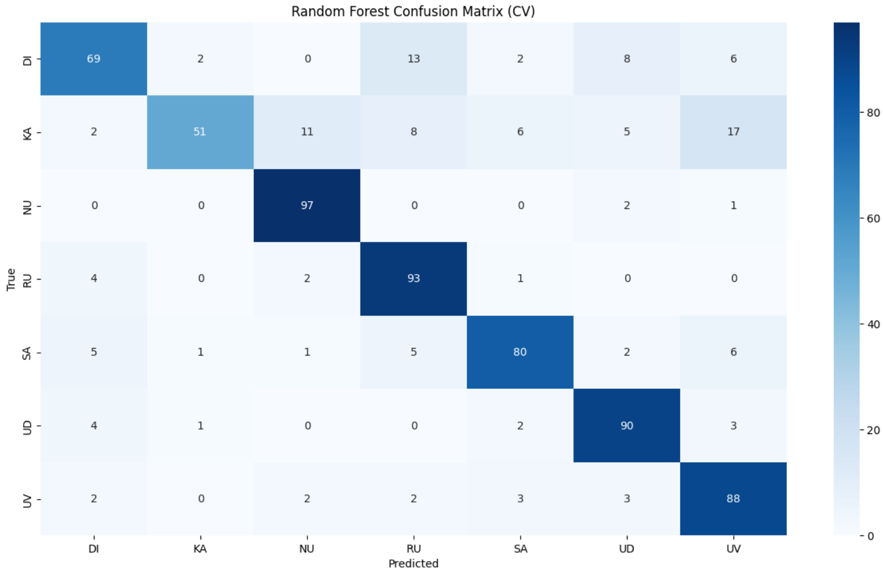

Figure 32: 4.3.7 Confusion Matrix - Random Forest

Figure 32: 4.3.7 Confusion Matrix - Random Forest illustrates that Random Forest model for classifying tea samples into seven tea growing regions in Sri Lanka. The confusion matrix represents week dominant diagonal structure which indicates lower number of samples are classified correctly.

Out of 700 tea liquor samples, there were 568 samples that were correctly classified across all regions. It has better classification performance than SVM. It has also demonstrated the effectiveness and robustness of the proposed machine learning approach. Among seven tea regions, Nuwara Eliya, Udapussellawa, Uva and Ruhuna have better classification. But other regions have more misclassifications.

The minimal off-diagonal values represent the negligible inter-class confusion. It has confirmed that liquor color serves as a highly discriminative feature for tea region classification. In conclusion, the confusion matrix illustrates that proposed Random Forest based model has achieved lower discrimination using liquor color attributes.

4.3.8. KNN

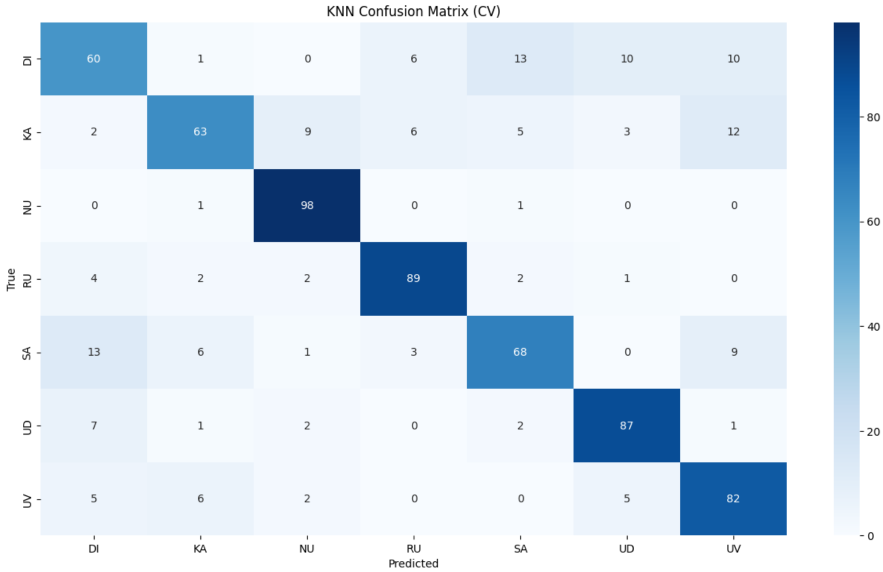

Figure 33: 4.3.8 Confusion Matrix - KNN

Figure 33: 4.3.8 Confusion Matrix – KNN illustrates that KNN model for classifying tea samples into seven tea growing regions in Sri Lanka. The confusion matrix represents week dominant diagonal structure which indicates lower number of samples are classified correctly.

Out of 700 tea liquor samples, there were 547 samples that were correctly classified across all regions. It has better classification performance than SVM, but not than Random Forest. It has demonstrated the effectiveness and robustness of the proposed machine learning approach. Among seven tea regions Nuwara Eliya, Ruhuna, Udapussellawa and Uva have better classification. But other regions have more misclassifications.

The minimal off-diagonal values represent the negligible inter-class confusion. It has confirmed that liquor color serves as a highly discriminative feature for tea region classification. In conclusion, the confusion matrix illustrates that proposed KNN based model has achieved lower discrimination using liquor color attributes.

4.3.9. Logistic Regression

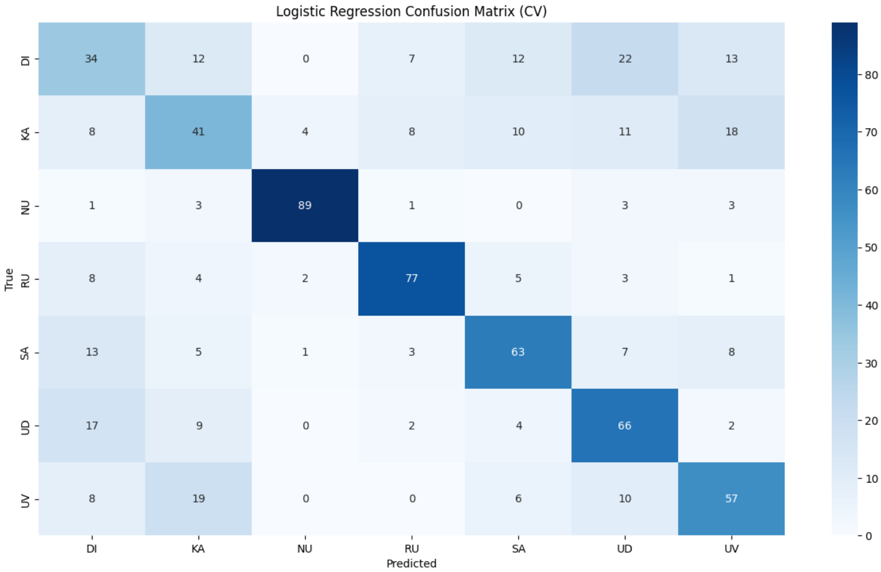

Figure 34: 4.3.9 Confusion Matrix - Logistic Regression

Figure 34: 4.3.9 Confusion Matrix - Logistic Regression illustrates that Logistic Regression model for classifying tea samples into seven tea growing regions in Sri Lanka. The confusion matrix represents week dominant diagonal structure which indicates lower number of samples are classified correctly.

Out of 700 tea liquor samples, there were 427 samples that were correctly classified across all regions. It has lower classification performance across other Machine learning models. It has demonstrated the effectiveness and robustness of the proposed machine learning approach. Among seven tea regions there were no regions that classified all the samples correctly.

The minimal off-diagonal values represent the negligible inter-class confusion. It has confirmed that liquor color serves as a highly discriminative feature for tea region classification. In conclusion, the confusion matrix illustrates that proposed Logistic Regression based model has achieved lower discrimination using liquor color attributes.

4.4. Roc Curve Analysis

Receiver Operating Characteristic (ROC) curves are used to evaluate the trade-off between true positive rates and false positive rates. It is conducted using one-vs-rest strategy in multiclass classifications. The CNN models such as RestNet-18, ShuffleNetV2, MobileNetV2, EfficientNetb0 and SqueezeNet are used for model comparisons among deep learning model architectures.

4.4.1. RestNet-18

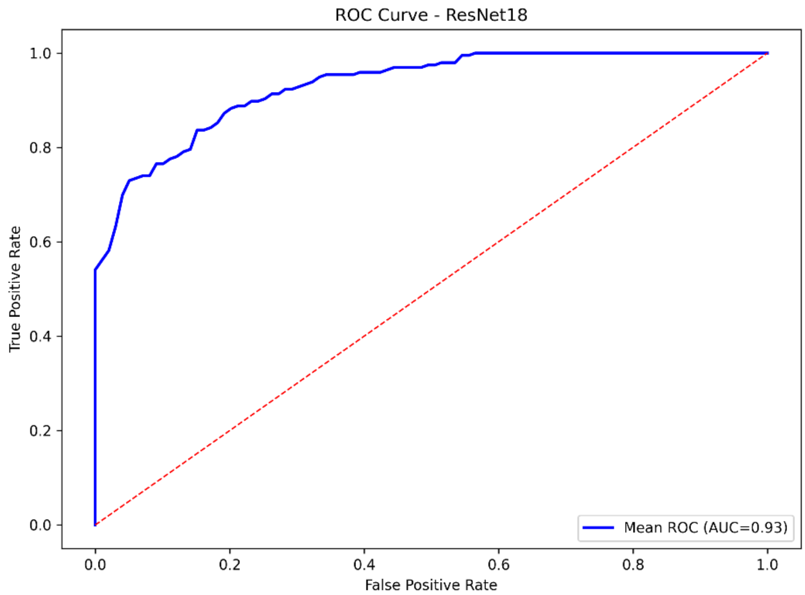

Figure 35: 4.4.1 Roc Curve - RestNet-18

Figure 35: 4.4.1 Roc Curve - RestNet-18 illustrates that ROC curve of RestNet-18 model for tea region classification based on tea liquor color. It represents the trade-off between true positive rate (TPR) and false positive rate (FPR) across different decision thresholds.

It has obtained mean Area Under the Curve (AUC) value is 0.93. It has excellent discriminative capability of the proposed model. The ROC curve remains consistently above the diagonal line which indicates that model achieved high true positive rate while maintaining low false positive rates for all seven tea region samples. These results have confirmed that liquor color provides strong classification among different tea growing regions.

4.4.2. ShuffleNetV2

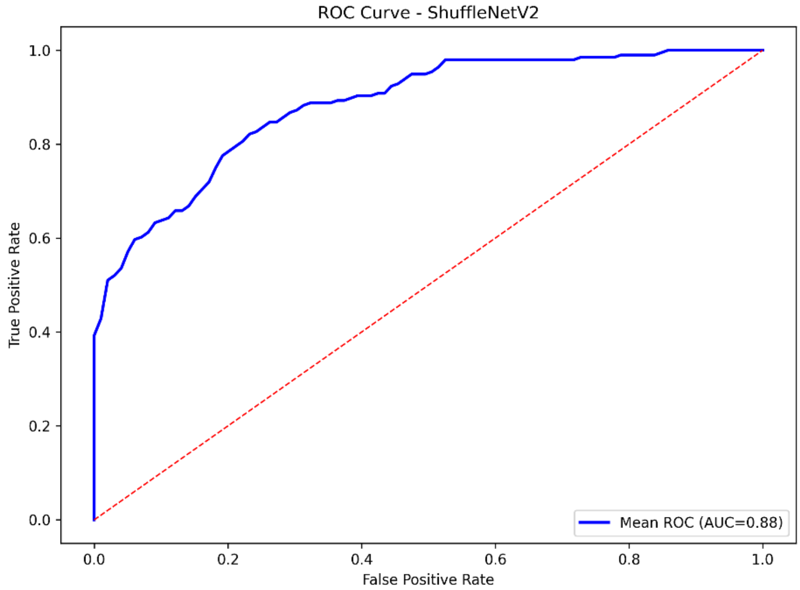

Figure 36: 4.4.2 Roc Curve - ShuffleNetV2

Figure 36: 4.4.2 Roc Curve - ShuffleNetV2 illustrates that ROC curve of ShuffleNetV2 model for tea region classification based on tea liquor color. It represents the trade-off between true positive rate (TPR) and false positive rate (FPR) across different decision thresholds.

It has obtained mean Area Under the Curve (AUC) value is 0.88. It has excellent discriminative capability of the proposed model. The ROC curve remains consistently above the diagonal line which indicates that model achieved high true positive rate while maintaining low false positive rates for all seven tea region samples. These results have confirmed that liquor color provides strong classification among different tea growing regions.

4.4.3. MobileNetV2

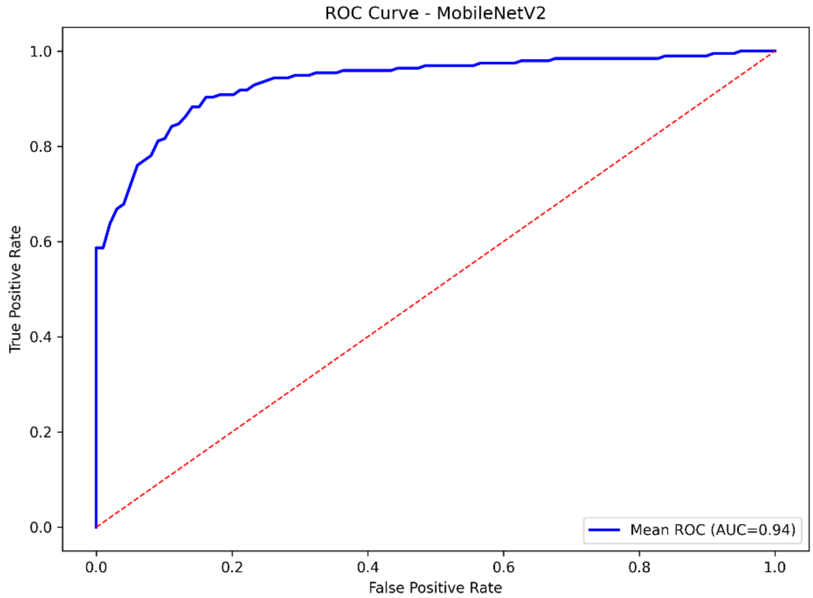

Figure 37: 4.4.3 Roc Curve - MobileNetV2

Figure 37: 4.4.3 Roc Curve - MobileNetV2 illustrates that ROC curve of MobileNetV2 model for tea region classification based on tea liquor color. It represents the trade-off between true positive rate (TPR) and false positive rate (FPR) across different decision thresholds.

It has obtained mean Area Under the Curve (AUC) value is 0.94. It has excellent discriminative capability of the proposed model. The ROC curve remains consistently above the diagonal line which indicates that model achieved high true positive rate while maintaining low false positive rates for all seven tea region samples. These results have confirmed that liquor color provides strong classification among different tea growing regions.

4.4.4. EfficientNetb0

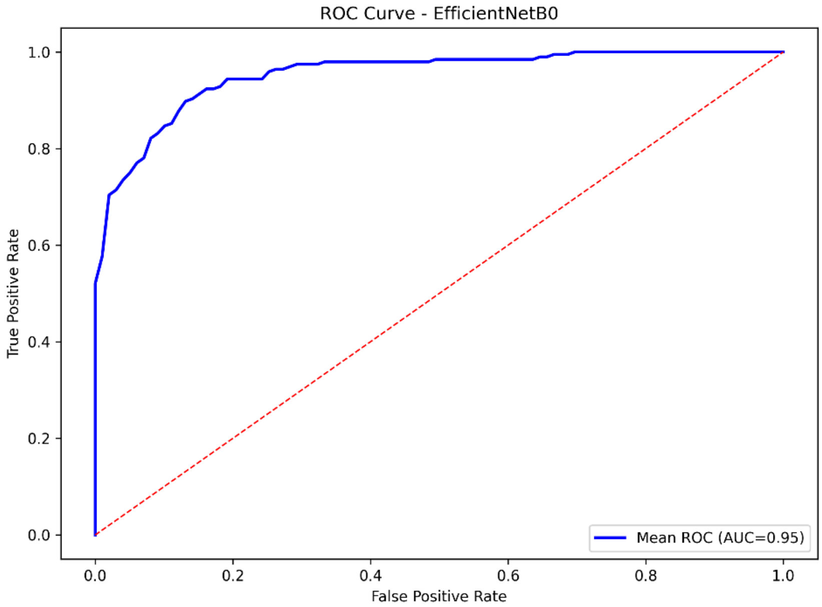

Figure 38: 4.4.4 Roc Curve - EfficientNetb0

Figure 38: 4.4.4 Roc Curve - EfficientNetb0 illustrates that ROC curve of EfficientNetb0 model for tea region classification based on tea liquor color. It represents the trade-off between true positive rate (TPR) and false positive rate (FPR) across different decision thresholds.

It has obtained mean Area Under the Curve (AUC) value is 0.95. It has excellent discriminative capability of the proposed model. The ROC curve remains consistently above the diagonal line which indicates that model achieved high true positive rate while maintaining low false positive rates for all seven tea region samples. These results have confirmed that liquor color provides strong classification among different tea growing regions.

4.4.5. SqueezeNet

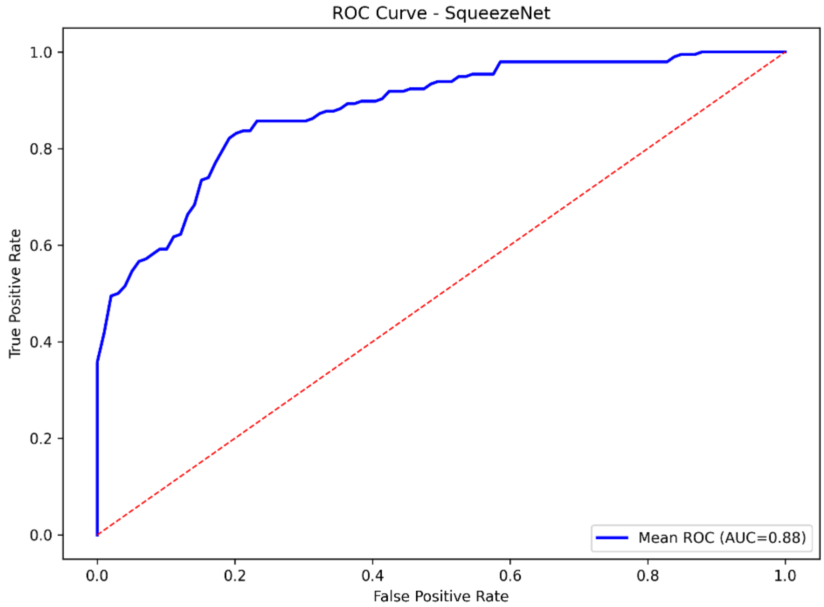

Figure 39: 4.4.5 Roc Curve - SqueezeNet

Figure 39: 4.4.5 Roc Curve – SqueezeNet illustrates that ROC curve of SqueezeNet model for tea region classification based on tea liquor color. It represents the trade-off between true positive rate (TPR) and false positive rate (FPR) across different decision thresholds.

It has obtained mean Area Under the Curve (AUC) value is 0.88. It has excellent discriminative capability of the proposed model. The ROC curve remains consistently above the diagonal line which indicates that model achieved high true positive rate while maintaining low false positive rates for all seven tea region samples. These results have confirmed that liquor color provides strong classification among different tea growing regions.

4.4.6. SVM

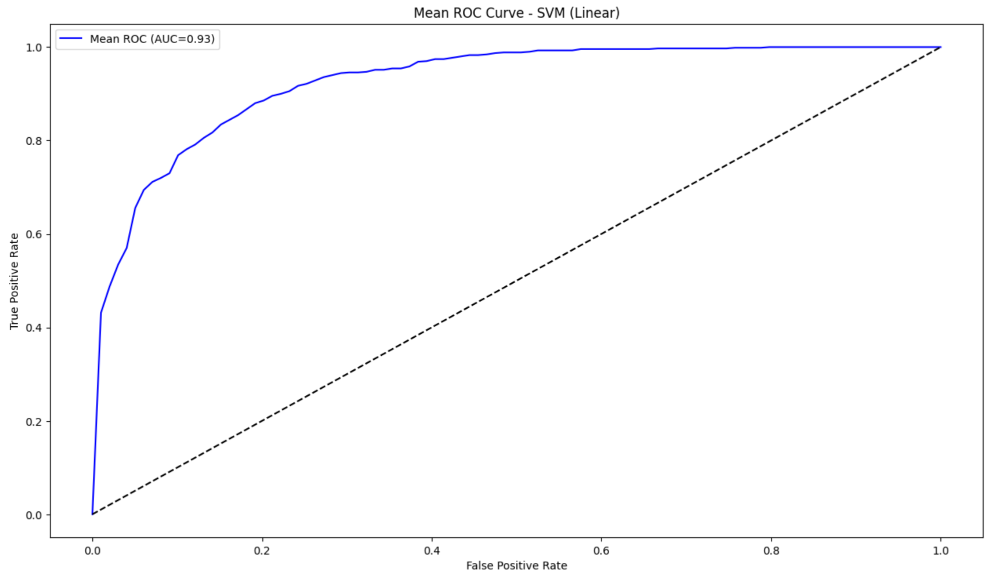

Figure 40: 4.4.6 Roc Curve - SVM

Figure 40: 4.4.6 Roc Curve – SVM illustrates that ROC curve of SVM model for tea region classification based on tea liquor color. It represents the trade-off between true positive rate (TPR) and false positive rate (FPR) across different decision thresholds.

It has obtained mean Area Under the Curve (AUC) value is 0.93. It has excellent discriminative capability of the proposed model. The ROC curve remains consistently above the diagonal line which indicates that model achieved high true positive rate while maintaining low false positive rates for all seven tea region samples. These results have confirmed that liquor color provides strong classification among different tea growing regions.

4.4.7. Random Forest

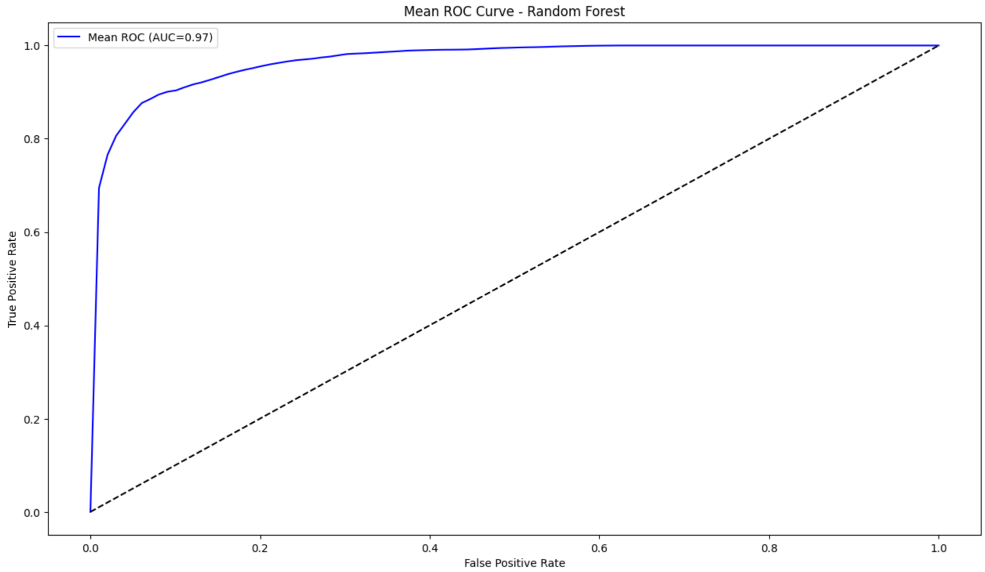

Figure 41: 4.4.7 Roc Curve - Random Forest

Figure 41: 4.4.7 Roc Curve - Random Forest illustrates that ROC curve of Random Forest model for tea region classification based on tea liquor color. It represents the trade-off between true positive rate (TPR) and false positive rate (FPR) across different decision thresholds.

It has obtained mean Area Under the Curve (AUC) value is 0.97. It has excellent discriminative capability of the proposed model. The ROC curve remains consistently above the diagonal line which indicates that model achieved high true positive rate while maintaining low false positive rates for all seven tea region samples. These results have confirmed that liquor color provides strong classification among different tea growing regions.

4.4.8. KNN

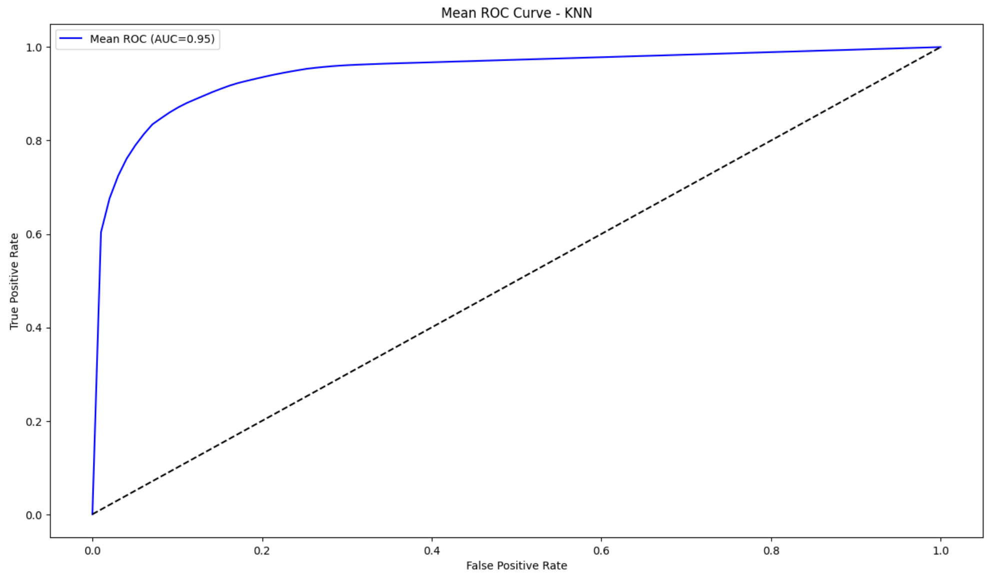

Figure 42: 4.4.8 Roc Curve - KNN

Figure 42: 4.4.8 Roc Curve – KNN illustrates that ROC curve of KNN model for tea region classification based on tea liquor color. It represents the trade-off between true positive rate (TPR) and false positive rate (FPR) across different decision thresholds.

It has obtained mean Area Under the Curve (AUC) value is 0.95. It has excellent discriminative capability of the proposed model. The ROC curve remains consistently above the diagonal line which indicates that model achieved high true positive rate while maintaining low false positive rates for all seven tea region samples. These results have confirmed that liquor color provides strong classification among different tea growing regions.

4.4.9. Logistic Regression

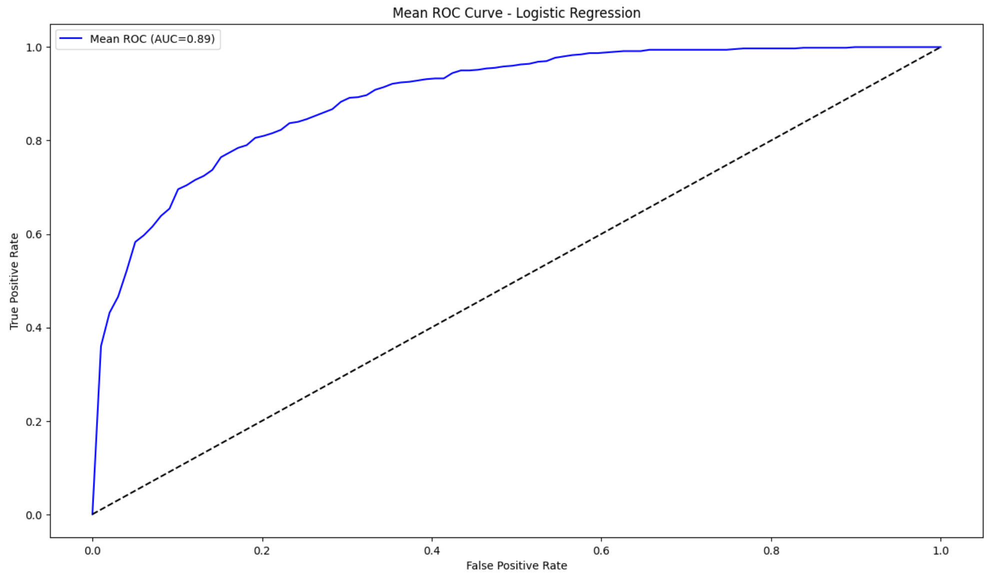

Figure 43: 4.4.9 Roc Curve - Logistic Regression

Figure 43: 4.4.9 Roc Curve - Logistic Regression illustrates that ROC curve of Logistic Regression model for tea region classification based on tea liquor color. It represents the trade-off between true positive rate (TPR) and false positive rate (FPR) across different decision thresholds.

It has obtained mean Area Under the Curve (AUC) value is 0.89. It has excellent discriminative capability of the proposed model. The ROC curve remains consistently above the diagonal line which indicates that model achieved high true positive rate while maintaining low false positive rates for all seven tea region samples. These results have confirmed that liquor color provides strong classification among different tea growing regions.

4.5. Principal Component Analysis

The Principal Component Analysis (PCA) represents the dimensionality reduction and visualization technique which analyze feature separability in learned representation space. The PCA shows well-separated clusters corresponding to different tea regions for the TEAQNET models. It has indicated strong discriminated feature learning. The CNN models such as RestNet-18, ShuffleNetV2, MobileNetV2, EfficientNetb0 and SqueezeNet are used for model comparisons among deep learning model architectures.

4.5.1. RestNet-18

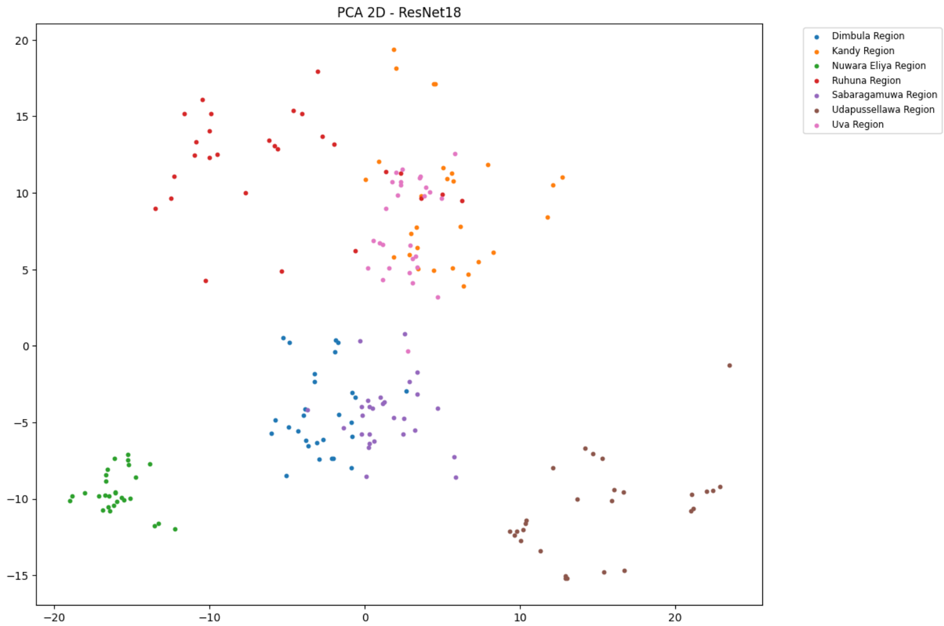

Figure 44: 4.5.1 PCA - RestNet-18

Figure 44: 4.5.1 PCA - RestNet-18 illustrates that 2D Principal Component Analysis of color features extracted from RestNet-18 model for tea region classification based on tea liquor color.

The PCA plot has demonstrated clear clustering patterns corresponding to seven tea regions of Sri Lanka. There are 196 tea samples and those are well separated into clusters. It has indicated strong intra-class similarity and effective feature discrimination across seven regions. There is average overlap observed between a few regions such as Dimbula , Kandy, Sabaragamuwa and Uva. Other regions are clearly classified.

The overall spatial separation among clusters represents that RestNet-18 model successfully captured the discriminative chromatic and textural features from tea liquor images. The PCA visualization has provided further evidence of robustness and effectiveness of tea liquor color which is reliable for feature extraction for automated tea region classification.

4.5.2. ShuffleNetV2

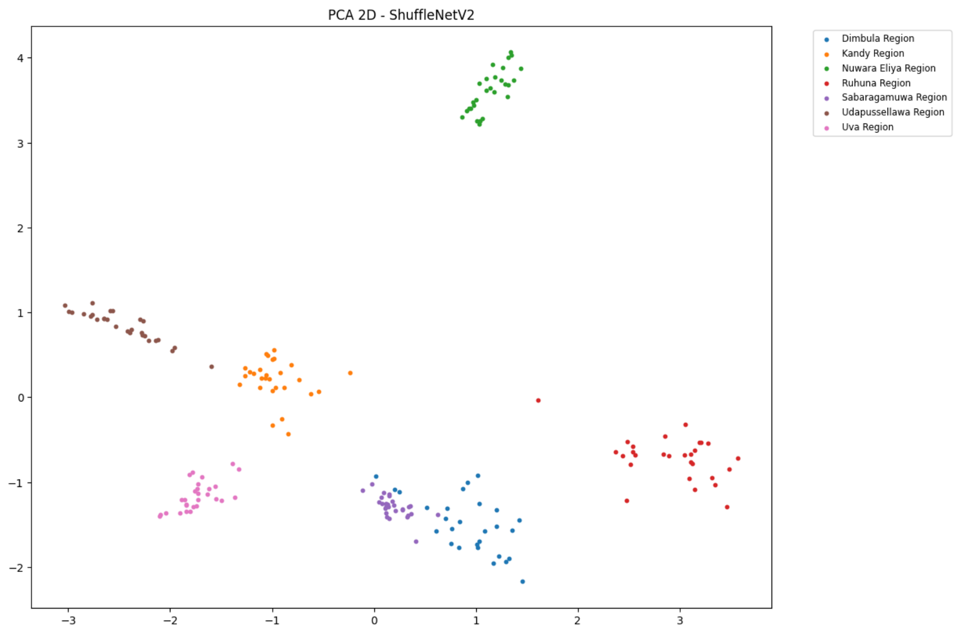

Figure 45: 4.5.2 PCA - ShuffleNetV2

Figure 45: 4.5.2 PCA - ShuffleNetV2 illustrates that 2D Principal Component Analysis of color features extracted from ShuffleNetV2 model for tea region classification based on tea liquor color.

The PCA plot has demonstrated clear clustering patterns corresponding to seven tea regions of Sri Lanka. There are 196 tea samples and those are well separated into clusters. It has indicated strong intra-class similarity and effective feature discrimination across seven regions. There is limited overlap observed between a few regions such as Dimbula and Sabaragamuwa. Other regions are clearly classified.

The overall spatial separation among clusters represents that ShuffleNetV2 model successfully captured the discriminative chromatic and textural features from tea liquor images. The PCA visualization has provided further evidence of robustness and effectiveness of tea liquor color which is reliable for feature extraction for automated tea region classification.

4.5.3. MobilenetV2

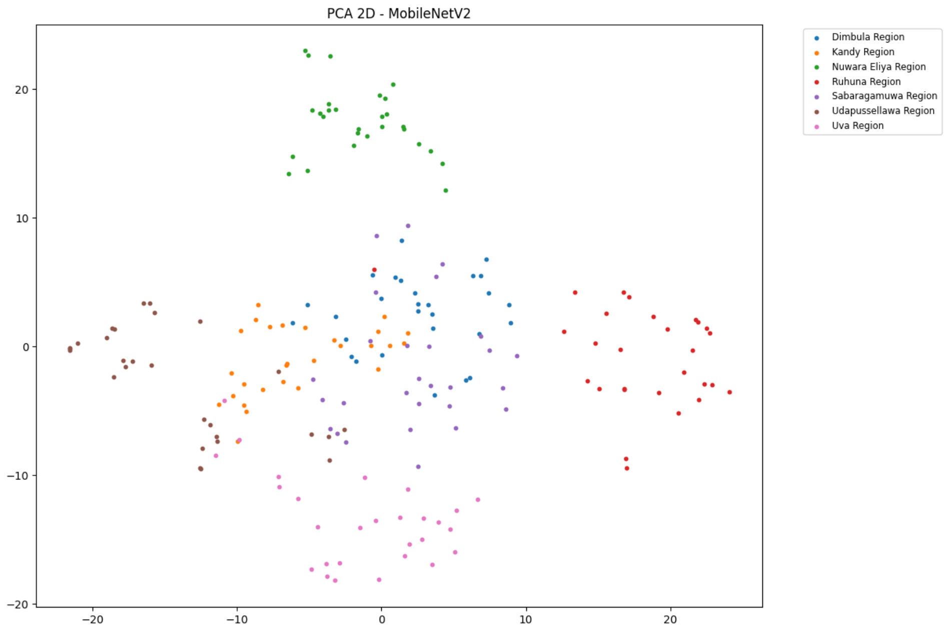

Figure 46: 4.5.3 PCA - MobileNetV2

Figure 46: 4.5.3 PCA - MobileNetV2 illustrates that 2D Principal Component Analysis of color features extracted from MobileNetV2 model for tea region classification based on tea liquor color.

The PCA plot has demonstrated clear clustering patterns corresponding to seven tea regions of Sri Lanka. There are 196 tea samples and those are well separated into clusters. It has indicated strong intra-class similarity and effective feature discrimination across seven regions. There is huge overlap observed between a few regions such as Dimbula , Kandy and Sabaragamuwa. Among other regions Nuwara Eliya and Ruhuna are clearly classified.

The overall spatial separation among clusters represents that MobileNetV2 model successfully captured the discriminative chromatic and textural features from tea liquor images. The PCA visualization has provided further evidence of robustness and effectiveness of tea liquor color which is reliable for feature extraction for automated tea region classification.

4.5.4. EfficientNetb0

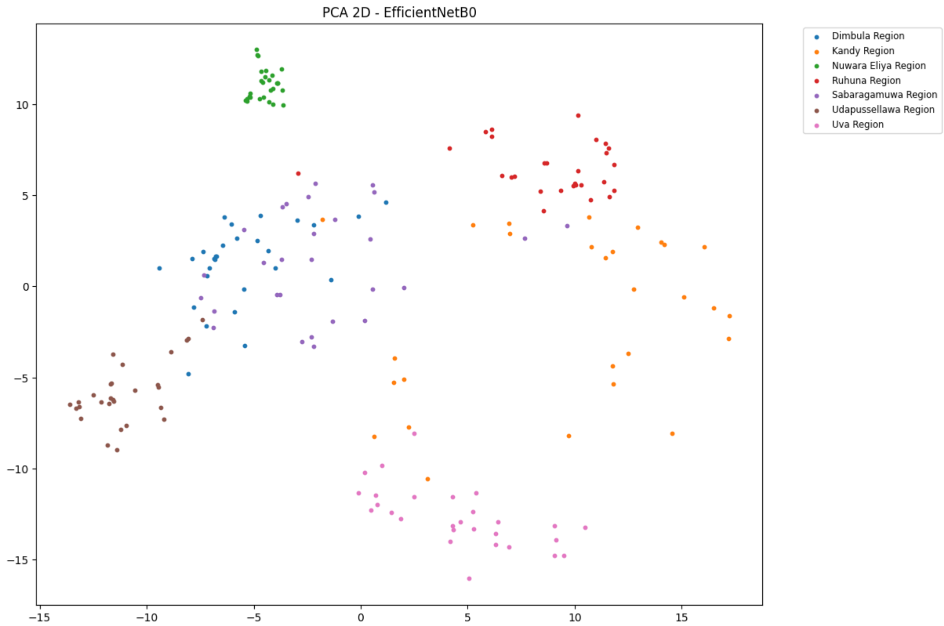

Figure 47: 4.5.4 PCA - Efficientnetb0

Figure 47: 4.5.4 PCA - Efficientnetb0 illustrates that 2D Principal Component Analysis of color features extracted from Efficientnetb0 model for tea region classification based on tea liquor color.

The PCA plot has demonstrated clear clustering patterns corresponding to seven tea regions of Sri Lanka. There are 196 tea samples and those are well separated into clusters. It has indicated strong intra-class similarity and effective feature discrimination across seven regions. There is limited overlap observed between a few regions such as Dimbula and Sabaragamuwa.

The overall spatial separation among clusters represents that Efficientnetb0 model successfully captured the discriminative chromatic and textural features from tea liquor images. The PCA visualization has provided further evidence of robustness and effectiveness of tea liquor color which is reliable for feature extraction for automated tea region classification.

4.5.5. SqueezeNet

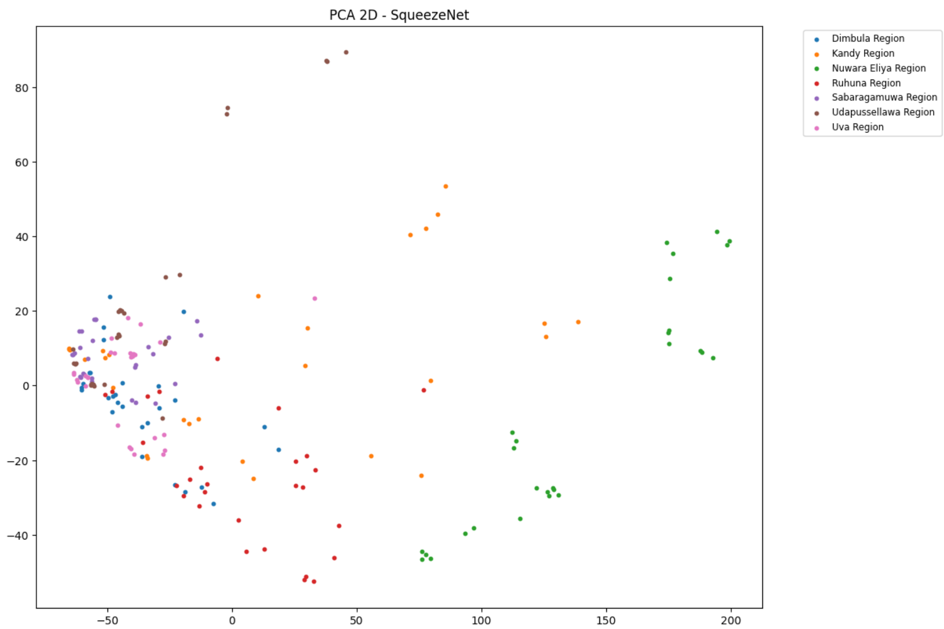

Figure 48: 4.5.5 PCA - SqueezeNet

Figure 48: 4.5.5 PCA – SqueezeNet illustrates that 2D Principal Component Analysis of color features extracted from SqueezeNet model for tea region classification based on tea liquor color.

The PCA plot has demonstrated clear clustering patterns corresponding to seven tea regions of Sri Lanka. There are 196 tea samples and those are not correctly separated into clusters except samples of Nuwara Eliya Region. It has indicated very week intra-class similarity and not effective feature discrimination across seven regions. There are huge overlaps observed between all regions except Nuwara Eliya.

The overall spatial separation among clusters represents that SqueezeNet model not fully-captured the discriminative chromatic and textural features from tea liquor images. The PCA visualization has provided further evidence of robustness and effectiveness of tea liquor color which is reliable for feature extraction for automated tea region classification.

4.5.6. SVM

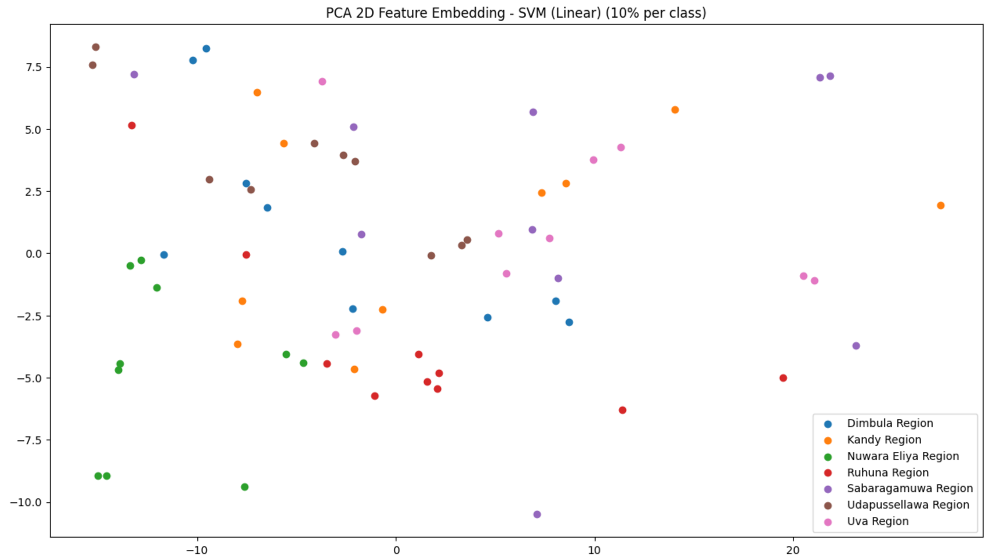

Figure 49: 4.5.6 PCA - SVM

Figure 49: 4.5.6 PCA – SVM illustrates that 2D Principal Component Analysis of color features extracted from SVM model for tea region classification based on tea liquor color.

The PCA plot has demonstrated clustering patterns corresponding to seven tea regions of Sri Lanka. There are 70 tea samples and those are on average separated into clusters. It has indicated week intra-class similarity and not effective feature discrimination across seven regions. There are huge overlaps observed between all regions.

The overall spatial separation among clusters represents that SVM model did not capture the discriminative chromatic and textural features from tea liquor images. The PCA visualization has provided further evidence of robustness and effectiveness of tea liquor color which is reliable for feature extraction for automated tea region classification.

4.5.7. Random Forest

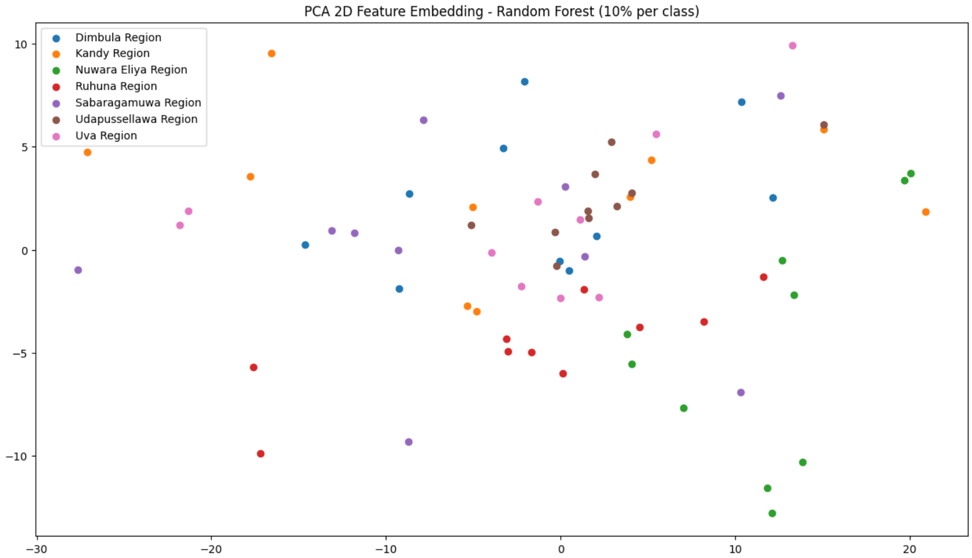

Figure 50: 4.5.7 PCA - Random Forest

Figure 50: 4.5.7 PCA - Random Forest illustrates that 2D Principal Component Analysis of color features extracted from Random Forest model for tea region classification based on tea liquor color.

The PCA plot has demonstrated clustering patterns corresponding to seven tea regions of Sri Lanka. There are 70 tea samples and those are on average separated into clusters. It has indicated week intra-class similarity and not effective feature discrimination across seven regions. There are huge overlaps observed between all regions.

The overall spatial separation among clusters represents that Random Forest model did not capture the discriminative chromatic and textural features from tea liquor images. The PCA visualization has provided further evidence of robustness and effectiveness of tea liquor color which is reliable for feature extraction for automated tea region classification.

4.5.8. KNN

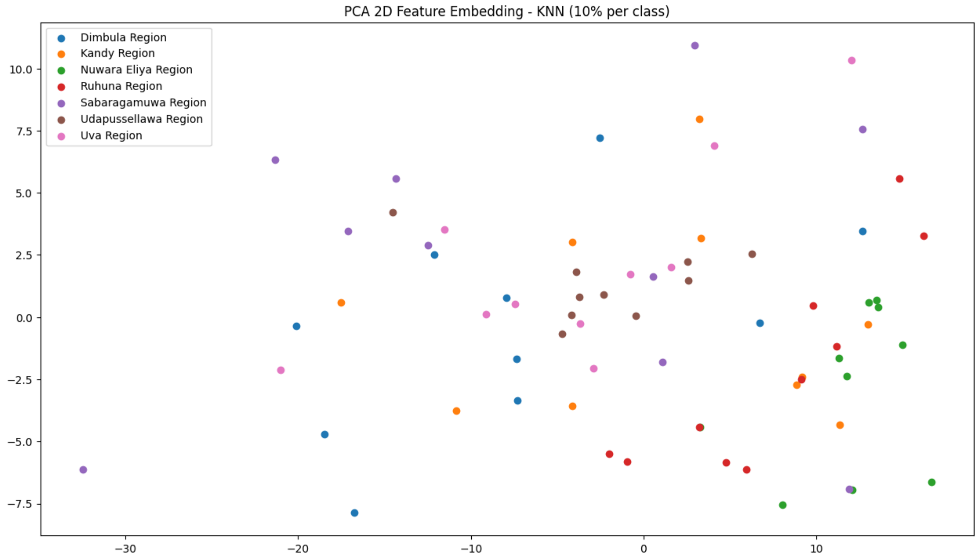

Figure 51: 4.5.8 PCA - KNN

Figure 51: 4.5.8 PCA – KNN illustrates that 2D Principal Component Analysis of color features extracted from KNN model for tea region classification based on tea liquor color.

The PCA plot has demonstrated clustering patterns corresponding to seven tea regions of Sri Lanka. There are 70 tea samples and those are on average separated into clusters. It has indicated week intra-class similarity and not effective feature discrimination across seven regions. There are huge overlaps observed between all regions.

The overall spatial separation among clusters represents that KNN model did not capture the discriminative chromatic and textural features from tea liquor images. The PCA visualization has provided further evidence of robustness and effectiveness of tea liquor color which is reliable for feature extraction for automated tea region classification.

4.5.9. Logistic Regression

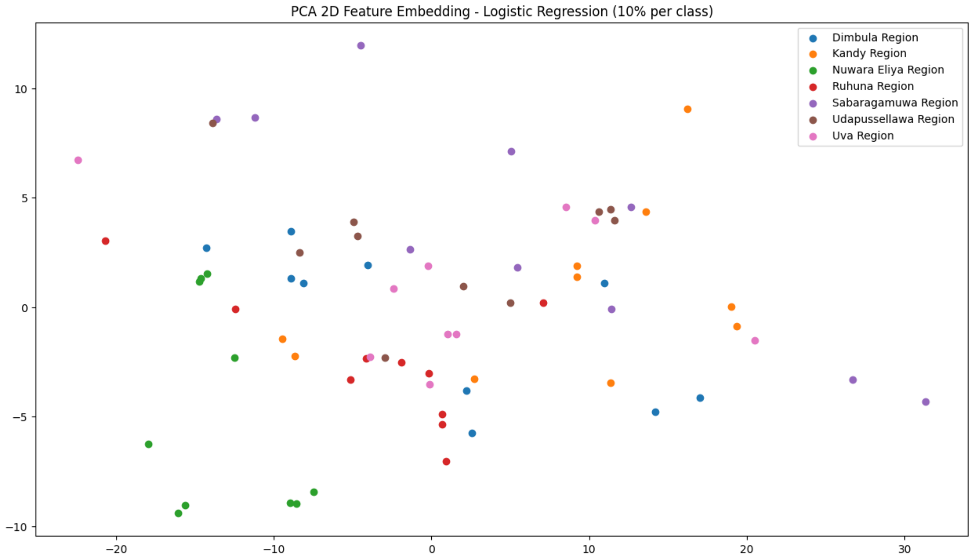

Figure 52: 4.5.9 PCA - Logistic Regression

Figure 52: 4.5.9 PCA - Logistic Regression illustrates that 2D Principal Component Analysis of color features extracted from Logistic Regression model for tea region classification based on tea liquor color.

The PCA plot has demonstrated clustering patterns corresponding to seven tea regions of Sri Lanka. There are 70 tea samples and those not correctly separated into clusters. It has indicated very week intra-class similarity and not effective feature discrimination across seven regions. There are huge overlaps observed between all regions.

The overall spatial separation among clusters represents that Logistic Regression model did not capture the discriminative chromatic and textural features from tea liquor images. The PCA visualization has provided further evidence of robustness and effectiveness of tea liquor color which is reliable for feature extraction for automated tea region classification.

4.6. Evaluation Metrics

The final evaluation is conducted using quantitative performance metrics such as accuracy, precision, recall, F1-score, and Intersection over Union (IOU). The Deep Learning models have achieved superior performance across all metrics compared to the traditional Machine Learning models. The CNN models such as RestNet-18, ShuffleNetV2, MobileNetV2, EfficientNetb0 and SqueezeNet are used for model comparisons among deep learning model architectures.

Table 21: 4.6 Evaluation Metrics-Deep Learning

Model / Metrics	Accuracy	Precision	Recall	F1-score	IOU
RestNet-18	0.9898	0.9901	0.9898	0.9898	0.9799
EfficientNetb0	0.9700	0.9700	0.9700	0.9700	0.9417
MobileNetV2	0.9949	0.9951	0.9949	0.9949	0.9900
ShuffleNetV2	0.9999	0.9999	0.9999	0.9999	0.9998
SqueezeNet	0.6327	0.7277	0.6327	0.5977	0.5115

Traditional models such as SVM, Random Forest, KNN and Logistic regression are used for model comparisons among machine learning model architectures.

Table 22: 4.6 Evaluation Metrics-Machine Learning

Model / Metrics	Accuracy	Precision	Recall	F1-score	IOU
SVM	0.6642	0.6745	0.6642	0.6621	0.5199
Random forest	0.8114	0.8272	0.8114	0.8046	0.6838
KNN	0.7814	0.7839	0.7814	0.7769	0.6513
Logistic Regression	0.6100	0.6156	0.6100	0.6062	0.4647

4.7. Chapter Summary

This chapter has represented quantitative and qualitative evaluation of the proposed TEAQNET based region classification framework. The experiment analysis has conducted through multiple deep learning and classical machine learning models get the classification metrics such as accuracy, precision, recall and F1-score. In this chapter has demonstrated about four types of plots such as learning curves, confusion matrix, ROC curve and PCA

The learning curve analysis has represented deep learning architectures for CNN models such as RestNet-18, ShuffleNetV2, MobileNetV2, EfficientNetb0 and SqueezeNet which convergence with minimal overfitting effective learning behavior and strong generalization.

There were techniques used such as early stopping to ensure optimal models are selected by preventing unnecessary training beyond convergence. Classical machine learning models represented comparatively weaker convergence patterns.

Confusion matrix analysis represented that CNN based models have achieved perfect classification across seven tea-growing regions. ShuffleNetV2 demonstrated the highest classification performance compared to other models. Traditional machine learning classifiers such as SVM, Random Forest, KNN and Logistic regression have higher misclassification rate.

ROC Curve analysis has validated the discriminative strength of the proposed approach. All CNN models have achieved high mean for AUC (Area under the Curve) values which greater than 0.88. It represents the trade-off between true positive rates across multiclass classification tasks. Among machine learning models Random Forest and KNN yield comparatively moderate discriminative capability.

The principal component analysis (PCA) has provided visual confirmation of feature separability across different images which belong to different regions. The CNN based models have defined well-defined compact clusters for tea regions. But traditional machine learning models show significant cluster overlaps. It has reflected weak feature discrimination. These findings have proved deep learning effectively captures the chromatic and textural attributes which are inherent to the liquor image clusters.

Finally, those models are evaluated using matrixes such accuracy, precision, recall, IOU and F1-score. It has represented the clear advantage of deep learning models. ShuffleNetV2 achieved the highest overall performance across all metrics. Classical machine learning models have achieved underperformed in comparison which emphasize their limitations and high dimensional visual classification tasks.

5. Chapter 05 – Discussion

5.1. Chapter Introduction

This chapter presents the discussion of experimental findings which are obtained from proposed tea region classification framework. The objective of this chapter is to interpret the results of deep learning models and machine learning models. It has emphasized on how tea liquor color features contribute to distinguishing tea samples based on tea regions in Sri Lanka. The outcomes are analyzed related to the research objectives, methodological choices and existing literature in tea quality assessment and region classification.

The discussion focused on identifying the effectiveness of the standardized brewing protocol (ISO 3103) use in the controlled imaging environment. The image preprocessing techniques have minimized subjectivity and environmental variability. The classification architecture of CNN has examined with different evaluation matrices such a accuracy, precision, recall and F1-score to identify the best deep learning model. Comparative insights are provided with relating strong observed results to traditional sensory evaluation methods and previously reported machine vision and chemometric approaches for tea region classification.

This chapter has identified the practical implications of adopting liquor color based automated classification system for quality control, origin verification and decision making in tea industry. The strength and limitations of this study have correctly identified. Those challenges are lighting constraint, sample preparation consistency and scalability to real-world deployments. Finally, the discussion has identified key insights of experimental results, recommendations and future research directions for tea region classification.

5.2. Findings

The study has investigated that tea region classification in Sri Lanka using tea liquor color. The findings have represented that tea liquor color contains discriminative visual information which is capable of distinguishing tea regions. The images are captured under standardized conditions and analyzed using computational techniques.

Experimental results are presented that proposed image preprocessing and classification pipeline achieved robust regional discrimination across the seven major tea regions. The analysis of confusion matrixes has shown that high classification accuracy, precision and specifying the effectiveness of color bases features for regional classification. Among seven tea regions, Nuwara Eliya, Uva and Dimbula have achieved strong classification performance due to their distinct agro climatic conditions which influence the pigment composition and infusion color.

The key finding of learning curve analysis is Deep learning models are better suited for capturing subtle, region-specific liquor color variations compared to classical Machine learning approaches. The analysis of learning curves has revealed clear differences between classical Machine Learning models and Deep Learning models. The ML models like SVM, Random Forest, KNN has shown early convergence with training and validation accuracies which stabilize at lower performance levels. It's indicated that limited representational capacity while handling complex, non-linear color variations in tea liquor images across regions.

Convolutional Neural Network based models have demonstrated that progressive improvement in validation accuracy with increasing training epochs. It has been accomplished by a gradual reduction in training validation loss gap. There exist minor overfittings which are observed in deeper CNN configurations that mitigate through normalization.

The key finding of confusion matrix analysis is CNN models significantly reduce the inter-region confusion which is especially for visually distinctive regions such as Nuwara Eliya. The confusion matrix analysis has provided detailed inside into class-wise performance across seven tea regions in Sri Lanka. It has identified through classification accuracy through each model in deep learning and machine learning approaches. This classified number of samples can be measured through diagonal entries of confusion matrixes.

Table 23: Model Accuracy Based on Correct Predictions.

Model	**Model Type**	**No. of Samples used**	**No. of Samples Classified**	**Accuracy (percentage %)**
ShuffleNetV2	DL	196	196	100
MobileNetV2	DL	196	195	99.48
RestNet-18	DL	196	194	98.97
EfficientNetb0	DL	196	190	96.93
Random Forest	ML	700	568	81.14

KNN	ML	700	547	78.14
SVM	ML	700	465	66.42
SqueezeNet	DL	196	127	63.26
Logistic Regression	ML	700	427	61.00

According to the confusion matrixes, ShuffleNetV2 model has achieved highest accuracy. It has classified all the samples correctly. As a result, Deep Learning models have achieved better accuracy than the Machine Learning models except SqueezeNet model. The SqueezeNet model has 63.26% accuracy, but it is less than the accuracy of SVM, Random Forest and KNN models.

The key findings of ROC curve analysis are high Area Under the Curve (AUC) values represented that deep learning models provide reliable and stable discrimination for tea region classification based on liquor color. It further validated superiority of deep learning approaches. Classical ML models have shown moderate AUC values and reflecting limited discriminative ability in multi-class classification. The CNN based models have achieved higher AUC values across all regions. It has indicated that strong sensitivity and specificity tradeoffs. The ROC curves for Nuwara Eliya and Uva regions were particularly steep and demonstrating excellent classification in learned feature space.

The key findings of Principal Component Analysis are clear PCA clustering exist for Nuwara Eliya region across all models. It confirms the discriminative strength of deep learning features derived from liquor color images. PCA visualization of extracted features has highlighted a critical distinction between Machine learning and Deep Learning models. It has applied handcrafted color features which has resulted in overlapping clusters in mid elevation region. The PCA revealed well-defined clustering. As an example, Nuwara Region has formed clear isolated cluster in different DL and ML models.

The experimental findings have been demonstrated that tea liquor color contains sufficient discriminative information to classify Sri Lankan tea regions based on tea liquor color in standardized environment. The Deep Learning approaches such ass CNN has successfully overcome the limitations by automatically learning hierarchical color and texture representations. The consistent separability of Nuwara Eliya tea across different plots such as confusion matrixes, ROC curves and PCA plots further validates both dataset quality and methodological rigor of proposed system.

5.3. Comparison with literature

The findings of the research are consistent with yet extend beyond the existing research on tea classification and quality assessment reported in the literature. The previous studies of tea evaluation have relied on chemical profiling and spectroscopic analysis. Traditional sensory based tea liquor color evaluation is documented in the literature. It is quantitative and heavily depends on expert tasters. The prior studies have identified the limitations in reproducibility and objectivity due to observer subjectivity and environmental variations. The purposed TEAQNET framework directly addresses these limitations by quantifying liquor color under ISO 3103 tea brewing standards. It has improved consistency and repeatability compared with manual sensory assessment.

Machine Learning models in literature has shown moderate accuracy and their effectiveness depends strongly on manual feature engineering and controlled experiment settings. Deep learning-based approaches such as CNN used in this study, learn automatically discriminative spatial color features from raw liquor images. It has resulted in higher overall accuracy, improved class separability and better generalization evidenced by learning curves, ROC analysis, and confusion matrices.

The PCA visualization reported in this study revel clear and compact clustering for Nuwara Eliya region. It has highlighted strong discriminative power of liquor color characteristics for High-elevation of Sri Lankan teas. The observation supports tea science literature that associates lighter, brighter liquor color with high-grown tea regions. It has demonstrated that such characteristics can be effectively captured and modeled using Deep Learning Models.

The existing studies on tea liquor analysis have rarely focused on real time classification or deployable systems. The integration of CNN based classification with standardized imaging environment has enabled web based and mobile based applications. The previous studies heavily focused on chemical composition-based tea evaluation. It's heavily dependent on polyphenols of tea such as theaflavins and thearubigins. It is needed in a lab environment for identification of tea regions. This purposed TEAQNET framework is robust, cost-effective and anyone can use without domain knowledge of tea classification.

5.4. Novel contributions

This research presents several novel scientific and technical contributions to the domain of tea region classification based on tea liquor color. In this study, first comprehensive attempt to classify Sri Lankan tea regions such as Dimbula, Ruhuna, Sabaragamuwa, Udapussellawa, Uva, Kandy and Nuwara Eliya with using tea liquor color and Deep Learning techniques. The prior studies focused on chemical composition, trace elements, aroma profiling and leaf-level imaging. This research has identified liquor color as the primary discriminative feature for tea region classification.

This study provides comparative evaluation of classical Machine Learning and Deep Learning models for tea region classification. Prior studies evaluated a single model for tea evaluation, but this study extended it into broader range of models to get the best results. This study demonstrated that CNN-based models significantly outperformed ML models and DL models captures the subtle color-texture interactions in tea liquor more efficiently. This research has contributed to the digital transformation of traditional tea evaluation which replaces subjective human judgment with objective, repeatable and scalable AI-based assessment. This aligns with global trends in smart agriculture industry which positioning Sri Lankan tea research with modern computational paradigm.

5.5. Limitations

The proposed tea region classification framework has several limitations. It is essential for identifying these limitations for correct interpretation of results and for guiding future research directions.

The proposed system highly depends on controlled imaging conditions which include lightbox, fixed camera distance, controlled exposure and white balance. The model performance has depended on environmental conditions. Those conditions are illumination intensity, color temperature, background reflectance and camera sensitivity.

The study is restricted to tea liquor samples prepared under ISO 3103 brewing conditions. This standard ensures reproducibility and fairness in comparison of tea regions. It doesn't account for variations in real-world brewing practices such as differences in water quality, infusion time and room temperature.

The scope of applicability is confined to tea color analysis only. It does not depend on other sensory and chemical attributes such as aroma, taste, polyphenol composition and trace element profiles. The dataset size is also limited. There were 700 tea liquor images which collected around seven tea regions of Sri Lanka. It is limited when compared to large-scale industrial datasets. The proposed system is currently limited for image acquisition camera. The image quality is highly dependent on the camera quality, Lense and resolution.

5.6. Industrial implications

The findings of this study have several practical and industrial implications for tea sector. The proposed classification system based on liquor color and it offers a rapid, non-destructive and cost-effective alternative to traditional sensory evaluation which conducted by expert tea tasters. In large scale manufacturing environments, human sensory evaluation is time consuming, subjective and difficult to scale. But automated image-based classification systems have enabled consistent and reproducible evaluation and reduce the dependence on limited human expertise. It also minimizes the inter and intra-observer variation.

The integration of standardized brewing conditions such as ISO 3103 and controlled imaging environment has marked proposed approach suitable for industrial deployment. The system can be implemented in tea factories, research laboratories and quality assurance units. The effectiveness of deep learning models highlights their potential for real time quality monitoring during production.

The proposed system supports the digital transformation initiatives in tea industry. It has enabled integration with web-based platforms and mobile applications. It can be extended to tea brokers, exporters and even consumers for preliminary quality verification. This promotes data-driven decision making across the tea value chain. It has also reduced the manual labour involved in tea classification. As summary the proposed tea region classification system demonstrates strong industrial applicability which offering scalable, standardized and intelligent solution for modern tea region classification.

5.7. Chapter Summary

This chapter has demonstrated the discussion of proposed tea classification framework for tea region classification based on tea liquor color. It has demonstrated that deep learning approaches such as CNN architecture which consistently outperformed classical ML models in terms of classification accuracy, robustness, and generalization ability. Learning curve analysis represented stable convergence and reduced overfitting when standardized imaging conditions and appropriate data augmentation strategies were applied. The confusion matrix and ROC curve analyses further highlighted the superior discriminate capability of CNN models for distinguishing visually similar tea regions. Principal Component Analysis represents the meaningful clustering patterns among tea regions. Among tea regions, Nuwara Eliya has exhibited clear and distinct separability which indicating strong regional signatures in tea liquor characteristics. This observation has aligned with known variations in altitude, climate and biochemical composition of teas which are produced in tea regions.

The findings were contextualized with prior research which demonstrating consistency with earlier studies have focused on spectroscopy, elemental profiling and sensory analysis for origin authentication. The limitations of this study related to the controlled lighting and camera dependency are clearly addressed in this chapter. Finally, the industrial relevance of the proposed system was discussed. It has highlighted applicability in quality assurance, authentication and decision support within the tea industry.

6. Chapter 06 - Conclusion & Future Works

6.1. Chapter Introduction

The chapter concludes the study of tea region classification based on tea liquor color analysis. It summarizes the key outcomes of proposed approach. It highlights the effectiveness of developed methodology and reflects the exceptional findings. In addition, this chapter outlines the future research directions which aimed at further improving model accuracy, robustness and applicability in real world tea industry settings.

6.2. Contributions

This research has contributed to computer vision based agricultural product classification which focuses on tea region classification in academic, methodological and practical fields. The proposed work resolves the key limitations in traditional tea evaluation methods. It also gives advances the application of deep learning techniques within the industry domain.

The one of primary contribution of this research is development of image-based framework for automated tea region classification using tea liquor color. Traditional tea evaluation heavily relies on expert sensory evaluation, chemical analysis and laboratory-based techniques. Those are often subjective, time consuming and costly. This research represented the visual cues which derive from tea liquor images. It served as a reliable and non-destructive alternative for regional classification.

This research contributed to the comprehensive analysis of multiple state of art in deep learning models. The experiments have been done with various CNN models such as Restnet-18, Efficientnetb0, MobileNetV2, ShufflenetV2 and SqueezeNet for tea region classification. These CNN architectures have studied generic image classification tasks and achieved exceptional performance on image classifications.

This research has contributed to creating effective application using transfer learning techniques to overcome data scarcity challenges commonly occurred in agricultural datasets. It has initialized models with pre-trained weight from large-scale datasets like ImageNet. This research has improved convergence speed, enhanced the feature generalization and reduced the overfitting.

6.3. Scientific impact

This research has contributed to computer vision, machine learning and agricultural informatics by representing the feasibility and effectiveness of tea region classification based on tea liquor color. This research has bridged domain knowledge from tea science with modern deep learning methodologies. This research has opened new directions for non-destructive agricultural product evaluation.

The one of key scientific impact is validating visual color information as a discriminative feature for regional authentication. Traditional tea classification techniques have depended on chemical composition, spectral analysis or expert evaluation. This study has provided evidence that fine-grained chromatic variations in tea liquor images can be learned by CNN models.

This research has used multiple evaluation metrics for evaluating different deep learning and machine learning models. This multi-metric classification approach has contributed to the model evaluation for fine-grained multi class classification problems.

6.4. Industrial value

This research has provided substantial industrial value by providing a practical, automated and scalable solution for tea region classification. The proposed architecture addresses several key challenges faced by tea industry. Those are subjectivity in tea region assessment, high dependence on tea tasters and cost and time associated with traditional laboratory-based evaluation methods.

The automated region classification is one of the primary benefits of this research. The accurate identification of tea region is essential for several things. Those are maintaining product consistency, protecting geographic indications and ensuring compliance with export standards.

This research has lightweight model architecture which enhances the industrial applicability in resource constraint environment. The proposed system has reduced the expenses of chemical analysis equipment and specialized laboratory personnel. The proposed system relies on mobile devices. Therefore, everyone can use it without any domain knowledge about tea region classification. It also enhanced the trust among exporters, regulators and consumers. It has contributed to the reputation and regional identification of Sri Lankan tea in international markets.

This research has contributed to the decision support and process optimization. The classification output can guide managers and quality controllers in identifying errors in processing conditions. The proposed system has bridged the gap between research innovation and industrial adaptation which contributes to digital transformation of tea industry.

6.5. Future research directions

This research has represented the effectiveness of tea region classification based on tea liquor color analysis. Future direction of this research is to enhance model accuracy, robustness and applicability. These directions are focused on current framework by incorporating richer data representations, advanced sensing techniques and broader classification objectives.

6.5.1. Multi-modal fusion

The future directions are opened for multi model fusion where visual information from tea liquor images are combined with additional complementary data sources. It can integrate tea liquor images with before brewing tea images (dry tea particles) and after brewing tea images (spent tea particles). The appearance of dry tea particles related to the processing styles and leaf grade. It reveals the structural and color changes influenced by regional growing conditions.

6.5.2. Spectral imaging

Another future direction is adaptation of spectral and hyperspectral imaging conditions. It captures the reflectance information across wider range of wavelength compared to the RGB imaging. This integration improves classification performance under varying lightning conditions and consistent region classification.

6.5.3. Larger dataset

The performance and generalization of deep learning models are strongly influenced by dataset size and diversity. The future work should focus on constructing larger-scale datasets. It includes tea sample collected across multiple seasons, different processing batches and geographic sub regions. A more diverse dataset can reduce model bias. It improves robustness to environmental variation using different sources which increase strengthen of model credibility and scientific validity.

6.6. Chapter Summary

This chapter represented the summarization of key findings, contributions and implications of proposed tea region classification. It has revisited the research objectives and demonstrated how methodology successfully developed. It has also addressed the limitations of traditional tea evaluation techniques through application of computer vision and deep learning.

The chapter represented the major contributions to image-based classification framework and experimental evaluation against both state of the art deep learning models and traditional machine learning classifiers. It has discussed the terms of advancing color-based feature learning, data-efficient model training and domain specific neural network design within agriculture and food quality assessment research.

This chapter also summarizes the industrial value of proposed system which emphasized and potential to support automated quality control and region authentication. This chapter also discussed the future directions of this research which aimed to extend the current work through multi-model fusion, spectral imaging and larger datasets. This chapter has provided a comprehensive conclusion for the industrial adoption of intelligent tea classification systems.

7. Chapter 07 - References

[1] K. G. N. P. Piyasena and L. S. K. Hettiarachchi, “Comparison of tea quality parameters of conventionally and organically grown tea, and effects of fertilizer on tea quality: A mini-review,” *Food Chem. Adv.*, vol. 3, p. 100399, Dec. 2023, doi: 10.1016/j.focha.2023.100399.

[2] S. Jin *et al.*, “Study on the correlation between color and taste of beauty tea infusion and the pivotal contributing compounds based on UV–visible spectroscopy, taste equivalent quantification and metabolite analysis,” *Food Chem. X*, vol. 21, p. 101192, Mar. 2024, doi: 10.1016/j.fochx.2024.101192.

[3] J. Zhou *et al.*, “Analysis of changes in flavor characteristics of congou black tea at different fermentation degrees under industrial production conditions using flavor compound weighted network co-expression method,” *Food Chem.*, vol. 468, p. 142241, 2025.

[4] L. Zhu, P. Spachos, E. Pensini, and K. N. Plataniotis, “Deep learning and machine vision for food processing: A survey,” *Curr. Res. Food Sci.*, vol. 4, pp. 233–249, 2021.

[5] F. Mendoza, P. Dejmek, and J. M. Aguilera, “Calibrated color measurements of agricultural foods using image analysis,” *Postharvest Biol. Technol.*, vol. 41, no. 3, pp. 285–295, 2006.

[6] Unsplash, “Herbal Tea Pictures | Download Free Images on Unsplash.” Accessed: Nov. 26, 2025. [Online]. Available: https://unsplash.com/s/photos/herbal-tea

[7] A. Valavanidis, “Tea, the Most Popular Beverage Worldwide, is Beneficial to Human Health. Studies on antioxidant polyphenolic constituents and epidemiological evidence for disease prevention,” vol. 1, pp. 1–35, June 2019.

[8] “Ceylon Tea Online Store,” Ceylon Tea Online Store. Accessed: Nov. 26, 2025. [Online]. Available: https://ceylontea.home.blog/

[9] “Tea industry overview January to December 2016 | Daily FT.” Accessed: Nov. 26, 2025. [Online]. Available: https://www.ft.lk/agriculture/tea-industry-overview-january-to-december-2016/31-611422

[10] S. C. Thushara, “Sri Lankan Tea Industry: Prospects and Challenges,” 2015.

[11] Y. Alahakoon, M. Peiris, and N. Gunarathne, “Sustainability Challenges and the Way Forward in the Tea Industry: The Case of Sri Lanka,” in *Case Studies on Sustainability in the Food Industry: Dealing With a Rapidly Growing Population*, S. O. Idowu and R. Schmidpeter, Eds., Cham: Springer International Publishing, 2022, pp. 271–292. doi: 10.1007/978-3-031-07742-5_11.

[12] "A Journey Through the Tea Fields of Sri Lanka | Alveus." Accessed: Nov. 26, 2025. [Online]. Available: https://www.alveus.eu/blog/tea-sri-lanka

[13] "saea.lk/wp-content/uploads/2023/12/SJAE_2023_01_Final.pdf." Accessed: Nov. 25, 2025. [Online]. Available: https://saea.lk/wp-content/uploads/2023/12/SJAE_2023_01_Final.pdf

[14] J. Yun *et al.*, "Use of headspace GC/MS combined with chemometric analysis to identify the geographic origins of black tea," *Food Chem.*, vol. 360, p. 130033, Oct. 2021, doi: 10.1016/j.foodchem.2021.130033.

[15] "Oral Halitosis (Bad Breath) and its management.pptx." Accessed: Nov. 26, 2025. [Online]. Available: https://www.slideshare.net/slideshow/oral-halitosis-bad-breath-and-its-management-pptx/280500298

[16] "cdn.standards.iteh.ai/samples/73224/79469b30ce154460b38ebb10af5565cf/ISO-3103-2019.pdf." Accessed: Nov. 25, 2025. [Online]. Available: https://cdn.standards.iteh.ai/samples/73224/79469b30ce154460b38ebb10af5565cf/ISO-3103-2019.pdf

[17] J. Bordim, C. C. Lise, C. Marques, T. C. Oldoni, P. Varela, and M. L. Mitterer-Daltoé, "Potential use of naturally colored antioxidants in the food industry—A study of consumers' perception and acceptance," *J. Sens. Stud.*, vol. 36, no. 4, p. e12657, 2021, doi: 10.1111/joss.12657.

[18] "Cover | Chemical Senses | Oxford Academic." Accessed: Nov. 25, 2025. [Online]. Available: https://academic.oup.com/chemse/article-abstract/42/1/NP/2667793?redirectedFrom=fulltext&login=false

[19] "'#Munnar Munnar is a town in the Western Ghats mountain range in India's Kerala state. A hill station and former resort for the British Raj elite, it's surrounded by rolling hills dotted with tea plantations established in the late 19th century.' - Picture of RG Holidays, Rameswaram - Tripadvisor." Accessed: Nov. 26, 2025. [Online]. Available: https://www.tripadvisor.com/LocationPhotoDirectLink-g303894-d15952680-i376601640-RG_Holidays-Rameswaram_Ramanathapuram_District_Tamil_Nadu.html

[20] "Tea Research Institute of Sri Lanka." Accessed: Nov. 25, 2025. [Online]. Available: https://www.tri.lk/

[21] "A tea picking robot is solving labour shortage crisis in the industry," masterstalk.online. Accessed: Nov. 26, 2025. [Online]. Available: https://masterstalk.online/2024/04/08/a-tea-picking-robot-is-solving-labour-shortage-crisis-in-the-industry/

[22] “The chemistry of tea flavonoids: Critical Reviews in Food Science and Nutrition: Vol 37, No 8.” Accessed: Dec. 02, 2025. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/10408399709527797

[23] “Black Tea Theraubigins - ProQuest.” Accessed: Dec. 02, 2025. [Online]. Available: https://www.proquest.com/openview/44dec6b28d0e040c9927a3a6bc574a6e/1?pq-origsite=gscholar&cbl=2026366&diss=y

[24] T. Gebely, “What is Withering in Tea Processing? · Tea Epicure,” Tea Epicure. Accessed: Dec. 02, 2025. [Online]. Available: https://teaepicure.com/tea-leaves-wither/

[25] “HARVESTING IMPROVING TEA QUALITY THROUGH IOT IN FERMENTATION PROCESS.” Accessed: Dec. 02, 2025. [Online]. Available: https://senzagro.com/blogs/harvesting-improving-tea-quality/

[26] Z. Feng *et al.*, “Tea aroma formation from six model manufacturing processes,” *Food Chem.*, vol. 285, pp. 347–354, July 2019, doi: 10.1016/j.foodchem.2019.01.174.

[27] Y. Zhang, X. Chen, D. Chen, L. Zhu, G. Wang, and Z. Chen, “Machine learning-based classification and prediction of typical Chinese green tea taste profiles,” *Food Res. Int.*, vol. 203, p. 115796, Feb. 2025, doi: 10.1016/j.foodres.2025.115796.

[28] L. Jin *et al.*, “Antioxidant properties and color parameters of herbal teas in China,” *Ind. Crops Prod.*, vol. 87, pp. 198–209, Sept. 2016, doi: 10.1016/j.indcrop.2016.04.044.

[29] Manikharda, V. E. Shofi, B. K. Betari, and Supriyadi, “Effect shading intensity on color, chemical composition, and sensory evaluation of green tea (*Camelia sinensis* var Assamica),” *J. Saudi Soc. Agric. Sci.*, vol. 22, no. 7, pp. 407–412, Oct. 2023, doi: 10.1016/j.jssas.2023.03.006.

[30] K.-W. Ng, Z.-J. Cao, H.-B. Chen, Z.-Z. Zhao, L. Zhu, and T. Yi, “Oolong tea: A critical review of processing methods, chemical composition, health effects, and risk,” *Crit. Rev. Food Sci. Nutr.*, vol. 58, no. 17, pp. 2957–2980, Nov. 2018, doi: 10.1080/10408398.2017.1347556.

[31] “Effect shading intensity on color, chemical composition, and sensory evaluation of green tea (Camelia sinensis var Assamica),” *J. Saudi Soc. Agric. Sci.*, vol. 22, no. 7, pp. 407–412, Oct. 2023, doi: 10.1016/j.jssas.2023.03.006.

[32] Y. Wang *et al.*, “Tea Analyzer: A low-cost and portable tool for quality quantification of postharvest fresh tea leaves,” *LWT*, vol. 159, p. 113248, Apr. 2022, doi: 10.1016/j.lwt.2022.113248.

[33] X. Zhai, L. Zhang, M. Granvogl, C.-T. Ho, and X. Wan, “Flavor of tea (Camellia sinensis): A review on odorants and analytical techniques,” *Compr. Rev. Food Sci. Food Saf.*, vol. 21, no. 5, pp. 3867–3909, 2022, doi: 10.1111/1541-4337.12999.

[34] C. Dong *et al.*, “Prediction of Congou Black Tea Fermentation Quality Indices from Color Features Using Non-Linear Regression Methods,” *Sci. Rep.*, vol. 8, no. 1, p. 10535, July 2018, doi: 10.1038/s41598-018-28767-2.

[35] A. Finger, S. Kuhr, and U. H. Engelhardt, “Chromatography of tea constituents,” *J. Chromatogr. A*, vol. 624, no. 1, pp. 293–315, Oct. 1992, doi: 10.1016/0021-9673(92)85685-M.

[36] A. Berna, “Metal Oxide Sensors for Electronic Noses and Their Application to Food Analysis,” *Sensors*, vol. 10, no. 4, pp. 3882–3910, Apr. 2010, doi: 10.3390/s100403882.

[37] Q. Chen, J. Zhao, and S. Vittayapadung, “Identification of the green tea grade level using electronic tongue and pattern recognition,” *Food Res. Int.*, vol. 41, no. 5, pp. 500–504, Jan. 2008, doi: 10.1016/j.foodres.2008.03.005.

[38] Z. Bojie, W. Dong, S. Weizhong, L. Yu, and W. Ke, “Research on tea bud identification technology based on HSI/HSV color transformation,” presented at the 2019 6th International Conference on Information Science and Control Engineering (ICISCE), IEEE Computer Society, Dec. 2019, pp. 511–515. doi: 10.1109/ICISCE48695.2019.00108.

[39] D. Liang, Q. Zhou, C. Ling, L. Gao, X. Mu, and Z. Liao, “Research progress on the application of hyperspectral imaging techniques in tea science,” *J. Chemom.*, vol. 37, no. 6, p. e3481, 2023, doi: 10.1002/cem.3481.

[40] C. Cui *et al.*, “Accurate and rapid quantification of tea saponins using a thin-layer chromatography (TLC) method based on hemolysis and machine vision,” *LWT*, vol. 199, p. 116139, May 2024, doi: 10.1016/j.lwt.2024.116139.

[41] “1114_col_foodsafetyquality.pdf.” Accessed: Dec. 07, 2025. [Online]. Available: https://www.ift.org/~/media/food%20technology/pdf/2014/11/1114_col_foodsafetyquality.pdf

[42] H. Wang, J. Gu, and M. Wang, “A review on the application of computer vision and machine learning in the tea industry,” *Front. Sustain. Food Syst.*, vol. 7, Apr. 2023, doi: 10.3389/fsufs.2023.1172543.

[43] T. Wu *et al.*, “Applications of deep learning in tea quality monitoring: a review,” *Artif. Intell. Rev.*, vol. 58, no. 11, p. 342, Aug. 2025, doi: 10.1007/s10462-025-11335-2.

[44] T. Tang, Q. Luo, L. Yang, C. Gao, C. Ling, and W. Wu, “Research Review on Quality Detection of Fresh Tea Leaves Based on Spectral Technology,” *Foods*, vol. 13, no. 1, p. 25, Jan. 2024, doi: 10.3390/foods13010025.

[45] A. Moreda-Piñeiro, A. Fisher, and S. J. Hill, “The classification of tea according to region of origin using pattern recognition techniques and trace metal data,” *J. Food Compos. Anal.*, vol. 16, no. 2, pp. 195–211, Apr. 2003, doi: 10.1016/S0889-1575(02)00163-1.
[46] “PROFESSIONAL TEA CUPPING -.” Accessed: Dec. 13, 2025. [Online]. Available: https://authory.com/dianarosen/PROFESSIONAL-TEA-CUPPING-add313dc2160e421187c8ff89cae0b788
[47] “ISO-3103-2019.pdf.” Accessed: Dec. 13, 2025. [Online]. Available: https://cdn.standards.iteh.ai/samples/73224/79469b30ce154460b38ebb10af5565cf/ISO-3103-2019.pdf
[48] “Introduction to Convolution Neural Network - GeeksforGeeks.” Accessed: Dec. 13, 2025. [Online]. Available: https://www.geeksforgeeks.org/machine-learning/introduction-convolution-neural-network/
[49] “ResNet18 from Scratch Using PyTorch - GeeksforGeeks.” Accessed: Dec. 13, 2025. [Online]. Available: https://www.geeksforgeeks.org/deep-learning/resnet18-from-scratch-using-pytorch/
[50] “Efficientnet Architecture,” GeeksforGeeks. Accessed: Dec. 13, 2025. [Online]. Available: https://www.geeksforgeeks.org/computer-vision/efficientnet-architecture/
[51] “Deep Learning Architecture Review: DenseNet, ResNeXt, MnasNet & ShuffleNet v2 | DigitalOcean.” Accessed: Dec. 13, 2025. [Online]. Available: https://www.digitalocean.com/community/tutorials/popular-deep-learning-architectures-densenet-mnasnet-shufflenet
[52] “Deep Learning Architectures Explained: ResNet, InceptionV3, SqueezeNet | DigitalOcean.” Accessed: Dec. 13, 2025. [Online]. Available: https://www.digitalocean.com/community/tutorials/popular-deep-learning-architectures-resnet-inceptionv3-squeezenet
[53] “Loss Functions in Deep Learning - GeeksforGeeks.” Accessed: Dec. 13, 2025. [Online]. Available: https://www.geeksforgeeks.org/deep-learning/loss-functions-in-deep-learning/
[54] “Evaluation Metrics in Machine Learning,” GeeksforGeeks. Accessed: Dec. 13, 2025. [Online]. Available: https://www.geeksforgeeks.org/machine-learning/metrics-for-machine-learning-model/
[55] F. Alvi, “PyTorch vs TensorFlow in 2025: A Comparative Guide of AI Frameworks,” OpenCV. Accessed: Dec. 13, 2025. [Online]. Available: https://opencv.org/blog/pytorch-vs-tensorflow/

8. APPENDIXES

8.1. Preprocessed Data Samples

Table 24: 8.1 Preprocessed Data sample

Region	Samples
Dimbula	 *Figure 53: 8.1 Liquor Sample-Dimbula*
Nuwara Eliya	 *Figure 54: 8.1 Liquor Sample-Nuwara Eliya*
Kandy	 *Figure 55: 8.1 Liquor Sample-Kandy*
Sabaragamuwa	 *Figure 56: 8.1 Liquor Sample-Sabaragamuwa*

Udapussellawa	 *Figure 57: 8.1 Liquor Sample-Udapussellawa*
Ruhuna	 *Figure 58: 8.1 Liquor Sample-Ruhuna*
Uva	 *Figure 59: 8.1 Liquor Sample-Uva*

8.2. Deep Learning Models

8.2.1. ResNet-18

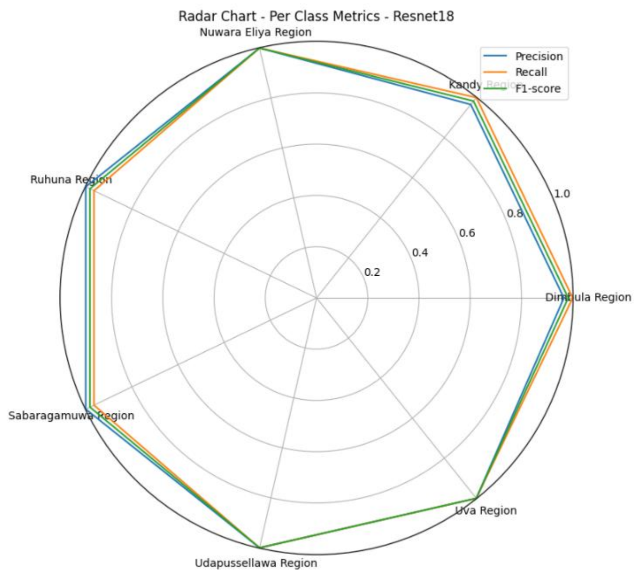

Figure 60: 8.2.1 Radar Graph - RestNet-18

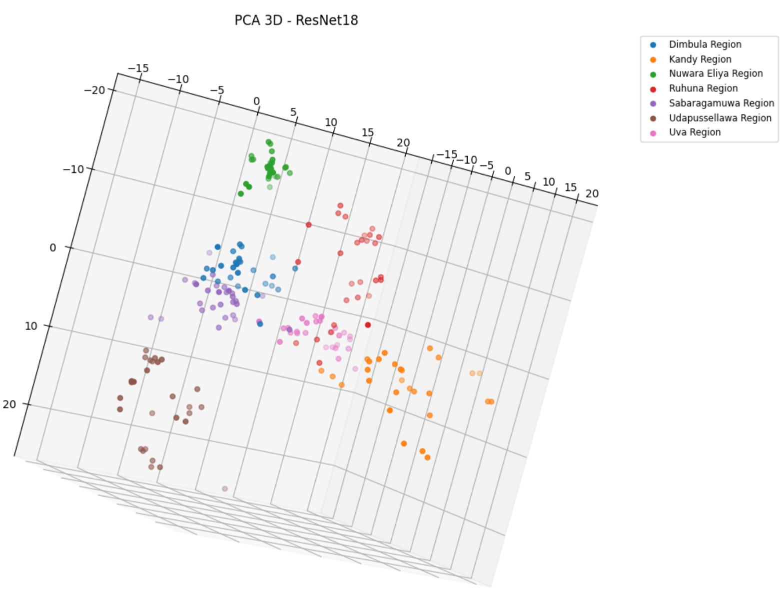

Figure 61: 8.2.1 PCA 3D -RestNet-18

8.2.2. ShuffleNetV2

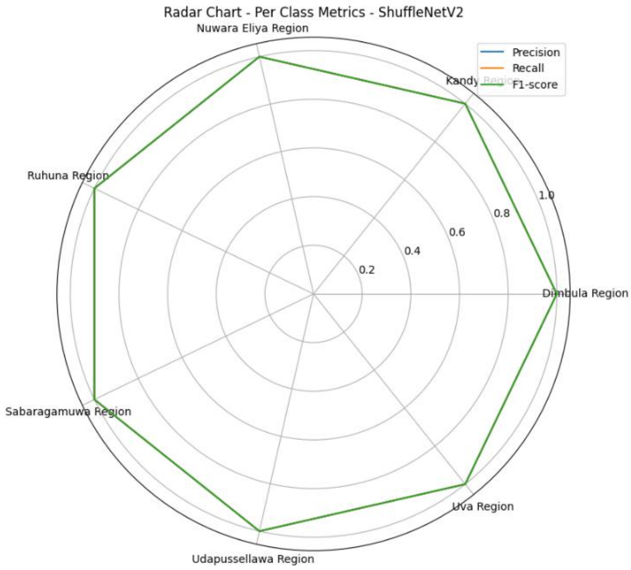

Figure 62: 8.2.2 Radar Graph - ShuffleNetV2

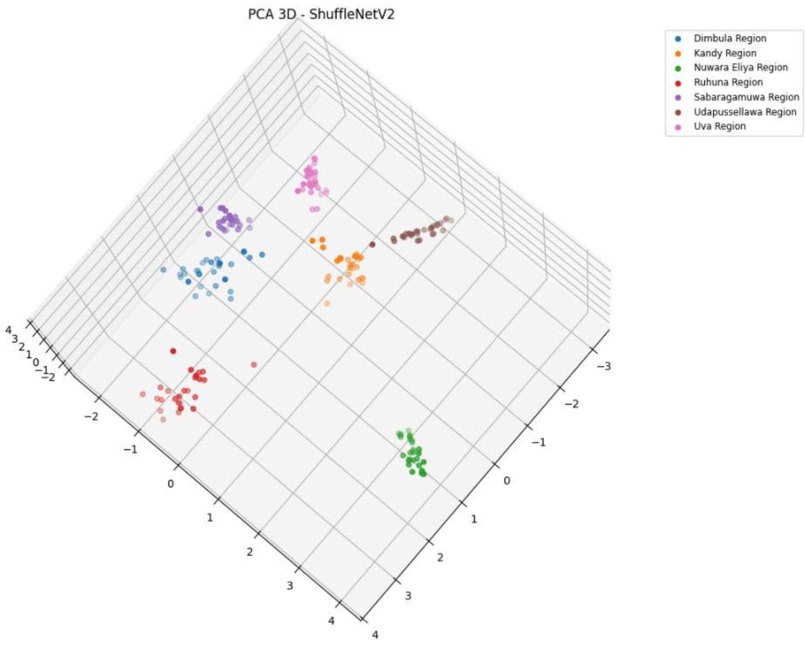

Figure 63: 8.2.2 PCA 3D - ShuffleNetV2

8.2.3. EfficientNetb0

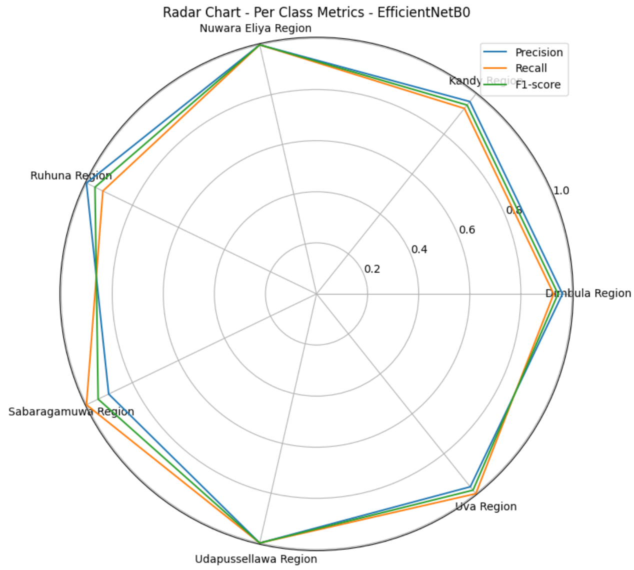

Figure 64: 8.2.3 Radar Graph - EfficientNetb0

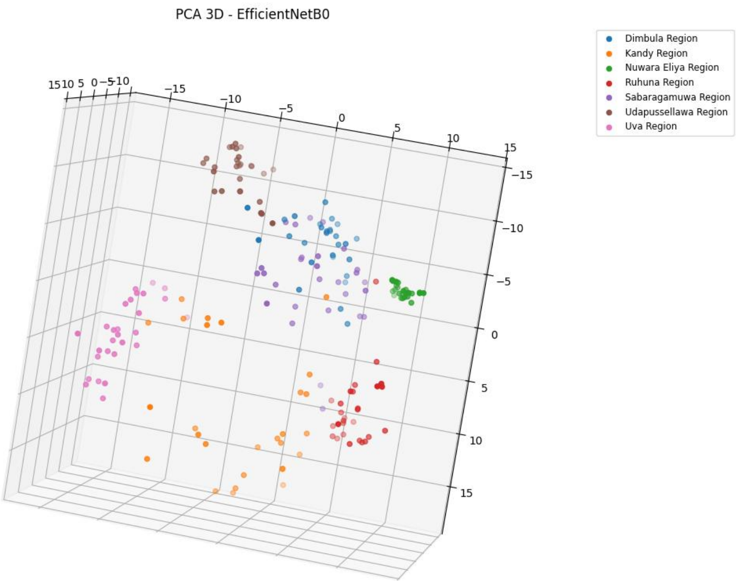

Figure 65: 8.2.3 PCA 3D - EfficientNetb0

8.2.4. MobileNetV2

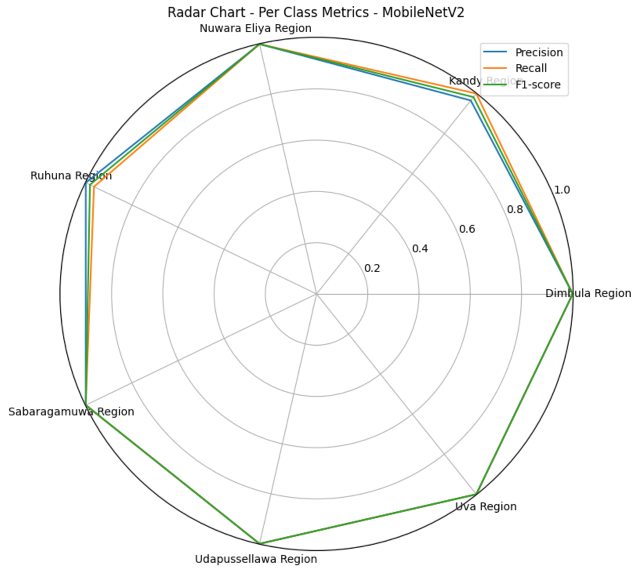

Figure 66: 8.2.4 Radar Graph - MobileNetv2

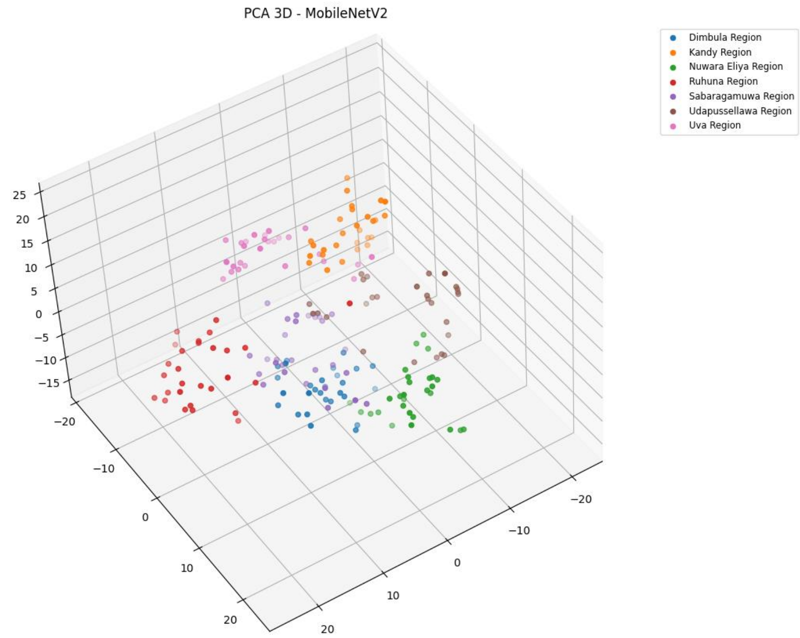

Figure 67: 8.2.4 PCA 3D - MobileNetV2

8.2.5. SqueezeNet

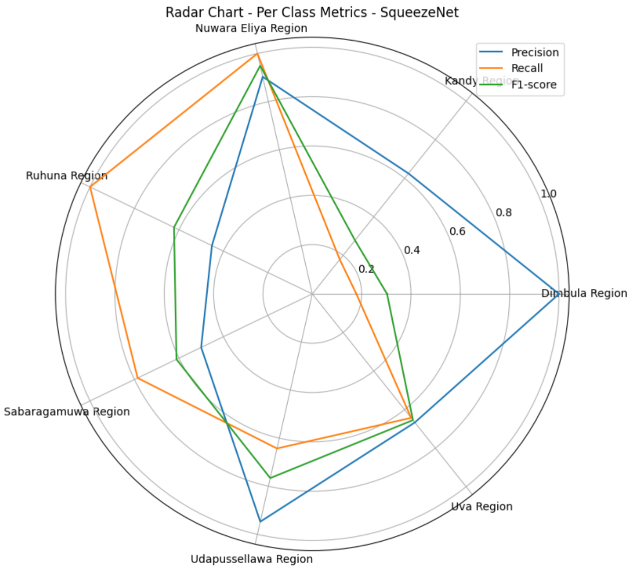

Figure 68: 8.2.5 Radar graph – SqueezeNet

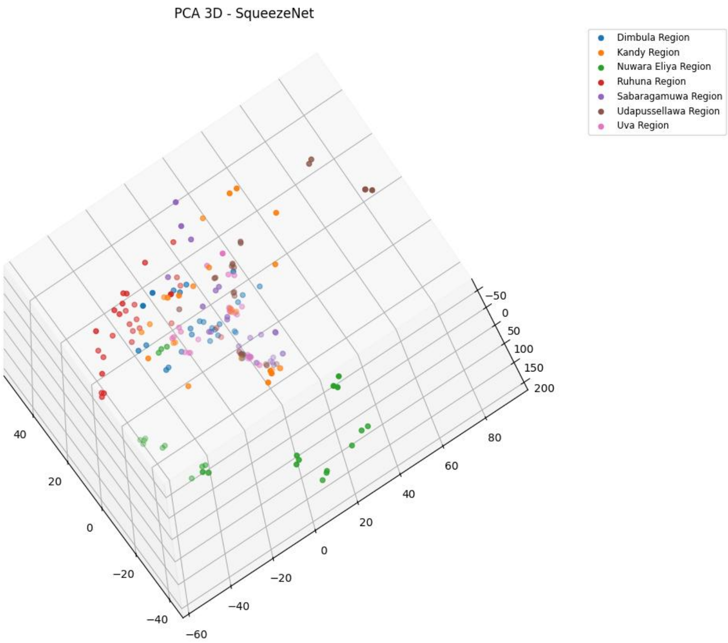

Figure 69: 8.2.5 PCA 3D - SqueezeNet

8.2.6. Custom CNN

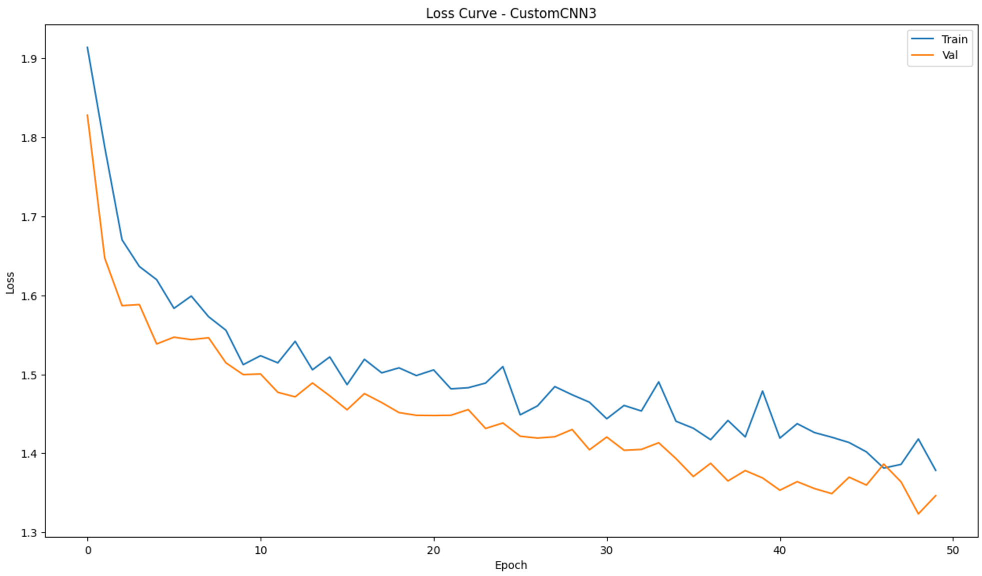

Figure 70: 8.2.6 Accuracy Curve - Custom CNN

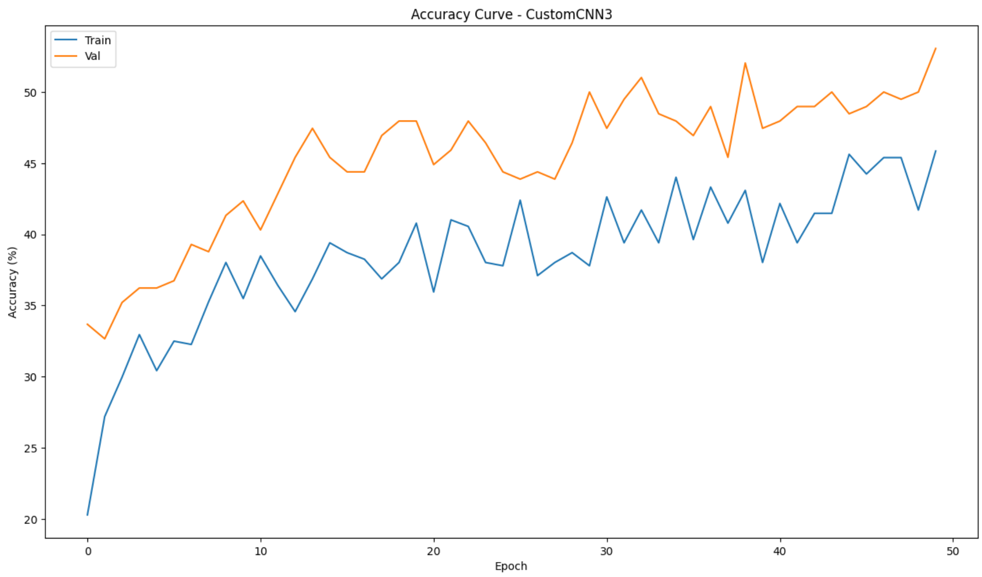

Figure 71: 8.2.6 Loss Curve - Custom CNN

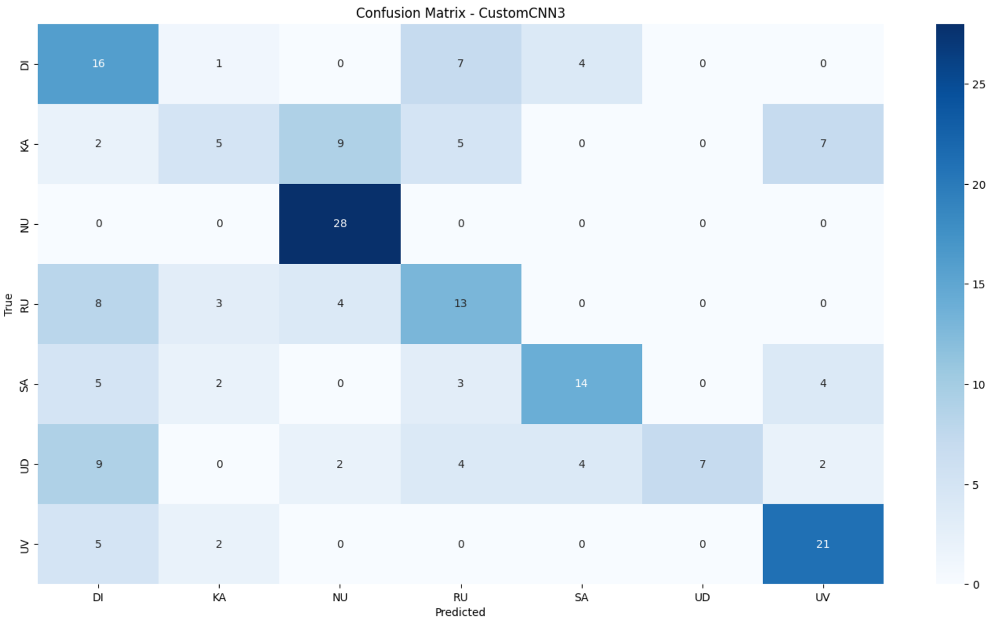

Figure 72: 8.2.6 Confusion Matrix - Custom CNN

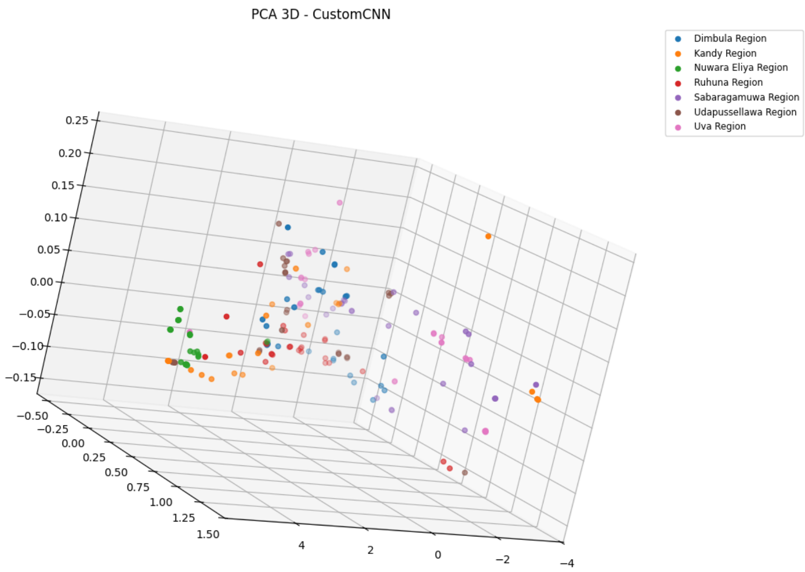

Figure 73: 8.2.6 PCA 32 - Custom CNN

8.3. Machine Learning Models

8.3.1. SVM

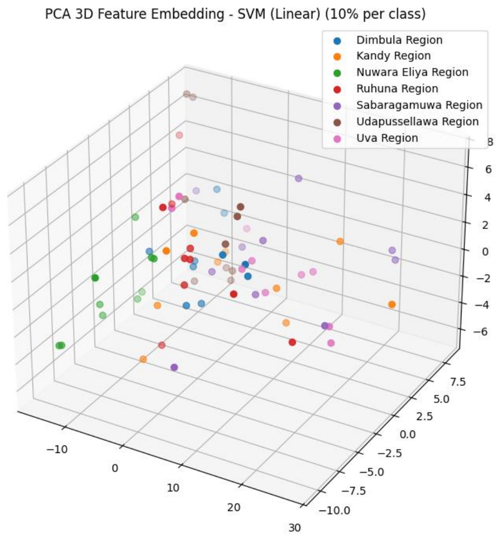

Figure 74: 8.3.1 PCA 3D – SVM

8.3.2. Random Forest

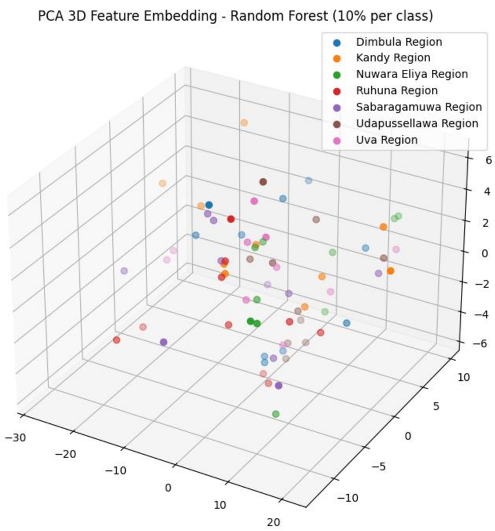

Figure 75: 8.3.2 PCA 3D – Random Forest

8.3.3. KNN

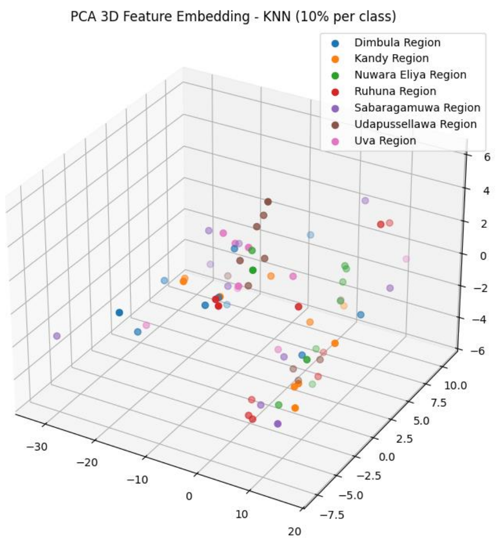

Figure 76: 8.3.3 PCA 3D – KNN

8.3.4. Logistic Regression

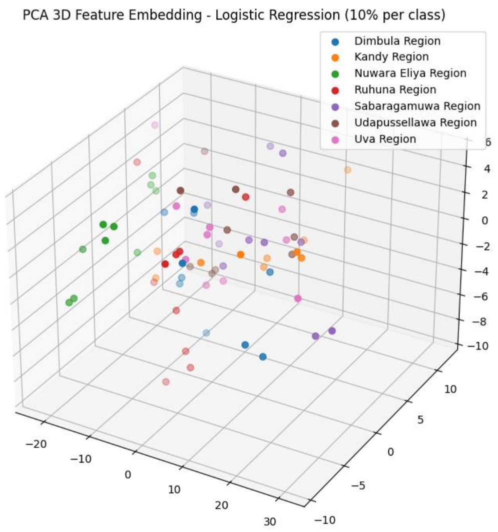

Figure 77: 8.3.4 PCA 3D – Logistic Regression

8.4. Composite Score Graphs

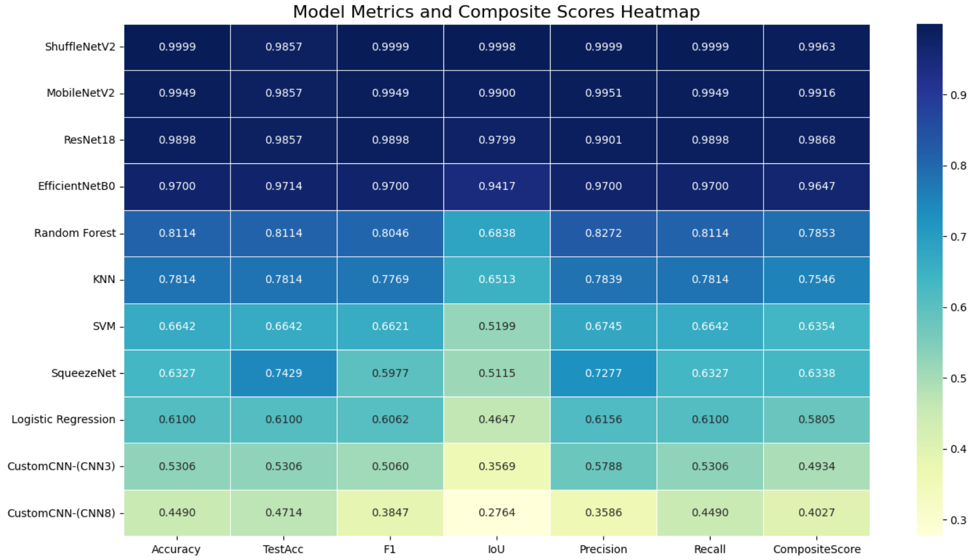

Figure 78: 8.4 CSG -1

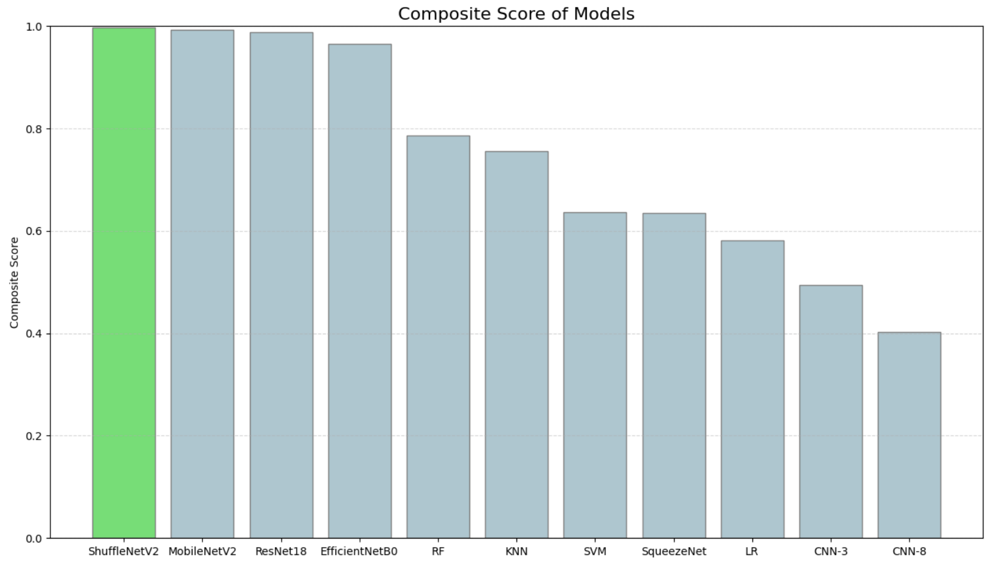

Figure 79: 8.4 CSG -2

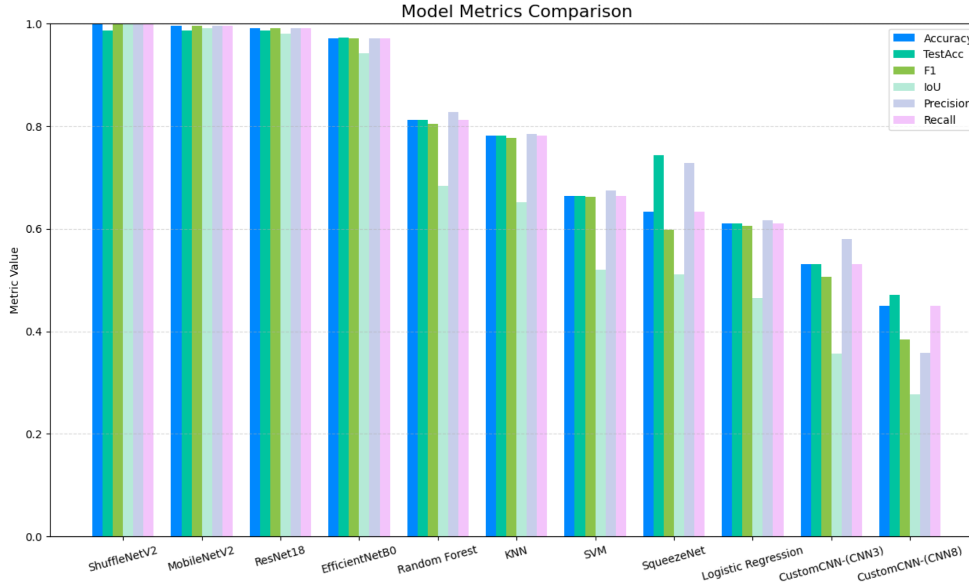

Figure 80: 8.4 CSG -3

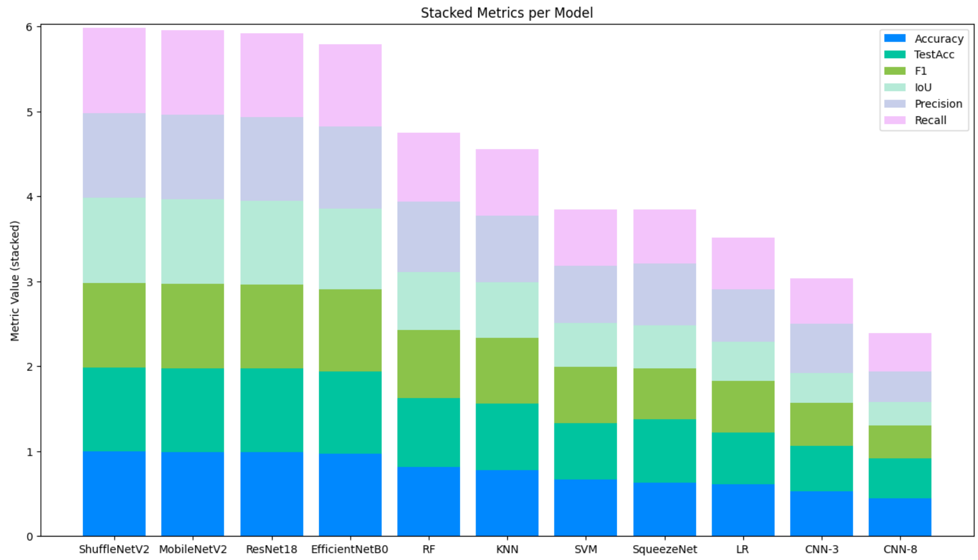

Figure 81: 8.4 CSG -4

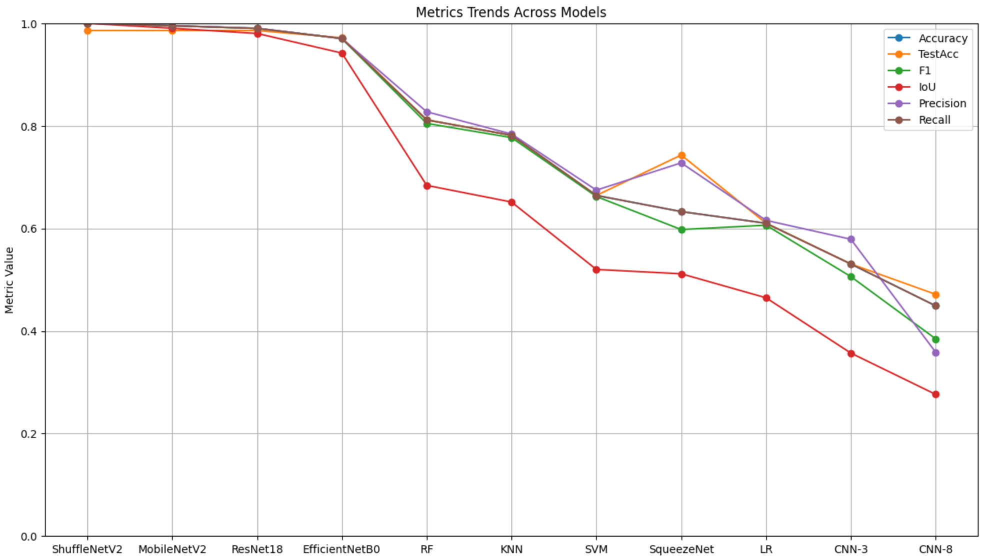

Figure 82: 8.4 CSG -5

8.5. Web Application

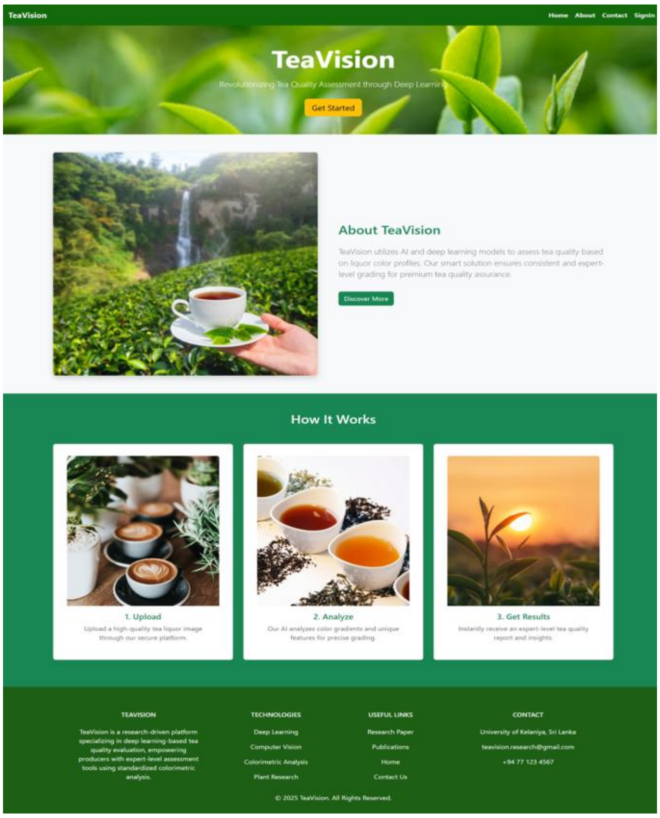

Figure 83: 8.5 Web Application -1

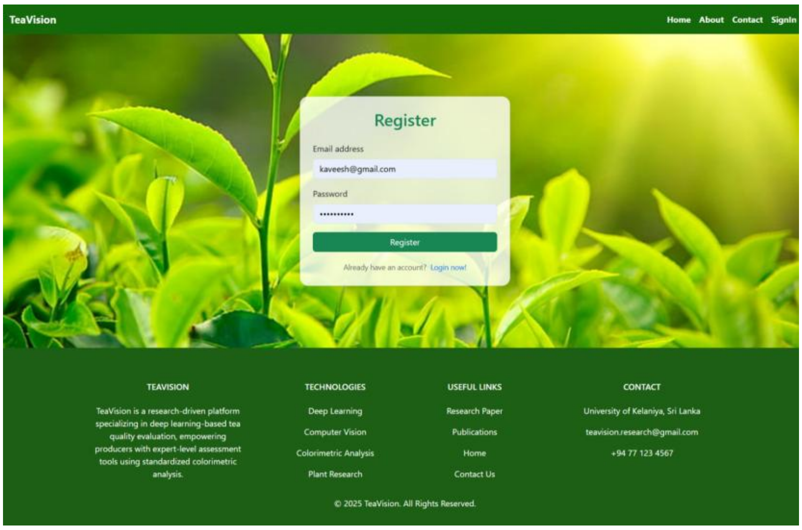

Figure 84: 8.5 Web Application -2

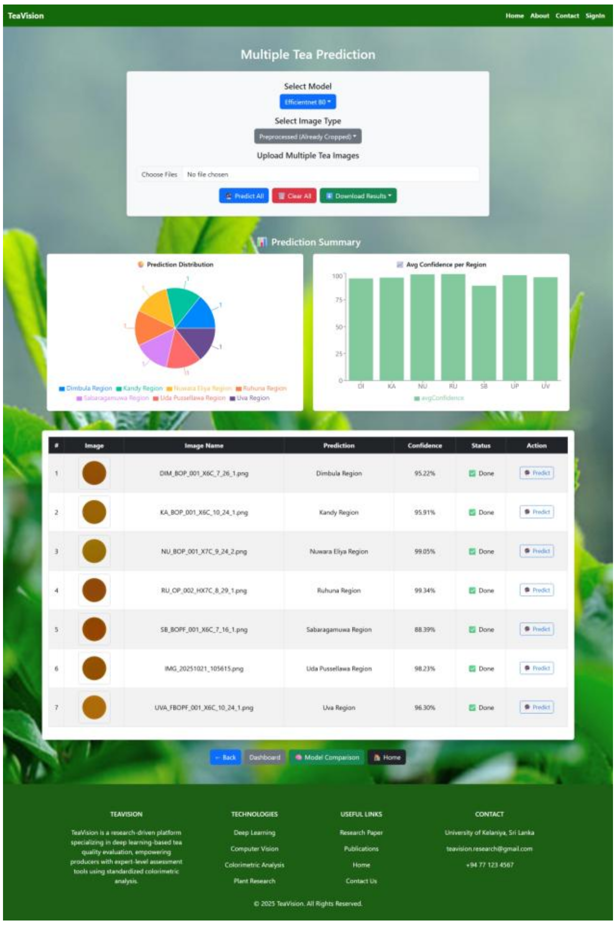

Figure 85: 8.5 Web Application -3

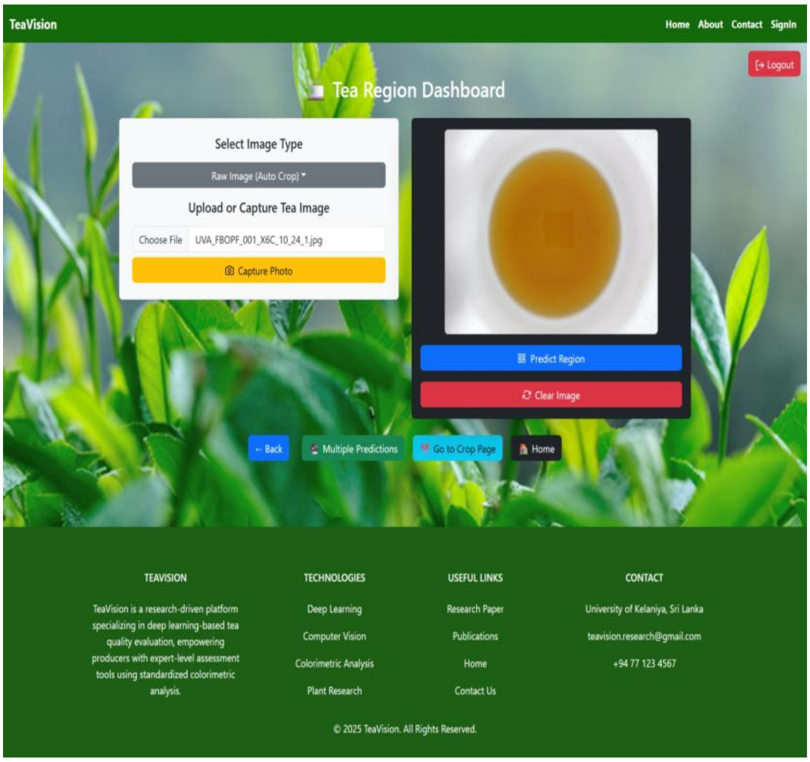

Figure 86: 8.5 Web Application -4

Figure 87: 8.5 Web Application -5

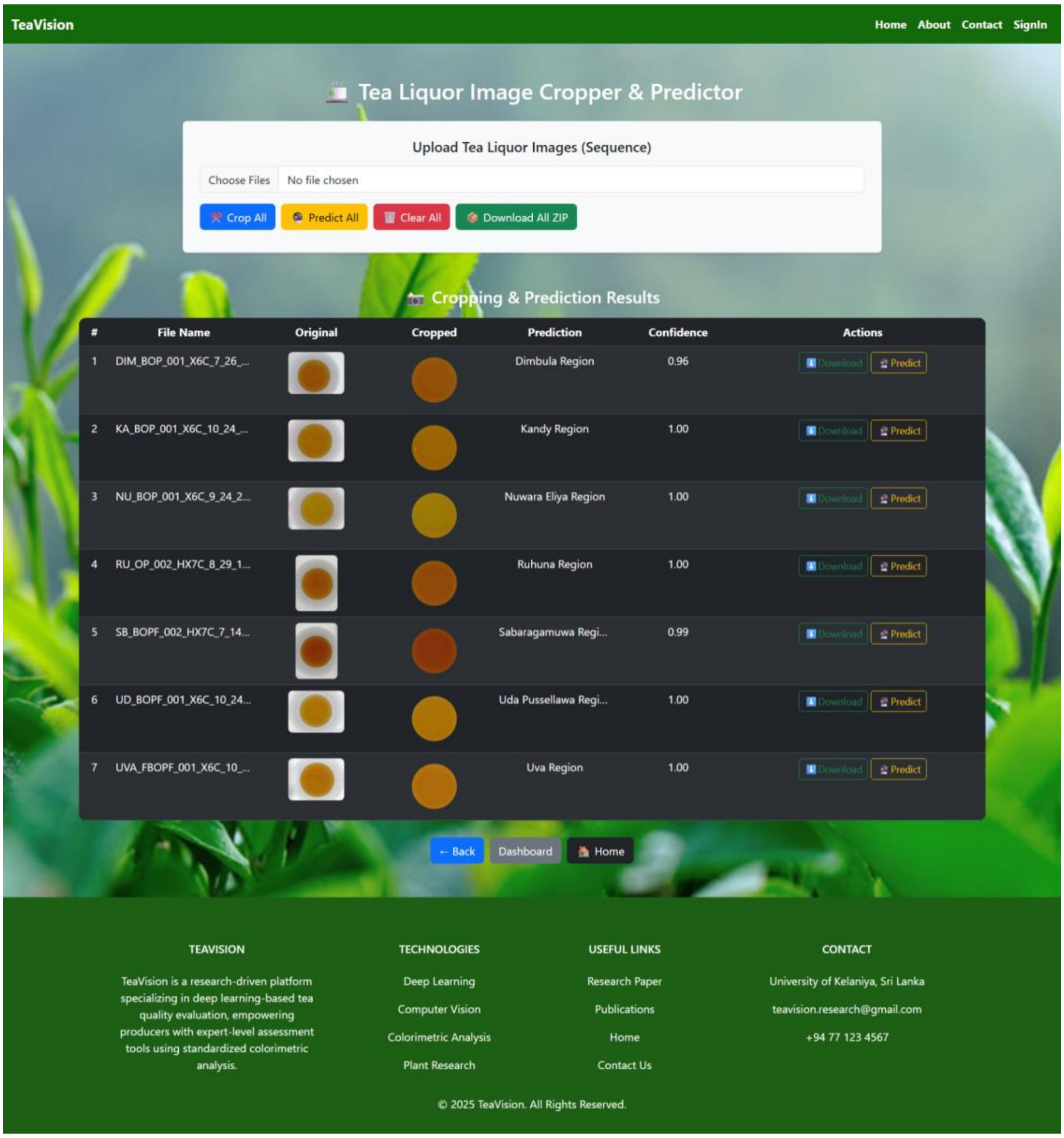

Figure 88: 8.5 Web Application -6

8.6. Android App

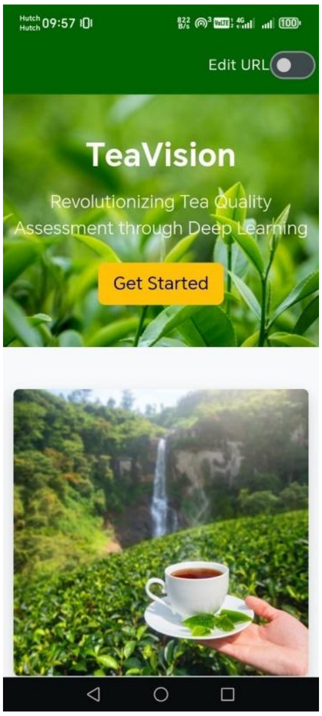

Figure 93: 8.6 Android App -1

Figure 94: 8.6 Android App -2

Figure 92: : 8.6 Android App -3

Figure 90: 8.6 Android App -4

Figure 89: 8.6 Android App -5

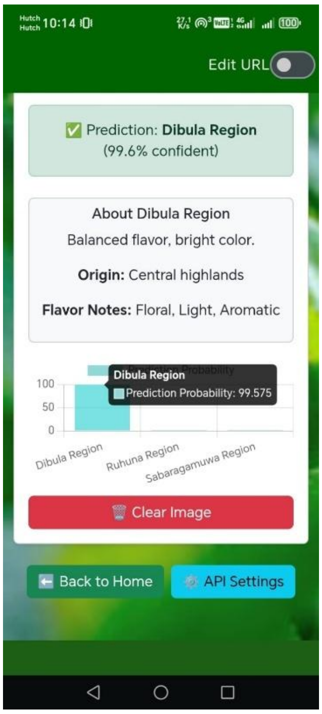

Figure 91: 8.6 Android App -6