Independent University **Bangladesh (IUB)**

IUB Academic Repository

Internship Reports 2026

2026-04-25

Machine-Vision-Based Bangladeshi Geographical Indication (GI) Dessert Recognitio

Bhuiyan, Shakibur Rahman

IUB

https://ar.iub.edu.bd/handle/11348/1186

Downloaded from IUB Academic Repository

2026_Spring____Internship_Report_Template.pdf

by Dip Kundu

Submission date: 28-Apr-2026 12:13PM (UTC+0600)
Submission ID: 2946384742
File name: 2026_Spring____Internship_Report_Template.pdf (52.08M)
Word count: 23007
Character count: 133747



An Undergraduate Senior Project on Topic

Machine-Vision-Based Bangladeshi Geographical Indication (GI) Dessert Recognition

By

Shakibur Rahman Bhuiyan

Student ID: **2221328**

Dip Kundu

Student ID: **2221683**

Md Tamzed Hossain Siam

Student ID: **2221390**

Spring, 2026

Department of Computer Science & Engineering

Independent University, Bangladesh

Supervised By:

Md Tarek Habib, PhD

Associate Professor

Department of Computer Science & Engineering

Independent University, Bangladesh

April 25, 2026

Attestation

We recognize that plagiarism is highly forbidden and is a breach of our university rules and regulations. This is an original, we say. Any copywritten information, models or intellectual input of other people that are involved in this project have been duly referred and referred to according to the internationally recognized academic conventions and standards.

Shakibur Rahman Bhuiyan Date: _____

Signature & Author

Dip Kundu Date: _____

Signature & Author

Md. Tamzed Hossain Siam Date: _____

Signature & Author

Acknowledgement

We would like to say first of all, that we are extremely thankful to the Almighty who enabled us to get the strength and patience to conduct this research to its successful conclusion.

We owe inestimable debt to our supervisor, **Md. Tarek Habib, PhD**, Associate Professor, Department of Computer Science & Engineering, Independent University, Bangladesh. His outstanding management, insight and patience in all levels of this project were key in determining the quality and the direction of our work.

We would like to acknowledge that **AKM Mahbubur Rahman, PhD**, Head of the Department of Computer Science & Engineering, Independent University, Bangladesh, provided a conducive and highly academic environment that was of much help to our development.

We would like to express our sincere gratitude to the owners and personnel of the approved dessert stores and retail merchants in the ten districts of Bangladesh who kindly allowed us to take pictures on-site. The collaboration was crucial to the reality of the BD-GI-Desserts13 dataset.

It is with great pleasure that we would like to thank our senior brother Zarif Wasif Bhuiyan, who has given such guidance, motivation and constant support during the period of this project. His counsel and motivation were able to guide us through difficult situations with confidence and keep us focused on our objectives. We also wish to express our deepest appreciation to our parents, friends, younger brother and sister, who have been our constant support, inspiration and strength during this time. Their unconditional love and active assistance in the data collection process were deeply cherished and greatly contributed to the success of this work.

We are also grateful to our peers and colleagues at Independent University, Bangladesh, whose encouragement and collaborative spirit kept us motivated throughout this project.

The Authors
Spring 2026

Letter of Transmittal

Shakibur Rahman Bhuiyan
Dip Kundu
Md. Tamzed Hossain Siam
Independent University, Bangladesh
Department of Computer Science & Engineering
April 25, 2026

To
Md. Tarek Habib, PhD
Assistant Professor, Department of Computer Science & Engineering
Independent University, Bangladesh

Subject: Submission of Undergraduate Senior Project Report

Dear Sir,

We are pleased to submit our undergraduate senior project report titled *"Machine-Vision-Based Bangladeshi GI Dessert Recognition"*, submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science & Engineering at Independent University, Bangladesh.

This research, conducted under your esteemed supervision, presents a machine vision framework for fine-grained visual recognition of thirteen Geographical Indication (GI)-certified traditional Bangladeshi desserts, serving as an origin-informed classification and decision-support tool. We constructed **BD-GI-Desserts13**, a curated dataset of 11,692 high-resolution RGB images collected from certified vendors across ten districts of Bangladesh. A Custom Convolutional Neural Network (CCNN) was developed and benchmarked against EfficientNet-B3, ShuffleNetV2, SVM, k-NN, and Decision Tree classifiers. The proposed CCNN achieved the highest test accuracy of **99.37%** and was deployed in a real-time web-based GI dessert recognition interface.

We sincerely thank you for your continuous guidance, constructive feedback, and generous investment of time throughout this project. We humbly request your kind acceptance of this report.

Sincerely,

Shakibur Rahman Bhuiyan	**Dip Kundu**	**Md. Tamzed Hossain Siam**
2221328	2221683	2221390

Evaluation Committee

Supervision Panel

.....

Md. Tarek Habib, **PhD**
Supervisor

Contents

Abstract xi

1 Introduction 1

1.1 Background and Motivation . . . 1
1.2 Problem Statement . . . 1
1.3 Objectives of the Study . . . 2
1.4 Research Gap . . . 2
1.5 Research Questions . . . 3
1.6 Research Outcome . . . 3

2 Literature Review 4

2.1 Existing Methods . . . 4
2.1.1 Bangladeshi Food and Dessert Recognition and Datasets . . . 4
2.1.2 Regional Dessert and Food Recognition Studies (Indian, Thai, Asian, etc.) . . . 5
2.1.3 General Food Image Recognition Techniques and Models . . . 6
2.1.4 Surveys and Reviews on Food Recognition and Computer Vision in Food Domain . . . 7
2.1.5 Foundational Deep Learning and Computer Vision Techniques . . . 8
2.1.6 Related Machine Vision Applications in Agriculture and Other Domains (Bangladeshi Research Context) . . . 8
2.2 Global AI-Based Food Authentication . . . 9
2.3 Gaps within the Existing Methods . . . 10
2.4 Problems with the Current Datasets. . . . 12

3 Dataset 14

3.1 Overview . . . 14
3.2 Data Collection . . . 14
3.3 Dataset Organization and Statistics . . . 15
3.4 Data Augmentation . . . 16
3.5 GI Dessert Classes . . . 17

4 Methodology **23**
4.1 The Methodological Framework . . . 23
4.2 Problem Formulation . . . 23
4.3 Classification Models . . . 24
4.3.1 Custom Convolutional Neural Network (CCNN) . . . 25
4.3.2 EfficientNet-B3 . . . 26
4.3.3 ShuffleNetV2 . . . 26
4.3.4 Traditional Machine Learning Classifiers . . . 27
4.4 Training Configuration . . . 28
4.5 Evaluation Strategy . . . 29
4.6 Software and Tools . . . 29
4.7 Ethical Considerations . . . 30

5 Experimental Results and Models Comparative Analysis **31**
5.1 Experimental Setup . . . 31
5.2 Deep Learning Model Performance . . . 32
5.2.1 Class-wise Quantitative Results . . . 32
5.2.2 Confusion Matrix Analysis . . . 33
5.2.3 Training and Validation Curves . . . 36
5.2.4 Grad-CAM Interpretability . . . 38
5.3 Traditional Machine Learning Baseline Performance . . . 40
5.4 Cross-Model Comparative Analysis . . . 41

6 Web-Based GI Dessert Recognition Framework **43**
6.1 System Architecture and Model Integration . . . 43
6.2 User Interface and Functional Modules . . . 44
6.2.1 Homepage . . . 45
6.2.2 About Page . . . 46
6.2.3 Recognition Page . . . 46
6.2.4 GI Dessert Map Page . . . 47
6.2.5 Dessert Gallery Page . . . 47
6.2.6 Dessert Details Page . . . 47
6.3 Real-Time Inference Capability . . . 48

7 Benchmarking and Comparative Analysis of GI Dessert Recognition Models **49**
7.1 Comparative Evaluation Against Published Studies . . . 49
7.2 Performance Analysis and Superiority of the Proposed Architecture . . . 51

92
8 Conclusion and Future Work **53**
8.1 Conclusion 53
8.2 Contributions and Impact 53
8.3 Limitations 54
8.4 Future Work 55

Publication Records **55**

Bibliography **61**

List of Figures

3.1 Geographic distribution of the 13 GI-certified dessert classes in BD-GI-Desserts13 at the district level. Each marker designates the primary source district of the respective product. 15

3.2 Typical on-site collection locations: (a) Cumilla Rashmalai, (b) Natore Kachagolla, and (c) Sherpur Chanapayesh. 16

3.3 Hierarchical folder structure of BD-GI-Desserts13 following the `Location_dessert[_variant]` naming convention. 16

4.1 Methodological pipeline of GI dessert recognition. 24

4.2 Architecture of the proposed Custom Convolutional Neural Network (CCNN) for GI dessert recognition 25

5.1 Confusion matrix of the Custom CNN on the BD-GI-Desserts13 test set. 34

5.2 Confusion matrix of ShuffleNetV2 on the BD-GI-Desserts13 test set. . . . 35

5.3 Confusion matrix of EfficientNet-B3 on the BD-GI-Desserts13 test set. . 36

5.4 Training and validation loss (left) and accuracy (right) curves of the Custom CNN over 30 epochs. 37

5.5 Training and validation loss (left) and accuracy (right) curves of ShuffleNetV2 over 30 epochs. 38

5.6 Training and validation loss (left) and accuracy (right) curves of EfficientNet-B3 over 30 epochs. 38

5.7 Grad-CAM activation maps for Custom CNN predictions across all GI dessert categories. Green labels indicate correct predictions; red labels indicate misclassified instances. Hot colors (red/orange) represent high activation regions; cool colors (blue/green) represent low activation. . . . 40

6.1 Three-tier microservice architecture of the web-based GI dessert recognition system, illustrating the Client Layer (React.js SPA), API Gateway Layer (Node.js/Express), and ML Inference Layer (Python/Flask). 44

6.2 Web application interface (Tablet View): (a) Homepage with an overview of the GI dessert recognition system and featured dessert thumbnails; (b) About Page presenting the project mission, BD-GI-Desserts13 dataset, and deployed deep learning models; (c) Recognition Page showing inference of Bogura Doi predicted with 100% confidence using the Custom CNN model. 45

6.3 Web application interface (Tablet View): (a) GI Dessert Map Page showing an interactive Leaflet-based geographical visualization of all 13 GI dessert origins in Bangladesh; (b) Dessert Gallery Page displaying all GI dessert categories; (c) Dessert Details Page presenting encyclopedic cultural, historical, and geographical information. 46

List of Tables

2.1 Summary of Key Gaps in Existing Food Recognition Methods 11
2.2 Summary of Existing Food and Dessert Image Datasets 13

3.1 Image counts per class, augmentation parameters, and dataset distribution. 17
3.2 Classes of traditional Bangladeshi sweets used in the dataset with descriptions and sample images. 18

4.1 Training Configuration for Deep Learning Models. 28

5.1 Class-wise Performance of Deep Learning Models on BD-GI-Desserts13 . 32
5.2 Class-wise Performance of Traditional ML Models on BD-GI-Desserts13 . 41
5.3 Overall Performance Summary of All Six Models 42

7.1 Comparative Analysis of Existing Food Image Recognition Approaches . 50

Abstract

Geographical Indication (GI) certification secures products where the quality, reputation and commercial value of the product is intrinsically connected with its geographical origin, and several traditional Bangladeshi desserts have recently received GI certification as part of a wider national policy on conserving heritage and regional branding. Nevertheless, their cultural and economic importance has not been accompanied by a standardized benchmark dataset or a strong machine-vision system devoted to the automated detection and authentication of Bangladeshi GI-certified desserts, putting consumers and honest producers at risk of mislabelling and counterfeiting. This gap is addressed in this thesis through the creation of a visual dataset, BD-GI-Desserts13, of 11,692 RGB images 512×512 pixels representing 13 official GI dessert classes across various districts, and proposes a complete machine-vision pipeline, which uses a combination of deep learning and classical machine learning methods to recognize fine-grained GI desserts. A computationally efficient Custom Convolutional Neural Network (CCNN) is developed and benchmarked as well as transfer-learning models such as ShuffleNetV2, EfficientNet-B3 and traditional machine learning classical classifiers SVM, k-NN, Decision Tree trained on deep feature embeddings, with class-stratified train/validation/test splits and standard metrics such as accuracy, precision, recall, F_1-score and confusion matrix. It has been demonstrated that the Custom CNN is able to achieve 99.37% test-accuracy and nearly perfect class-wise F_1-scores on most types of desserts, and is also able to outperform transfer-learning baselines as well as conventional techniques without incurring high computational complexity, which is highly conducive to real-time web and mobile application. Based on these results, this thesis proposes a conceptual framework for GI dessert visual recognition that serves as a reproducible baseline for future studies in origin-informed food classification and GI product awareness systems in Bangladesh. It is explicitly acknowledged that this system performs visual pattern-based recognition, not legal geographic origin verification, and should be interpreted as a decision-support tool rather than a certification authority.

Keywords— Geographical Indication (GI), Bangladeshi Desserts, Fine-Grained Image Classification, Deep Learning, Convolutional Neural Network (CNN), Transfer Learning, Food Authentication, BD-GI-Desserts13 Dataset

1 Introduction

1.1 Background and Motivation

Geographical Indication (GI) is a form of intellectual property protection that recognizes products originating from a specific region where a particular quality, reputation, or characteristic is essentially attributable to that geographical origin. In Bangladesh, thirteen traditional desserts have received official GI recognition — including Bogura Doi, Cumilla Rashmalai, Tangail Chomchom, Gopalgonj Rosogolla, Muktagacha Monda, Natore Kachagolla, Netrokona Balish Mishti, Tangail Sondesh, and Kushtia Tiler Khaza. These are not just foodstuff but symbolise the local and cultural background, the geographical location of the region and the craftsmanship of the region and the GI designation holds a lot of legal and business importance in verifying the geographical provenance and the genuineness of production. Non-GI products that are visually identical to certified GI desserts are often indirect mislabeled and marketed, amounting to consumer fraud and direct economic damage to bona fide regional producers, despite their cultural and economic significance. The technical problem here is that visual differences between GI dessert classes are extremely fine-grained - distinctions very narrow, sometimes confined to minute variations in color shade, surface texture, syrup consistency, and morphological boundaries. The brown and white colors of Muktagacha Monda and Tangail Sondesh such as these have a high intra-class visual similarity and cannot be reliably sorted out by standard food recognition systems.

New developments in Convolutional Neural Networks (CNNs) have shown good performance in fine-grained visual classifications through direct learning of hierarchical feature representations using raw pixel data. Nevertheless, these approaches are yet to be implemented into Bangladeshi GI dessert recognition, mainly because of the lack of a dedicated benchmark dataset and customized machine vision framework. The fusion of these technological opportunities and the unfulfilled need to have an automated GI dessert authentication system has led to this study.

1.2 Problem Statement

There is no provenance-conscious, automated machine vision system to detect and verify GI-certified Bangladeshi desserts. This classification problem has not been able to systematically evaluate its model because of the lack of a specific benchmark dataset. In the business world, such a loophole allows the commercial mass mislabeling of non-GI products as certified regional desserts, which harms the economic interests of the legitimate manufacturers and destroys con-

sumer confidence. It restricts digital culturally. attempts to maintain Bangladesh cuisine on an industrial level. This paper tackles the issue by making two complementary contributions: (1) the creation of BD-GI-Desserts13 a high-resolution image dataset of all 13 officially recognized GI dessert classes with district-level provenance annotations; and (2) the design, training, and systematic evaluation of a personal machine vision system to classify fine-grained, with provenance, GI desserts.

1.3 Objectives of the Study

The overall aim is to design and test a machine vision system that can perform fine-grained visual recognition of GI-authenticated Bangladeshi desserts, basing on their learned visual features, and can be used as a decision-support system to classify them by origin. The specific objectives are:

1. **To Construct a Dedicated GI Dessert Image Dataset:** Build BD-GI-Desserts13.— a high-resolution, RGB image set of 13 GI-certified classes of desserts in Bangladesh produced in their geographical areas of origin, labeled with a standardized *Location_dessert_variant* naming system, and divided into 70/15/15 repro- ducible model development and evaluation.

2. **To Design and Evaluate a Custom CNN Architecture:**

   A lightweight Custom CNN with convolutional filter banks, batch normalization, max-pooling, dropout, and early stopping, which is trained end-to-end on BD-GI-Desserts13 and tested on test accuracy, precision, recall, and F_1-score.

3. **To Conduct a Comparative Evaluation of Deep Learning and Traditional ML Models:** Compare the Custom CNN with transfer learning models (EfficientNet-B3 and ShuffleNetV2) and classic classifiers (SVM, k-NN, and Decision Tree) trained on Custom CNN embeddings, with the same preprocessing, training, and evaluation setup.

4. **To Develop a Web-Based GI Dessert Recognition Interface:** Deploy the trained Custom CNN in a real-time web-based interface enabling users to upload dessert images and receive immediate GI class predictions with confidence scores. The interface serves as a visual recognition aid and does not constitute a legally binding GI certification or geographic provenance verification tool.

1.4 Research Gap

Although deep learning has been widely applied to general food recognition, Bangladeshi GI-certified dessert classification has not previously been addressed in the machine vision literature. Existing Bangladeshi food datasets, including FoodBD, do not distinguish between GI-certified and non-GI products, nor do they record district-level provenance metadata, making them unsuitable for origin-sensitive authentication systems. Furthermore, no prior study

has conducted a systematic multi-architecture comparison across both deep learning and traditional ML paradigms for this fine-grained classification problem. This study simultaneously fills all three gaps.

1.5 Research Questions

1. Can a curated GI dessert image dataset collected under realistic conditions serve as a reliable and reproducible benchmark for fine-grained visual recognition research?
2. Does a domain-specific Custom CNN outperform state-of-the-art transfer learning architectures on GI dessert classification?
3. How do end-to-end deep learning models compare to traditional ML classifiers trained on CNN-extracted feature embeddings for this task?
4. Do deep learning models attend to visually meaningful, class-relevant regions when classifying GI desserts, and do these correspond to the authentic characteristics that define each GI product?
5. Is the proposed framework practically deployable as a real-time web-based GI authentication tool?

1.6 Research Outcome

The study successfully constructed BD-GI-Desserts13, comprising 11,692 high-resolution RGB images (512×512 pixels) across 13 GI-certified dessert classes, captured at verified production sites across ten districts of Bangladesh. The proposed Custom CNN achieved the highest test accuracy of 99.37% with a macro-averaged F_1-score of 0.9937, outperforming EfficientNet-B3 (93.17%), ShuffleNetV2 (98.12%), and all traditional ML baselines — k-NN (91.18%), SVM (91.01%), and Decision Tree (86.74%). Grad-CAM visualizations confirmed that the model reliably attends to class-relevant discriminative features such as surface texture, color patterns, and structural boundaries, validating its suitability for real-world GI authentication and establishing a scalable infrastructure for the digital preservation of Bangladesh's certified culinary heritage.

2 Literature Review

Recent developments in computer vision and deep learning have greatly contributed to improved automated food image recognition, nutritional estimation, and dietary assessment systems. The technologies are especially useful in Bangladesh because of the increasing rates of diabetes and obesity prevalence, which have posed an immediate demand on the accurate and culturally sensitive tools to aid in monitoring daily nutrition. Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) have proven to be highly successful in hierarchical visual feature extraction and real-world conditions including varying lighting, cluttered backgrounds, and multiple dishes on the same image. There is currently a broad body of research on model architectures, hybrid methods, regional data, background techniques, surveys, and application to specific domains. This literature review categorizes the available work into concise sections and discusses every contribution individually. The review notes the development of food recognition methods beginning with early CNN-based systems to more recent hybrid systems, and the significance of culturally relevant datasets and real-world relevance of dietary assessment in Bangladesh. Through a systematic analysis, the review establishes the strengths of the existing methods and the unresolved areas that drive the ongoing senior project on a hybrid CNN-ViT-based system to recognise food and estimate nutritional content of Bangladeshi food. The synthesis also highlights how previous research has increasingly gravitated towards more practical, explainable, and lightweight solutions, but which nonetheless have much more to offer in terms of culturally specific applications applicable to the dietary practices of the Bengali population and the overall public health issues that affect them.

2.1 Existing Methods

2.1.1 Bangladeshi Food and Dessert Recognition and Datasets

A number of studies have specifically dealt with the recognition of Bangladeshi food and dessert. In the research, deep convolutional networks were trained on an original dataset of traditional Bengali desserts and obtained good classification results [1]. This initial experiment proved that CNNs can be effective in differentiating visually similar local sweets given typical conditions of Bangladeshi cuisine and formed a significant basis of using deep learning to culturally specific desserts which commonly have similar visual appearances because of similar ingredients and cooking techniques. This kind of contributions is particularly pertinent due to Bengali sweets which are a fundamental component of meals and celebratory events in Bangladesh and automated recognition is a viable solution to enhanced dietary monitoring. A detailed dataset of the Bangladeshi dessert classification includes images of all types

of traditional sweets and allowed MobileNet-based models to achieve 98% test accuracy with a lot of data augmentation and thorough testing [2]. The dataset has emerged as a significant benchmarking resource in future studies of desert regions, as it presents a consistent set of data to use to train and compare lightweight models to be used in a mobile environment with low resources. The polygon-annotated Meal image dataset FoodBD is a collection of 3523 smartphone-photographed images of Bangladeshi cuisines with 67 categories and detailed polygon annotation and nutritional data (carbohydrates, protein, fat, fiber, calories, and glycemic load) estimated by experts [3]. The data set was particularly planned to assist diabetic nutrition evaluation as well as culturally various AI teaching, responding to the absence of nutritionally designated domestic data that is vital in practical health use in a nation where the diabetes prevalence is ever increasing. Bangladeshi Street Food Calorie estimation with Improved YOLOv8 and Regression Model involved the combination of an improved YOLOv8 detector and a regression head that has been used to calculate accurate real-time detection and nutritional quantification of popular street foods in uncontrolled real-world environments [4]. This offered a good move to deployable systems to be used in daily life in urban and rural environment where street food makes a much larger part of daily food intake and emphasized the possibilities of object detection models to be used in dynamic and outdoor setting typical of Bangladesh.

2.1.2 Regional Dessert and Food Recognition Studies (Indian, Thai, Asian, etc.)

Similar issues in other countries have been discussed in regional dessert and food recognition studies which offer valuable information in the Bangladeshi contexts. Non-destructive characterization of Indian sweet meat spongy Rosogulla by image analysis was used to extract Gray-Level Co-occurrence Matrix (GLCM) texture features at various offsets and compared the statistical value of the features in order to evaluate the quality parameter reliably without any destruction of the sample [5]. This texture-based approach validated the utility of computer vision to Indian traditional sweets and suggested methods that can be used to augment deep learning in quality control tasks, particularly those that share similar appearance. Deep neural networks were used to detect and localize food-objects in real-time in Indian cuisines with high localization accuracy of different Indian dishes in unconstrained settings [6]. The article highlighted the significance of real-time performance to apply in real circumstances in busy meal environments and provided learning points to manage crowded or overlapping food items. The time-delay machine learning model of shelf life of Burfi used regression methods to predict freshness using time-delay parameters [7]. This offered a handy quality control instrument in conventional sweet manufacturing that can motivate the equivalent freshness estimation solutions to Bangladeshi desserts. KediNet was a hybrid deep learning model that recognizes Thai desserts, which consists of CNN and transformer blocks and provides high classification accuracy of Thai sweets, being able to process their visual complexity [8]. The model showed the advantages of fine-grained dessert classification using a hybrid architecture. A comparative analysis of CNN and Vision Transformers on the classification of Indonesian traditional cakes demonstrated the

competitive advantages of transformer-based models to recognize culturally specific desserts [9]. The Asian Food Image Classification Based on Deep Learning combined Convolutional Block Attention Module (CBAM) with MobileNetV2, VGG16, and ResNet50 along with Mixup augmentation and achieved a top-1 accuracy of 87.33 on a test set of various dishes in Asia [10]. This demonstrated that the attention mechanisms have the potential to enhance performance on various Asian food types. A fine-grained recognition method to identify Chinese food images applied optimized CNNs to effectively cope with fine intra-class variations [11]. Enhanced food image recognition through the use of deep learning and data-driven techniques with an application to the Central Asian Food Scene had greater generalization and sturdy performance on regional food [12]. VTnet+Handcrafted based methodology of food cuisines classification by using Vision Transformer ensembles with hand-crafted features (GIST, HoG, LBP) over a 13-class Indian food dataset and achieved 94.63 percent accuracy, 84.42 sensitivity, and 0.93 kappa coefficient outperforming ensemble CNN architectures [13]. The approach demonstrated a benefit of integration between transformer models with traditional characteristics to classify better in culturally diverse food scenes.

2.1.3 General Food Image Recognition Techniques and Models

General food image recognition methods and models have designed strong international standards on which localized systems are based. An inceptionv3-based deep learning system for dessert recognition to offer automated dietary evaluation based on transfer learning on a dessert versus non-dessert split of Food-101 achieved 95.89% validation accuracy [14]. Deep learning in automatic classification of fruits in industrial applications demonstrated the efficiency of CNNs in real-time quality sorting of fruits, which has similar methodological approaches with food freshness assessment [15]. Experiments and results combined with a new dataset of food recognition presented a new benchmark and tested various CNN architectures to classify food in general [16]. The use of CNN-based features in retrieval and classification of food images has affirmed that ResNet-50 is a very robust backbone when fine-tuned using food-specific data [17]. The deep feature aggregation of food images as a state recognition was also confirmed [18]. A food recognition system that was used to track the diet of diabetic patients was an optimized bag-of-features model [19]. Convolutional neural networks were optimized on fine-grained food image recognition to achieve better results on fine detail differences [20]. Convolutional neural network and web crawlers were used to classify food images and extract data with 97% accuracy on 20 classes and combines nutritional data retrieval by automated web scraping [21]. Res-VMamba on fine-grained visual classification of food category with selective state space models with deep residual learning outperformed ViT and CNN baselines on CNFOOD-241 dataset [22]. Multi-level attention feature fusion with self-attention and maximization of the KL-divergence coupled with multi-level fusion and attention mechanism (MAF-Net) enhanced food image recognition with Top-1 accuracies of 90.22% on CETH Food-101 b[23], 89.86% on Vireo Food-172 and 90.61% on UEC Food-100 General categorization based on food classification with the use of transfer learning technique was applied [24]. Food image recognition using noisy vision transformer was enhanced to become better robust to noisy settings [25].

The performance of food recognition based on hybrid transformer models and image preprocessing methods was improved [26]. Combination of deep residual neural network features and supervised machine learning algorithms was used to classify various food image datasets [27]. Modern CNN architectures were used to create large scale visual food recognition benchmarks [28]. A CNN and vision transformer-based method with explainable real-time food recognition incorporated the ResNet50 with ViT, applied the system to mobile platforms, and presented the interpretability by using the Grad-CAM and LIME [29]. Precise and real-time and automated food recognition using pretrained MobileNetV2 in food industry highlighted lightweight implementation [30]. Mish activation was applied to the YOLOv4 to achieve a better classification strategy towards freshness of fruits and vegetables and recorded better average precision of freshness detection [31].

2.1.4 Surveys and Reviews on Food Recognition and Computer Vision in Food Domain

Food recognition surveys and reviews have generalized progress and emphasized unmet challenges in the food domain. Food image recognition and nutrition analysis using deep learning with a focus on chronic disease monitoring placed an emphasis on its use in the healthcare domain and systematically reviewed how image-based systematically analyzed how food image-based systems could aid in supporting long-term dietary management of chronic diseases such as diabetes and obesity [32]. Such an overall perspective is especially applicable to Bangladeshi settings since it ties the visual recognition directly with the nutritional effects, providing a framework that the current senior project is based on regarding the local assessment of diabetic diets. A systematic review of mobile computer vision-based food recognition and volume/calorific estimation applications reported that 90.9 percent of systems cannot differentiate between food and non-food and that the explainability of most systems is too low to be trusted by the user [33]. The review also identified severe deficiencies in practice, including inadequate handling of unconstrained environments and inadequate integration of volume estimation, which is necessary to the correct counting of calories in the daily Bangladeshi meals. Food-recognition systems based on images on dietary assessment were assessed by a systematic review of already existing pipelines, which showed both achievements under controlled conditions and constant limitations when used to different cultural foods [34]. An overview of image-based systems of food recognition and volume estimation artificial intelligence was given with a wide range of approaches, including both conventional feature detection and current deep learning techniques, and their applicability to real-world dietary tools [35]. Food classification and volume estimation using images to assess diet was surveyed considering methodological developments and their effect on chronic disease surveillance [36]. It provided a general overview of food computing, which offers an interdisciplinary viewpoint, incorporating computer vision and nutrition science and shows future directions of culturally adaptive systems [37]. An extensive review of computer vision-based food grain recognition was provided, which summarized methods of grain-specific recognition capable of guiding the larger food category recognition initiatives, such as those required in Bangladeshi food staples [38]. Together, these surveys offer a critical background to

the comprehension of the state of the art and clearly emphasizes the importance of more localized, explainable and resource-efficient responses to the context of developing countries such as Bangladesh.

2.1.5 Foundational Deep Learning and Computer Vision Techniques

Basic deep learning and computer vision methods are still an important part of modern architectures and are still fundamentals in food recognition systems. The convolutional neural networks survey was conducted, including the analysis, applications, and perspectives, with the detailed discussion of the architectural evolution, training techniques, and new trends that directly guide the design of the current food image classifier [39]. This survey is a pillar reference since it outlines the manner in which CNNs evolved beyond mere image classification challenges to fine-grained visual differences needed by food products. Image retrieval through aggregation of deep convolutional features was proposed as a method to provide a better representation of features and similarity matching, which offers techniques often used in food recognition pipelines to generate higher retrieval accuracy with large datasets [40]. Small vegetable seed sowing was also recognized online using the machine vision technique which was demonstrated using practical vision-based systems which can be used in real-time with varying field conditions [41]. The work is related to agricultural seeds, but it provides the important lessons in lightweight, robust computer vision deployment that can be similar to the needs of mobile food recognition applications in resource-constrained environments like rural Bangladesh. All of these contributions can be considered the theoretical and technical foundations upon which more specialized food recognition models are developed, with modern systems enjoying the advantage of principles of feature extraction, feature aggregation, and real-time processing that are well understood.

2.1.6 Related Machine Vision Applications in Agriculture and Other Domains (Bangladeshi Research Context)

Similar machine vision uses in agriculture and other areas (Bangladeshi research context) have already applied the technology to local issues and show that computer vision can be applied to the problems of Bangladesh in a wider context. Papaya disease identification with CNNs was employed in machine vision which was developed to recognize papaya disease locally, to help detect common papaya diseases by analyzing images and providing a model which can be used to detect food quality in agricultural produce [42]. Correct machine vision local fish recognition of freshwater species typical of Bangladesh was attained, which can be used as an effective classification system of native varieties of fish that are staples of Bangladeshi diets [43]. Eye disease recognition using convolutional neural network considered the most common eye diseases in Bangladesh, demonstrating how deep learning may be applied to healthcare imaging problems and supporting the viability of implementing similar models in applications that monitor diets [44]. Potato species recognition of the local potato varieties was done using machine vision to provide the high-accuracy classification, which can help in supporting the agricultural productivity and has the potential to motivate similar methods in separating

similar-looking Bangladeshi food items [45]. This work was later continued by deep learning models of potato breed recognition, which includes more complex architectures to address the small breed-level variations, again demonstrating the potential of deep learning to fine-grained recognition problems in the context of local agriculture and food systems [46]. In Bangladesh, convolutional neural network-based tomato pest recognition achieved an accuracy of 95.49% with five prevalent tomato pests in mind, thus proving the usefulness of CNNs in the realistic agricultural contexts in Bangladesh [47]. A study of machine vision based insect recognition in agriculture has synthesized the world and regional insect recognition applications and identified the best of these applications which can be applied in food recognition and pest detection within Bangladesh farming environments [48]. The artificial intelligence food authentication process through a label attribute to fraud detection surveyed the changes between the traditional label-based system to the advanced deep learning system to authenticate food origin and quality which provided important insights of authenticity verification that complemented food recognition systems and ensured safety in the Bangladeshi food supply chain [49]. These associated applications demonstrate the degree of maturity of machine vision technologies in Bangladesh, as well as they offer transferable methodologies, datasets, and evaluation plans that underpin the rationale of the current senior project, focusing on culturally specific food recognition. To reinforce this trend, deep CNN-based methods have proven to be very accurate in plant recognition of leaf images and this underscores the generalizability of such models to more domain specific classification problems [50].The machine vision techniques have been widely used in monitoring and identifying diseases in fruits and vegetables, and comparative studies have helped to identify the state of the art and provided future research directions [51].

2.2 Global AI-Based Food Authentication

Beyond food recognition, AI has been extensively applied to food authentication and geographical origin verification — tasks that directly parallel the GI certification problem addressed in this thesis. The review article by Feng et al. [52] examined the use of VIS/NIR spectroscopy and hyperspectral imaging (HSI) and machine learning in various food categories and showed that spectral-spatial data fusion always performs better than single-modality methods in terms of food variety and origin discrimination. Their comparison of CNN-based models to identify HSI food stated that their accuracy was over 97 percent in identifying the type of rice, maize, and tea varieties which supports that CNNs are effective in origin-sensitive classification. When trained on image models, at the image-processing end of this spectrum Magdas et al. [49] demonstrated that CNN models achieved up to 97% accuracy on food adulteration detection tasks, outperforming LS-SVM (91%) and PLS-DA (79%) on identical datasets, especially on subtle cases of fraud that demanded fine-grained visual discrimination. Deng et al. [53] also overviewed the field of deep learning to food authenticity, stating that image-based DL represents the fastest-growing trend and that it is independent of human-engineered features and has better scalability. The three independent international reviews combined substantiate the qualitative methodology used in this thesis namely the use of CNN-based image classification

as a regionally conscious authentication tool of visually distinct regional food products.

2.3 Gaps within the Existing Methods

Despite the significant advancement in food image recognition, as shown by the reviewed literature, there are still a number of important gaps that hinder the practical implementation of the existing research, especially in the Bangladeshi context. The majority of current approaches are based on relatively low-sized datasets or images under laboratory conditions, which limits their applicability to real-life Bangladeshi settings with different lighting conditions, background clutters, and overlapping dishes [1],[2], [3],[4], [29], [13]. High-performing models, such as hybrid CNN-ViT models and transformer models, are highly accurate in benchmarking environments but have not been optimized adequately to real-time nutritional estimation or mobile systems with limited resources that are a frequent feature in rural Bangladesh [30], [31]. Many traditional machine vision methods based on handdesigned features and PCA tend not to be robust to morphological changes between closely related dishes, and do not work in unconstrained settings [5], [20]. It is often the case, even with the more sophisticated hybrid models that have attention mechanisms, that they need pre-segmented or pre-processed images, which are not feasible in real-world dietary monitoring systems where users take images at random [23]. In addition, the majority of systems are computationally expensive and do not have in-built explainability features that are critical in fostering user confidence in the applications of healthcare like diabetes management. There is limited research that has incorporated both proper food identification and valid nutritional estimation in one culturally sensitive model to Bangladeshi food. Although recent publications on GI desserts have started to consider the fine-grained local recognition, they remain infantile and are yet to integrate the recognition with extensive nutritional examination or complete mobile implementation that can be used in everyday life. These shortcomings point at the necessity of lightweight, explainable, and culturally specific systems that can be applied in a real-life situation in Bangladesh. Other limitations are a lack of support with multi-dish scenes, lack of emphasis on estimating glycemic loads in diabetic patients, and the lack of end-to-end pipelines that integrate detection, classification, and nutritional regression together in one analysis. Another issue brought about by the use of large scale international benchmarks is the domain shift problem whereby the models are applied to food items in Bangladesh that vary widely in terms of appearance, texture and presentation. To easily explain these limitations, Table 2.1 lists the key gaps identified in the studies reviewed.

The gaps outlined in Table 2.1 indicate institutional flaws that, when combined, make existing food recognition systems ineffective in assessing the real-world diet in Bangladesh. An example is that poor generalization is a direct consequence of small or controlled datasets since when trained on homogeneous backgrounds or in isolation, models fail to perform when faced with the heterogeneous visual nature of Bangladeshi thalis or street-food plates. Equally, lack of explainability mechanisms that is emphasized in various reviews undermines the confidence of users in healthcare applications, where patients and clinicians need access to clear explanations of why nutritional predictions are made. The absence of a built-in nutritional estimates is

Table 2.1: Summary of Key Gaps in Existing Food Recognition Methods

Gap Category	**Description**	**Affected Papers**	**Impact on Bangladeshi Context**
Limited Dataset Size & Diversity	Reliance on small or controlled-condition datasets with insufficient environmental variation	[1], [2], [3], [4], [5], [13], [29]	Reduces generalization to real Bangladeshi meals with variable lighting, clutter, and multi-dish scenes
Lack of Real-Time & Mobile Optimization	Models are computationally heavy and not optimized for low-resource mobile devices	[30], [31], [23]	Hinders deployment in rural Bangladesh where smartphones are the primary access tool
Absence of Explainability	No built-in mechanisms (e.g., Grad-CAM, LIME) for model interpretability	[29], [33]	Reduces clinical trust for diabetes management applications
Domain Mismatch for Local Cuisine	Heavy dependence on Western/East-Asian benchmarks with limited Bangladeshi representation	[2], [3], [10], [11], [12]	Causes significant performance drop when transferred to culturally specific Bangladeshi foods
No Integrated Nutritional Estimation	Focus primarily on classification; lacks end-to-end calorie/glycemic load regression	[4], [14], [19], [21]	Prevents practical dietary assessment for diabetic patients
Poor Handling of Unconstrained Scenes	Models require pre-segmented images and struggle with occlusion/multi-dish plates	[6], [23], [29]	Impractical for spontaneous, everyday Bangladeshi meal photography
Computational Inefficiency & Lack of Lightweight Design	Heavy reliance on large backbones (ResNet50, ViT, etc.)	[30], [22], [23]	Unsuitable for battery-constrained mobile devices in low-resource settings
Insufficient Focus on Glycemic Load	Limited or absent glycemic index/load estimation	[3], [32]	Critical gap for diabetes management in Bangladesh, where high-glycemic local foods are common

another contributing factor to the issue, as without the data on the calorie and glycemic load, classification will not allow meaningful diabetes treatment. The mismatch of domains between international standards and Bangladeshi cuisine exacerbates these restrictions, with the fact that small visual distinctions between local sweets and curries necessitating culturally sensitive models. Altogether, although the current body of literature has contributed to the development of the technical basis of food recognition, it is yet to provide practical, lightweight, explainable, and culturally-adapted solutions that can be widely applied in Bangladesh. The need to fill these gaps is the essence of the present senior project that generates a big, diverse dataset of Bangladeshi food and a hybrid CNN-ViT model to be trained on real-time, mobile, and explainable dietary assessment.

2.4 Problems with the Current Datasets.

The current datasets have several limitations that can be identified to identify food images and track the food intake that can interfere with the development of highly accurate and generalized models. The publicly accessible data is mostly related to Western or East-Asian cuisines rather than to the Bangladeshi dishes creating an incongruence to domain that reduces the adaptability of the models [2, 3]. Moreover, these types of datasets are often small, sparse and lack environmental variation, not capturing the whole range of natural stimuli (e.g. variation in lighting, cluttered backgrounds, occlusion and multiple dishes) that naturally vary in the world [4]. Many of the previous studies utilized synthetics or in-laboratory images to train and test deep learning models that may be very accurate when tested but fail to transfer to real world scenarios [1], [5]. These datasets commonly include neat, isolated food images on homogeneous backgrounds disregarding the complicated visual environment of Bangladeshi meals where several dishes can be observed at the same time, or part of a dish can be hidden behind other objects. Moreover, the unequal distribution of data with overrepresenting and underrepresenting of certain categories introduces bias, leading to overfitting of the models and poor recognition of the infrequent or traditional dishes. Another major problem is the second challenge, the lack of annotated datasets or confirmed datasets dedicated to the Bangladeshi foods with matching nutritional labels. Due to the lack of high-quality and reliable images of local dishes, it is hard to train the models that are able to produce accurate detections and calorie estimations. Additionally, the existing datasets, on the whole, do not consider the morphology and visual dissimilarity of adjacent-looking dishes. To overcome these problems associated with the dataset, our senior project creates a big dataset of the genuine Bangladeshi food images gathered in different districts and of high-quality. The data set is well endowed with diverse lighting, background and positional differences to enable it to be robust and generalizable. Our work presents a good base of creating effective deep learning frameworks which would improve dietary assessment and nutritional monitoring by overcoming the issue of data imbalance and incorporation of local dishes.

Table 2.2 summarizes the most pertinent food and dessert image datasets that are covered in the literature. It makes it quite clear that the majority of the available datasets are either region-focused (e.g., Thai, Chinese or Central Asian cuisines) or general-purpose benchmarks that are not very relevant to the context of Bangladeshi foods. Furthermore, the table highlights common limitations such as small dataset sizes, lack of environmental diversity, and absence of nutritional annotations. These gaps demonstrate the necessity to develop a large, diversified, and context-specific dataset unique to Bangladeshi cuisine, which is the immediate trigger in the development of the dataset in this work.

Table 2.2: Summary of Existing Food and Dessert Image Datasets

Dataset Name	**Ref.**	**Classes**	**Images**	**Annotations**	**Domain / Focus**	**Key Features**	**Limitations**
Bengali Sweets Dataset	[1]	Multiple	Custom	Class labels	Bengali sweets	Culturally specific to Bangladesh	Small scale; limited environmental variation
Bangladeshi Dessert Dataset	[2]	Multiple	Thousands	Class labels + augmentation	Traditional Bangladeshi desserts	High-quality images; 98% MobileNet accuracy	Limited to desserts; no nutritional labels
FoodBD	[3]	67	3,523	Polygon segmentation + nutritional labels	Bangladeshi cuisines	Polygon annotations; nutritional info (calories, macros, glycemic load)	Smaller dataset; limited to meals
Indian Cuisines Dataset	[6] [13]	13	∼400	Class labels	Indian cuisines	Real-time detection; hybrid handcrafted + transformer features	Very small dataset; limited diversity
Thai Dessert Dataset (KediNet)	[8]	Multiple	Not specified	Class labels	Thai desserts	Hybrid CNN–ViT architecture	Region-specific; not Bangladeshi
Indonesian Traditional Cakes Dataset	[9]	Multiple	Not specified	Class labels	Indonesian cakes	CNN vs. Vision Transformer comparison	Limited to cakes only
Asian Food Dataset	[10]	Multiple	Not specified	Class labels	Asian foods	CBAM + Mixup augmentation; 87.33% accuracy	Broad Asian focus; not Bangladesh-specific
Chinese Food Dataset	[11]	Multiple	Not specified	Class labels	Chinese foods	Fine-grained recognition	Culturally different from Bangladeshi cuisine
Central Asian Food Scene Dataset	[12]	Multiple	Not specified	Class labels	Central Asian foods	Deep learning + data-driven methods	Not relevant to Bangladeshi context
Food-101 (Benchmark)	[14] [16]	101	∼101,000	Class labels	General food	Large-scale benchmark dataset	Western/East-Asian bias; no Bangladeshi dishes
UEC Food-100 / Vireo Food-172	[23]	100 / 172	Large	Class labels	General & fine-grained food	Multi-level attention fusion	Not culturally aligned with Bangladesh

This table summarizes key datasets from the literature, highlighting the lack of large-scale, diverse, and nutritionally annotated datasets for Bangladeshi cuisine—motivating the proposed work.

3 Dataset

3.1 Overview

BD-GI-Desserts13 is a benchmark dataset which consists of 11,692 high-resolution RGB pictures belonging to 13 formally recognized GI-certified Bangladeshi dessert classes. It is the earliest publicly accessible machine vision dataset focused on origin-aware, fine-grained detection of the Bangladeshi GI desserts, an objective with subtle chromatic, textural, and morpho-logical differences among the visually similar subclasses [54]. The dataset is publicly available on Kaggle with no restrictions to access to facilitate reproducible research in the field of GI product authentication and preservation of the Bangladeshi cultural heritage.

The 13 classes are desserts having quality and identity registered in their district of origin according to Bengali GI certification. The incorporation of visually near-identical subclass pairs, especially the pair of Muktagacha Monda (Brown vs. White) and Tangail Sondesh (Brown vs. White) make BD-GI-Desserts13 a very fine-grained classification standard with high inter-class similarity and naturally occurring intra-class variation.

3.2 Data Collection

The image acquisition was only taken at known retail outlets and GI-approved dessert stores in their respective areas of operation in Bangladesh, and each sample was an original product at the site where it received GI certification. The photographs were taken in situ at actual places of production and retail- they were not purchased by third parties or internet repositories to ensure that they were geographically and product-wise authentic. To enhance the diversity of the dataset and recreate natural imaging conditions, lighting conditions, viewing angles and background environments were purposely varied when collecting the data. All photographs were taken using consumer-grade smartphones (iPhone 13, iPhone 16, iPhone 16 Pro Max) under natural ambient light.

As shown in the Figure 3.1 geographical distribution of the 13 GI dessert classes across Bangladesh. All markers designate the primary source district of their respective class—Bogura Doi, Natore Kachagolla, Kushtia Tiler Khaza, Tangail Chomchom, Tangail Sondesh (Brown and White versions), Sherpur Chanapayesh, Cumilla Rashmalai, Gopalgonj Rosogolla, and others—directly reflecting the location prefix in the dataset's naming convention and underscoring the intrinsic link between each dessert and its GI origin district.

In the Figure 3.2 depicts sample on-site collection locations for three dataset classes, illustrating the authentic retail and production settings where data collection occurred.

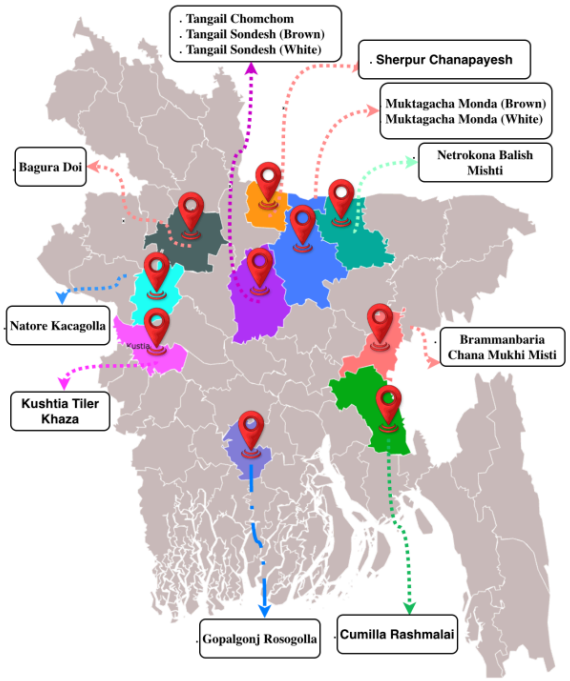


Figure 3.1: Geographic distribution of the 13 GI-certified dessert classes in BD-GI-Desserts13 at the district level. Each marker designates the primary source district of the respective product.

3.3 Dataset Organization and Statistics

Collected images were organized into a hierarchical folder structure as shown in Figure 3.3. At the top level, the repository is divided into three disjoint subsets: `train`, `val`, and `test`. Within each subset, 13 class-specific subdirectories follow the `Location_dessert[_variant]` naming convention—e.g., `Bogura_doi`, `Muktagacha_monda_brown`, `Tangail_sondesh_white`. The location prefix serves as a geographic provenance marker; the dessert suffix identifies the product type and, where applicable, visual or form variants (Brown/White). The complete dataset consists of 11,692 JPEG RGB images totaling approximately 1 GB of storage. All images are standardized to 512×512 pixels to provide uniform input resolution for deep learning models. The corpus is nearly balanced with each of the 13 classes containing approximately 900 images (range: 894–900 per class). Class-stratified sampling partitions the dataset into training (70%), validation (15%), and test (15%) subsets as shown in Table 3.1.



Figure 3.2: Typical on-site collection locations: (a) Cumilla Rashmalai, (b) Natore Kachagolla, and (c) Sherpur Chanapayesh.



Figure 3.3: Hierarchical folder structure of BD-GI-Desserts13 following the `Location_dessert[_variant]` naming convention.

A dataset-level integrity check confirmed all 11,692 files are readable, match the target 512× 512 resolution, and appear in exactly one partition. All images were manually annotated using the district/dessert naming convention with visual verification cross-checks to ensure labeling consistency. The presence of visually near-identical subclass pairs—Monda (Brown/White) and Sondesh (Brown/White)—makes BD-GI-Desserts13 a particularly challenging fine-grained recognition benchmark that generic food recognition datasets cannot adequately address.

3.4 Data Augmentation

Data augmentation was only used in the training subset offline to augment intra-class variability and strengthen the model to real world imaging variations. Validation and test subsets were not changed to make sure that the metrics of evaluation of performance are based on

raw,field-collected images. All the transformations were done with the help of Pillow library in Python. Table 3.1 summarizes the augmentation operations applied. Augmented images

Table 3.1: Image counts per class, augmentation parameters, and dataset distribution.

Class	Original	Data Augmentation Settings	Augmented	Train	Val	Test
Bogura_doi	312	Rotation: −20° to 20° Brightness: 0.7–1.3 Contrast: 0.7–1.3 Zoom: 0.8–1.0 Horiz. Flip: 0.5 Fill: nearest Rescale: 1/256	900	630	135	135
Brahmanbaria_chana_mukhi_misti	310		900	630	135	135
Cumilla_rashmalai	395		900	630	135	135
Gopalgonj_rosogolla	333		900	630	135	135
Muktagacha_monda_brown	326		899	629	134	136
Muktagacha_monda_white	301		900	630	135	135
Natore_kacagolla	301		894	625	134	135
Netrokona_balish_mishti	341		900	630	135	135
Sherpur_chanapayesh	372		899	629	134	136
Tangail_chomchom	331		900	630	135	135
Tangail_sondesh_brown	456		900	630	135	135
Tangail_sondesh_white	356		900	630	135	135
Kushtia_tiler_khaza	303		900	630	135	135
Total	**4137**		**11692**	**8183**	**1752**	**1757**

are stored alongside original samples in their class directories with transformation-specific filename suffixes: `_rotate.jpg` for geometric transformations, `_brightness.jpg` for photometric changes, and `_zoom.jpg` for spatial transformations. This naming convention enables clear differentiation between augmented and field-collected original samples at any pipeline stage.

3.5 GI Dessert Classes

The description and representative are given in Table 3.2 sample image of each of the 13 GI-certified dessert classes in BD-GI-Desserts13. The description of the dessert includes those

of each of the classes district of origin, important ingredient composition, typical texture and flavour profile, and GI registration year.

Table 3.2: Classes of traditional Bangladeshi sweets used in the dataset with descriptions and sample images.

Class Name	Description	Sample Image
Bogura Doi	Bogura Doi, also called Bograr Mishti Doi, is a thick, sweet yogurt from Bangladesh that has a creamy, custard-like texture. It's made by slowly boiling cow's milk and letting it set in traditional clay pots, which gives it a nice caramel-colored top and a balanced sweet-and-sour taste. It is a very popular dessert and even became an officially recognized regional product in 2023.	
Brahmanbaria Chhanamukhi Mishti	Brahmanbaria Chhanamukhi Mishti is also known as Chhanamukhi is a famous chhena-based confection from Brahmanbaria, eastern Bangladesh. It has a hard, crystalline sugar coating and a sweet, moist dairy center. Made from fresh chhena shaped into small quadrilateral pieces, cooked in sugar syrup, and dried, it gains a unique crunchy-yet-soft texture and rich milky flavor. Carrying approximately two centuries of regional heritage, it was registered as a GI product in 2024.	
Cumilla Rasmalai	Cumilla Rasmalai (also known as Cumillar Roshomalai) is a famous milk-based dessert from Cumilla, eastern Bangladesh. It has a soft, spongy texture and rich, sweet dairy flavor. Made from small chhena dumplings soaked in thick, slowly boiled milk (malai), it gains a velvety consistency and melts in the mouth. Registered as a GI product in 2024.	

Continued on next page

Table 3.2 continued from previous page

Class Name	Description	Sample Image
Gopalganj Rosogolla	Gopalganj Rosogolla (also known as Gopalganjer Rosogolla) is a famous syrupy dessert from Gopalganj, central Bangladesh. It has a soft, spongy texture and rich, sweet flavor. Made from fresh chhena kneaded into smooth balls and boiled in light sugar syrup, it gains perfectly porous consistency and melt-in-the-mouth quality. Registered as a GI product in 2024.	
Muktagacha Monda Brown	Muktagacha Monda (also known as Gopal Paler Monda) is a famous traditional sweet from Muktagacha, Mymensingh, northern Bangladesh. Known for its rich brown color, it has a soft, slightly grainy texture and a deep, earthy sweet flavor. Made from fresh chhena (curd) slowly cooked with jaggery (gur) instead of sugar and shaped into small, flat patties, it gains a unique melt-in-the-mouth consistency and a warm, molasses-like aroma. Registered as a GI product in 2024.	
Muktagacha Monda White	Muktagacha Monda (also known as Gopal Paler Monda) is a famous traditional sweet from Muktagacha, Mymensingh, northern Bangladesh. Known for its distinct white color, it has a soft, slightly grainy texture and a pure, rich dairy flavor. Made from fresh chhena (curd) slowly cooked with white sugar and shaped into small, flat patties, it gains a unique melt-in-the-mouth consistency and a delicate sweet aroma. Registered as a GI product in 2024.	

Continued on next page

Table 3.2 continued from previous page

Class Name	Description	Sample Image
Natore Kachagolla	Natore Kachagolla (also known as Natore-er Kachagolla) is a famous chhena-based sweet from Natore, northern Bangladesh. It has a soft, crumbly texture and a mildly sweet dairy flavor. Made by mixing fresh raw chhena with sugar and lightly cooking it without forming it into tight shapes, it gains a unique granular consistency and rich taste. Registered as a GI product in 2023.	
Netrokona Balish Mishti	Netrokona Balish Misti is a famous Geographical Indication (GI) sweet from Netrokona, northern Bangladesh. Shaped like a soft pillow, it is made from pure chhana (cottage cheese), milk, and sugar. It features a delicate spongy texture with a rich creamy malai layer on top. This melt-in-the-mouth sweet is a beloved symbol of Netrokona's culinary heritage. Registered as a GI product in 2025.	
Sherpur Chanapayesh	Sherpurer Chhanar Payesh is a famous milk-based dessert from Sherpur, northern Bangladesh. It has a soft, granular texture and a rich, sweet dairy flavor with a hint of cardamom. Made by forming fresh chhena into tiny balls, soaking them in light sugar syrup, and simmering them in thick, slowly boiled milk (kheer), it gains a wonderfully creamy and melt-in-the-mouth consistency. Registered as a GI product in 2024.	

Continued on next page

Table 3.2 continued from previous page

Class Name	Description	Sample Image
Tangail Chomchom	Tangail Chomchom (also known as Porabari Chomchom) is a famous traditional sweet from Porabari, Tangail, central Bangladesh. It has a firm, deeply caramelized outer crust and a soft, porous, syrup-drenched center. Made from fresh chhana slow-cooked in boiling sugar syrup and often rolled in dry mawa (crumbled milk solids), it gains a unique melt-in-the-mouth consistency and an intensely rich, sweet flavor. Registered as a GI product in 2024.	
Tangail Sondesh Brown	Tangail Sondesh (also known as Tangailer Jamurkir Sandesh) is a famous traditional sweet from Jamurki, Mirzapur, Tangail, central Bangladesh. Known for its rich brown color, it has a soft, melt-in-the-mouth texture and a deep, earthy sweet flavor with the distinct aroma of date palm jaggery. Made from fresh chhena slowly cooked with patali gur (date jaggery) after boiling high-quality cow's milk for hours, it gains a unique velvety consistency and a highly praised, warm molasses-like aroma. Carrying a legacy from the British era, it was registered as a GI product in 2025.	

Continued on next page

Table 3.2 continued from previous page

Class Name	Description	Sample Image
Tangail Sondesh White	Tangail Sondesh (also known as Tangailer Jamurkir Sandesh) is a famous traditional sweet from Jamurki, Mirzapur, Tangail, central Bangladesh. Known for its pristine white color, it has a soft, melt-in-the-mouth texture and a pure, rich dairy flavor with hints of cardamom. Made from fresh chhena slowly cooked with white sugar after boiling high-quality cow's milk for hours, it retains its bright white appearance while gaining a unique velvety consistency and a highly praised sweet aroma. Carrying a legacy from the British era, it was registered as a GI product in 2025.	
Kushtia Tiler Khaza	Kushtiar Tiler Khaja (also known as Tiler Khaja) is a famous traditional sweet snack from Kushtia, western Bangladesh. It has a crispy, flaky texture and a sweet, nutty sesame flavor. Made by boiling sugar or jaggery syrup, stretching it into thin layers, and coating it with roasted sesame seeds, it gains a uniquely crunchy consistency and delightful taste. Registered as a GI product in 2024.	

4 Methodology

4.1 The Methodological Framework

The study methodology provides a reproducible and systematic approach through a comprehensive pipeline that controls all steps of the work—from raw image acquisition to real-time model deployment—with the aim of creating a scientifically rigorous, provenance-conscious machine vision model for GI-certified desserts of Bangladesh.

The entire pipeline is divided into five consecutive stages,

- **Data Collection**: Multi-site field data collection where images of all 13 GI dessert classes are gathered from verified retail sources and certified dessert shops in their respective districts of origin in Bangladesh.
- **Data Preparation**: Manual quality screening, top-down organization of folders in "district/dessert" format, resizing to a standard size of 512×512 pixels (RGB color space), deletion of EXIF metadata, and stratified partitioning of classes into training (70%), validation (15%), and test (15%) subsets.
- **Data Augmentation**: Offline augmentation of the training subset with geometric and photometric changes to enhance model generalization.
- **Model Development**: Design, implementation, and training of a Custom Convolutional Neural Network (CCNN), transfer learning baselines, and traditional machine-learning classifiers.
- **Model Evaluation & Deployment**: Multi-metric analysis of all models and implementation of the most effective model in a web-based real-time recognition interface.

Each step is outlined in the subsections that follow.

4.2 Problem Formulation

The GI dessert recognition task is formalized as a multi-class supervised image classification problem. Considering an input image x of a Bangladeshi GI-certified dessert, the task is to learn a classification function $f : x \mapsto y$ that maps the input to one of $K = 13$ discrete class labels $y \in \mathcal{Y}$, where $\mathcal{Y}$ represents the official BD-GI-Desserts13 categories of GI desserts.

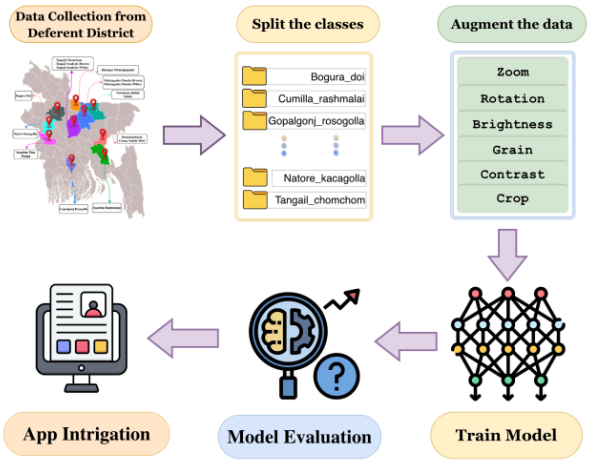


Figure 4.1: Methodological pipeline of GI dessert recognition.

The dataset is represented as a collection of labeled pairs $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$, where $x_i \in \mathbb{R}^{H \times W \times C}$ is the RGB image of the i-th sample (with height H, width W, and $C = 3$ color channels), and y_i is the corresponding ground-truth class label. The model is trained by minimizing the categorical cross-entropy loss on the training partition:

$$\mathcal{L}(\theta) = -\sum_{i} \sum_{j=1}^{K} y_{ij} \log(\hat{y}_{ij}) \tag{1}$$

where y_{ij} is the ground-truth one-hot encoded class indicator for class j and image i, $\hat{y}_{ij}$ denotes the predicted class probability from the Softmax layer, and θ is the set of learnable model parameters.

This classification problem is challenged by two primary sources of variability: (1) *inter-class similarity* between subvariants (most notably Muktagacha Monda Brown vs. White and Tangail Sondesh Brown vs. White), exhibiting minimal variations in color distribution, surface reflectance, and ingredient-based textural changes (e.g., jaggery vs. white sugar); and (2) *intra-class diversity* due to variations in illumination conditions, camera distance and angle, background environment, and natural batch-to-batch product variation. The data augmentation technique, outlined in Section 3.4, is specifically designed to build model resilience to these intra-class sources of variation.

4.3 Classification Models

There were six classification models that were first implemented and tested that were based on three architectural paradigms: a custom CNN (CCNN), pretrained transfer learning models

(EfficientNet-three traditional machine-learning classifiers (SVM, k-NN, and Deci- sion Tree). This breadth allows principled comparison of domain specific deep learning, general transfer learning, and embedding based classical strategies to the same fine-grained GI dessert recognition problem.

4.3.1 Custom Convolutional Neural Network (CCNN)

The CCNN is a specifically designed feed-forward deep learning network that offers an effective trade-off between performance and use of the GI dessert recognition problem. The design comprises of five convolutional blocks and a fully connected classification head.

The first two convolutional blocks have consecutive 3×3 Conv2D layers with same-padding, Batch Normalization, ReLU activation and 2×2 MaxPooling.. The depth of the filter gradually increases through the five blocks, 32, 64, 128, 256 and 512 and allows hierarchical learning of features with low level color and edge information in the initial layers to high level semantic textures and in deeper layers morphological patterns.

Input images are resized to 512×512 pixels in RGB format prior to network entry. The five 2×2 pooling operations gradually reduce the feature map resolution from 512×512 at the input to $16 \times 16 \times 512$ at the output of the fifth convolutional block. After the fifth convolutional block, a Spatial Dropout layer (rate = 0.25) is applied to reduce co-adaptation among feature maps. The output is then passed through a Flatten layer, converting the $7 \times 7 \times 512$ feature volume into a 25,088-dimensional vector. This is followed by a fully connected Dense layer of 500 units with ReLU activation, a second Dropout layer (rate = 0.40) for regularization, and a final Dense output layer of 13 units with Softmax activation for multi-class classification. The total trainable parameter count is 17,267,205.

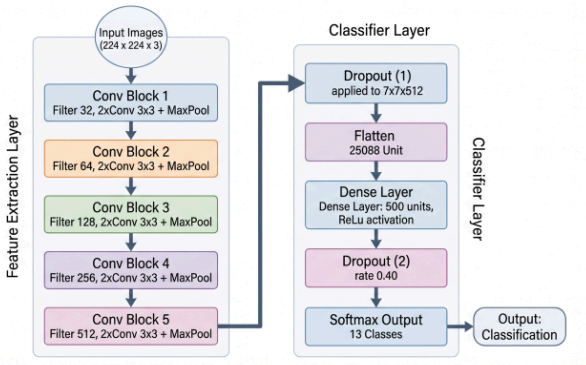


Figure 4.2: Architecture of the proposed Custom Convolutional Neural Network (CCNN) for GI dessert recognition

For a convolutional layer with input spatial dimension W_{in}, kernel size F, padding P, and

stride S, the output spatial dimension is:

$$W_{out} = \left\lfloor \frac{W_{in} - F + 2P}{S} \right\rfloor + 1 \tag{2}$$

After max-pooling with pool size F_p and stride S_p:

$$W' = \left\lfloor \frac{W_{in} - F_p}{S_p} \right\rfloor + 1 \tag{3}$$

The final class probabilities are computed by the Softmax activation:

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}, \quad K = 13 \tag{4}$$

where z_i denotes the pre-Softmax logit for class i.

4.3.2 EfficientNet-B3

EfficientNet-B3 is a transfer learning model pretrained with ImageNet weights. It employs a principled compound scaling strategy that concurrently upscales network depth, width, and input resolution using scaling coefficients:

$$\text{depth} = \alpha^{\phi}, \quad \text{width} = \beta^{\phi}, \quad \text{resolution} = \gamma^{\phi} \tag{5}$$

subject to the constraint:

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \tag{6}$$

where ϕ is the user-defined compound scaling coefficient. This principled scaling enables EfficientNet to achieve high accuracy with fewer parameters than one-dimensionally scaled models. EfficientNet-B3 is fine-tuned on BD-GI-Desserts13 using the same preprocessing pipeline and stratified data splits as the CCNN, with batch size 24 and learning rate of 1×10^{-6} over 30 epochs.

4.3.3 ShuffleNetV2

ShuffleNetV2 is a computationally efficient transfer learning model pretrained on ImageNet weights, specifically designed for inference with limited resources. Its efficiency is achieved through channel split and channel shuffle operations: an input tensor X is split into two branches X_1 and X_2, where one branch retains an identity mapping while the other undergoes lightweight transformation $F(X_2)$. The results are concatenated and channel shuffle is applied to enable cross-channel information communication:

$$\text{Output} = \text{ChannelShuffle}\left(\text{Concat}(X_1, F(X_2))\right) \tag{7}$$

ShuffleNetV2 is fine-tuned on BD-GI-Desserts13 with batch size 24 and learning rate 1×10^{-5} over 30 epochs.

4.3.4 Traditional Machine Learning Classifiers

Three traditional machine-learning classifiers—Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), and Decision Tree—are applied as baseline comparators. The selection of CNN as the main structure is justified by the vast comparative literature in food authentication. Magdas et al. [53] demonstrated that CNN models always performed better than SVM, PLS-DA and k-NN in the food adulteration detection tasks - CNN was 97% accurate as compared to 91% accuracy of LS-SVM and 79% accuracy of PLS-DA on the same spectroscopic data. Feng et al. [49] also reported that deep CNN was better than SVM and logistic regression in classifying food varieties based on hyperspectral imaging. These results alone support the hypothesis that a deep-learning end-to-end approach will be more successful compared to the traditional ML baselines tested in this paper. Rather than training on raw pixel data, all three are trained on high-dimensional deep feature embeddings extracted from the penultimate fully connected layer of the trained CCNN. For an input image x, the CCNN feature extractor $f(x)$ produces a d-dimensional embedding:

$$z = f(x), \quad z \in \mathbb{R}^d \tag{8}$$

These embeddings encode discriminative color, texture, and structural properties of each dessert image.

Support Vector Machine (SVM)

The SVM classifier finds the optimal separating hyperplane in the embedding space that maximizes the inter-class margin:

$$y_i(w^T z_i + b) \geq 1 \tag{9}$$

where w is the weight vector and b is the bias term. A one-versus-rest strategy is adopted to extend the binary SVM formulation to the 13-class setting. On the BD-GI-Desserts13 test set, SVM achieves an overall accuracy of 91.01% with a macro-averaged F_1-score of 0.9097. It delivers well-balanced performance across classes, with particularly strong results for Netrokona Balish Mishti ($F_1 = 0.9963$) and Muktagacha Monda Brown ($F_1 = 0.9815$). Performance degrades for visually similar pairs such as Gopalgonj Rosogolla ($F_1 = 0.7863$) and Natore Kacagolla ($F_1 = 0.8529$), where high inter-class similarity in embedding space reduces hyperplane separability.

k-Nearest Neighbors (k-NN)

The k-NN classifier assigns a query embedding z_q to the majority label among its k nearest neighbors, using Euclidean distance as the proximity measure:

$$d(z_i, z_j) = \|z_i - z_j\|_2 \tag{10}$$

k-NN attains the highest accuracy among the three traditional baselines at 91.18%, with a macro-averaged F_1-score of 0.9119. Strong per-class results are observed for Netrokona Balish Mishti ($F_1 = 0.9925$), Tangail Chomchom ($F_1 = 0.9818$), and Muktagacha Monda White ($F_1 = 0.9695$), reflecting well-clustered CCNN embeddings for these categories. However, k-NN is sensitive to intra-class spread and background variation; Brammanbaria Chana Mukhi Misti ($F_1 = 0.8312$) and Cumilla Rashmalai ($F_1 = 0.8571$) yield lower F_1-scores, consistent with greater visual variability within those classes.

Decision Tree

The Decision Tree classifier recursively partitions the embedding space with binary splits that minimize weighted Gini impurity:

$$\text{Gini}(S) = 1 - \sum_{i=1}^{K} p_i^2 \tag{11}$$

where p_i is the proportion of samples in subset S belonging to class i. The Decision Tree achieves the lowest accuracy of the three baselines at 86.74% with a macro-averaged F_1-score of 0.8674. While it performs reasonably on visually distinct classes such as Netrokona Balish Mishti ($F_1 = 0.9670$) and Tangail Chomchom ($F_1 = 0.9442$), it struggles with fine-grained texture discrimination. Classes with subtle surface-color or syrup-coating variations—Brammanbaria Chana Mukhi Misti ($F_1 = 0.7726$), Natore Kacagolla ($F_1 = 0.7605$), and Gopalgonj Rosogolla ($F_1 = 0.7770$)—yield notably lower scores, highlighting the model's limited capacity to capture continuous texture gradients through axis-aligned splits in high-dimensional space.

4.4 Training Configuration

All deep learning models are trained using categorical cross-entropy loss and the Adam optimizer via backpropagation. The architecture-specific training hyperparameters are outlined in Table 4.1.

Table 4.1: Training Configuration for Deep Learning Models.

Model	Batch Size	Learning Rate	Epochs	Regularization
CCNN	32	5×10^{-4}	30	Dropout(0.25, 0.40), L2 1×10^{-4}
ShuffleNetV2	24	1×10^{-5}	30	Architecture built-in
EfficientNet-B3	24	1×10^{-6}	30	Architecture built-in

Validation accuracy at the end of each epoch monitors training progress. Model checkpointing saves the weights producing the best validation accuracy for final test set evaluation. Early stopping terminates training if validation accuracy plateaus for 5 consecutive epochs, preventing overfitting by halting optimization before the model memorizes training samples.

The CCNN classification head employs Dropout regularization (rate = 0.25 and 0.40) and L2 weight decay (10^{-4}) to discourage co-adaptation of feature detectors and enhance generalization

to unseen samples. Training the three classical ML classifiers involves a single-pass fitting step on the pre-extracted CCNN embedding matrix, requiring no epoch-based optimization or GPU resources.

4.5 Evaluation Strategy

Model performance is assessed using a multi-metric evaluation framework on the withheld test set. Overall classification accuracy reports the percentage of correctly classified test instances. Per-class and macro-averaged precision, recall, and F_1-score provide class-sensitive views of predictive performance:

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} \tag{4.1}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{4.2}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{4.3}$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4.4}$$

Recall holds particular operational significance for GI authentication contexts: false negatives (failing to identify true GI desserts) carry direct commercial and legal implications favoring local manufacturers. High recall values indicate effective minimization of this error type. Accuracy remains crucial to prevent counterfeit or non-GI products from incorrectly passing authentication.

All six models are evaluated on the test set, with confusion matrices generated to reveal detailed per-class classification patterns and inter-class confusions, enabling systematic identification of misclassification trends. Gradient-weighted Class Activation Mapping (Grad-CAM) visualizations qualitatively validate that the CCNN focuses classification decisions on discriminative dessert-specific surface textures—such as caramelization patterns, syrup boundary pooling, and chhena granularity—rather than background artifacts or non-discriminative image regions.

4.6 Software and Tools

Python 3.x is used to implement data preprocessing, evaluation pipelines and all model development. TensorFlow and Keras are used to build and train deep learning architectures with high-level APIs to design, train, and evaluate convolutional neural networks. The Pillow library is used to convert image color space, resize and expand data offline. The scikit-learn library uses the classical ML classifiers (SVM, k-NN, Decision Tree), stratified data splitting and classification metric calculation. NumPy and pandas deal with managing data in the pipeline. Training curves and confusion matrices are created in Matplotlib and seaborn. The Grad-CAM visualizations are generated using the standard Keras gradient computation utilities.

4.7 Ethical Considerations

Data collection and model development were done in adherence to the appropriate ethical standards. The BD-GI-Desserts13 data was shot in known retail sites - certified sweet shops and dessert stores in various places in Bangladesh. None of the research stages included data collection involving any human subjects, and none of the research stages examined any personal data or sensitive information.

All the images had been checked manually before being included to make sure that no identifiable face of a human or any other personal information was visible, and the privacy concerns were eliminated by the dataset. The publicly-available BD-GI-Desserts13 dataset on Kaggle has no restrictions to access, which adheres to the spirit of open science and allows the replication of GI product authentication studies in Bangladesh.

Notably, the suggested machine vision is a tool of automated recognition, rather than a legally binding GI certification tool. The only authorities that can make official GI authentication and product legality determinations are specific entities of the government and regulatory bodies. The system helps consumers, retailers and researchers to determine the type of GI desserts; its estimates are not a legal certification and should not be used in business or in regulatory practices.

5 Experimental Results and Models Comparative Analysis

The chapter is a quantitative analysis of the six classification models on the BD-GI-Desserts13 data. The experimental design, measures of class-wise performance, confusion-matrix diagnostics and comparison with the literature in the field are reported. Each of the models was tested on the same held-out test set of 1,757 images so that there would be a fair like-for-like comparison.

5.1 Experimental Setup

All three deep learning models were trained on cloud-based GPU infrastructure using identical preprocessing: images were resized to 512×512 pixels in RGB format, followed by per-channel normalization and a fixed augmentation protocol comprising random rotation, horizontal flipping, and brightness adjustment. All models were optimized with the Adam optimizer and categorical cross-entropy loss:

$$\mathcal{L} = -\sum_{i=1}^{K} y_i \log(\hat{y}_i), \quad K = 13 \tag{12}$$

where y_i is the ground-truth one-hot label and $\hat{y}_i$ is the predicted class probability. Parameters were updated via backpropagation:

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla \mathcal{L}_t \tag{13}$$

Training ran for a maximum of 30 epochs with early stopping based on validation accuracy and model checkpointing to retain the best-performing weights. The architecture-specific hyperparameters are detailed in Table 4.1.

The three traditional machine learning classifiers—SVM, k-NN, and Decision Tree—required no iterative GPU-based training. After fully training the custom CNN, its penultimate fully connected layer was used as a fixed feature extractor. The 8,183 training images were passed through the frozen network to produce d-dimensional embedding vectors, denoted as $z = f(x)$, where $z \in \mathbb{R}^d$. These embeddings were then used to train each classical classifier directly. Hyperparameters, including the SVM regularization parameter C, the k-NN neighbor count k, and the Decision Tree maximum depth, were selected using the validation embedding set. Final evaluation was conducted exclusively on the held-out test embeddings.

5.2 Deep Learning Model Performance

5.2.1 Class-wise Quantitative Results

As shown in Table 5.1, the class-wise precision, recall, and F_1-score for Custom CNN, ShuffleNetV2, and EfficientNet-B3 on the held-out test set. The Custom CNN achieved the highest overall accuracy of 99.37%, outperforming ShuffleNetV2 (98.12%) and EfficientNet-B3 (93.17%).

Table 5.1: Class-wise Performance of Deep Learning Models on BD-GI-Desserts13

Class	Custom CNN			ShuffleNetV2			EfficientNet-B3		
	Prec.	Rec.	F_1	Prec.	Rec.	F_1	Prec.	Rec.	F_1
Bogura Doi	1.0000	0.9333	0.9655	0.9574	1.0000	0.9783	0.9846	0.9481	0.9660
Brahmanbaria Chhanamukhi Mishti	1.0000	0.9852	0.9925	0.9922	0.9407	0.9658	0.9669	0.8667	0.9141
Cumilla Rashmalai	1.0000	1.0000	1.0000	0.9310	1.0000	0.9643	0.9328	0.9259	0.9294
Gopalgonj Rosogolla	1.0000	1.0000	1.0000	0.9783	1.0000	0.9890	0.9710	0.9926	0.9817
Muktagacha Monda Brown	1.0000	1.0000	1.0000	0.9926	0.9926	0.9926	0.8766	0.9926	0.9310
Muktagacha Monda White	1.0000	1.0000	1.0000	0.9926	0.9926	0.9926	0.9911	0.8222	0.8988
Natore Kachagolla	1.0000	1.0000	1.0000	0.9848	0.9630	0.9738	0.9712	1.0000	0.9854
Netrokona Balish Mishti	1.0000	1.0000	1.0000	1.0000	0.9926	0.9963	0.9580	0.8444	0.8976
Sherpur Chanapayesh	1.0000	1.0000	1.0000	0.9627	0.9485	0.9556	0.8767	0.9412	0.9078
Tangail Chomchom	1.0000	1.0000	1.0000	0.9926	1.0000	0.9963	0.9420	0.9630	0.9524
Tangail Sondesh Brown	1.0000	1.0000	1.0000	1.0000	0.9926	0.9963	0.9504	0.8519	0.8984
Tangail Sondesh White	0.9854	1.0000	0.9926	0.9779	0.9852	0.9815	0.8037	0.9704	0.8792
Kushtia Tiler Khaza	0.9375	1.0000	0.9677	1.0000	0.9481	0.9734	0.9437	0.9926	0.9675
Overall Accuracy	**99.37%**			**98.12%**			**93.17%**		
Macro Avg F_1	**0.9937**			**0.9812**			**0.9315**		

The Custom CNN achieved near-perfect F_1-scores ($\geq$0.99) across 10 of 13 classes, including perfect scores ($F_1 = 1.0000$) for Cumilla Rashmalai, Gopalgonj Rosogolla, Muktagacha Monda (both variants), Natore Kachagolla, Netrokona Balish Mishti, Sherpur Chanapayesh, Tangail Chomchom, and Tangail Sondesh Brown. This class-wide consistency confirms high intra-class compactness and robust inter-class separability across the full GI dessert spectrum.

ShuffleNetV2 achieved 98.12% accuracy with strong F_1-scores on most categories, but exhibited marginal degradation on visually similar subclasses — specifically the brown/white variants of Muktagacha Monda and Tangail Sondesh — where the lightweight channel-shuffle mechanism could not fully resolve fine textural gradients under varied imaging conditions.

EfficientNet-B3 recorded the lowest classification accuracy among the deep learning models at 93.17%, with notable Recall deficits observed for Netrokona Balish Mishti (Recall = 0.8444) and Muktagacha Monda White (Recall = 0.8222). Despite its compound-scaled architecture and the integration of Squeeze-and-Excitation (SE) attention modules, the generic ImageNet pretrained feature hierarchy failed to transfer effectively to the fine-grained visual domain of Bangladeshi GI desserts.

5.2.2 Confusion Matrix Analysis

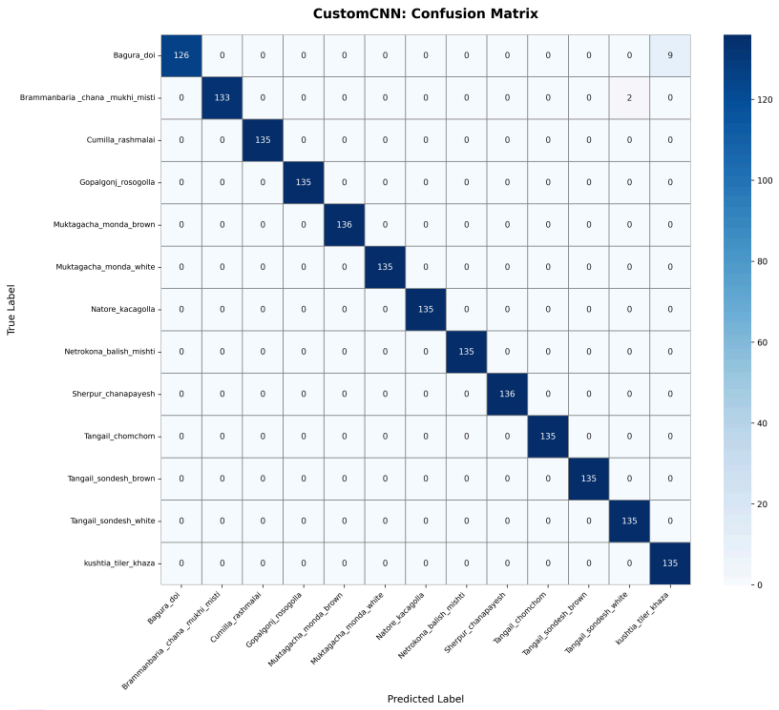
The confusion matrices of the three deep learning models are presented in Figures 5.1, 5.2, and 5.3. All three architectures exhibit strong diagonal dominance across all 13 GI dessert classes, indicating effective multi-class separation overall. However, notable differences exist in the magnitude and distribution of misclassification errors among the three models.

Custom CNN:

As illustrated in Figure 5.1, the Custom CNN achieves near-perfect classification across all 13 classes. Ten out of thirteen classes–including Cumilla_rashmalai, Gopalgonj_rosogolla, Muktagacha_monda_brown, Muktagacha_monda_white, Natore_kacagolla, Netrokona_balish_mishti, Tangail_chomchom, Tangail_sondesh_brown, Tangail_sondesh_white, and Sherpur_chanapayesh—register zero misclassifications, producing a perfectly clean diagonal. The only observable errors occur in two classes: Bagura_doi (126 correct, 9 misclassified as Kushtia_tiler_khaza) and Brammanbaria_chana_mukhi_misti (133 correct, 2 misclassified as Tangail_sondesh_white). The Bagura_doi, Kushtia_tiler_khaza confusion can be visually explained, as both desserts share a pale, creamy surface appearance that may seem similar under certain lighting conditions during image acquisition. Overall, the Custom CNN confusion matrix is the most concentrated and cleanest among all three models, indicating superior generalization capability on fine-grained GI dessert classes.

ShuffleNetV2:

The confusion matrix of ShuffleNetV2, shown in Figure 5.2, demonstrates strong overall performance but with a broader and more distributed error pattern compared with the Custom CNN. The most significant misclassification occurs in Sherpur_chanapayesh, where 7 out of 135 samples are incorrectly predicted as Cumilla_rashmalai. This confusion is visually understandable, as both desserts are milk-based, soft-textured, and syrup-soaked, making them among the most difficult pairs to distinguish in the dataset. Similarly, Natore_kacagolla and Kushtia_tiler_khaza record 5 and 6 errors predicted as Sherpur_chanapayesh and Bagura_doi, respectively, again reflecting inter-class visual similarity among pale-toned dessert variants. Brammanbaria_chana_mukhi_misti shows the broadest error spread, with 3 samples misclassified as Cumilla_rashmalai, 3 as Gopalgonj_rosogolla, and 2 as Tangail_sondesh_white. The Muktagacha_monda brown/white subclass pair produces one bidirectional confusion in each direction (1 brown predicted as white and 1 white predicted as brown), which is expected given the subtle chromatic distinction between these visually similar variants. Despite these errors, ShuffleNetV2 maintains strong diagonal dominance, with 8 out of 13 classes achieving 133 or more correct predictions.



123 Figure 5.1: Confusion matrix of the Custom CNN on the BD-GI-Desserts13 test set.

EfficientNet-B3:

In the Figure 5.3 illustrated the widest distribution of errors among all three models, consistent with the lower overall test accuracy of EfficientNet-B3 (93.17%). The most severe misclassification occurs in Tangail_sondesh_brown, where 20 out of 135 samples are incorrectly predicted as Tangail_sondesh_white, indicating that EfficientNet-B3 struggles to consistently differentiate the subtle color-based boundary between the brown and white Sondesh subclass variants. Muktagacha_monda_white records 18 confusions with Muktagacha_monda_brown, further suggesting that brown/white intra-class variant pairs are among the most challenging classification frontiers for transfer learning architectures. Cumilla_rashmalai produces 10 misclassifications directed toward Sherpur_chanapayesh, reflecting the same milk-based visual similarity observed in ShuffleNetV2. The broadest error spread appears in Brammanbaria_chana_mukhi_misti, where 9 samples are misclassified as Tangail_sondesh_white and 5 as Sherpur_chanapayesh. Netrokona_balish_mishti also shows notable dispersion, with 7 samples predicted as Tangail_chomchom and 7 as Kushtia_tiler_khaza, suggesting that EfficientNet-B3's pretrained ImageNet represen-

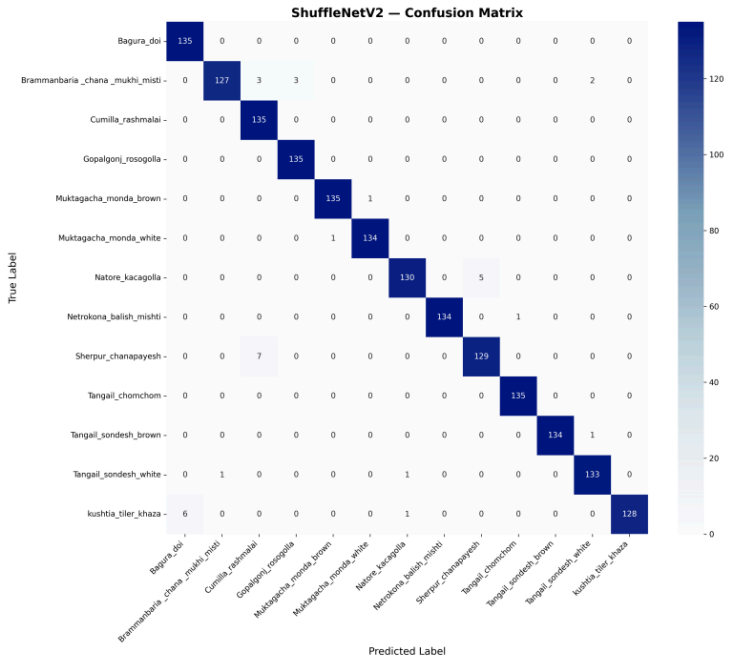


Figure 5.2: Confusion matrix of ShuffleNetV2 on the BD-GI-Desserts13 test set.

tations do not adequately capture the pillow-shaped morphological characteristics unique to Netrokona_balish_mishti. These findings indicate that while EfficientNet-B3 performs well on visually distinct classes, its performance declines on fine-grained subclass pairs and structurally similar dessert categories.

Across all three models, the most persistent sources of misclassification are the brown/white subclass pairs of Muktagacha_monda and Tangail_sondesh, as well as the milk-based dessert cluster comprising Cumilla_rashmalai, Sherpur_chanapayesh, and Brammanbaria_chana_mukhi_misti. These errors do not indicate systematic model failure; rather, they reflect the inherent fine-grained visual ambiguity of these class pairs, which constitute the most challenging discrimination task in BD-GI-Desserts13. The Custom CNN consistently outperforms both transfer learning models on these difficult pairs, demonstrating that its end-to-end domain-specific training yields more discriminative feature representations than pretrained ImageNet weights adapted to this task.

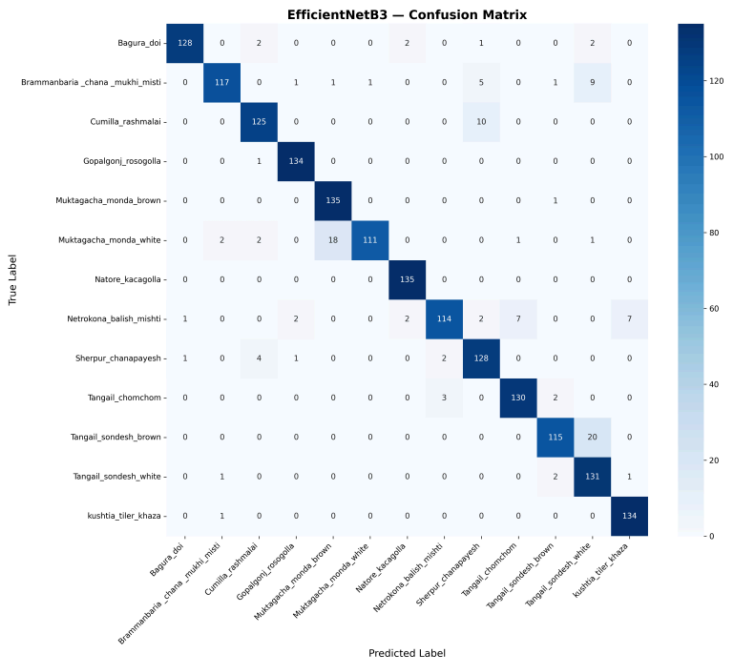


Figure 5.3: Confusion matrix of EfficientNet-B3 on the BD-GI-Desserts13 test set.

5.2.3 Training and Validation Curves

The loss and accuracy curves of all three deep learning models, i.e., Custom CNN, ShuffleNetV2, and EfficientNet-B3 are plotted during 30 training epochs. Together, these curves give diagnostic information on the convergence behaviour, generalisation capacity and the propensity of each model to overfitting of the BD-GI-Desserts13 data.

Custom CNN:

The training and validation curves for the Custom CNN are illustrated in Figure 5.4. gives the training and validation curves of the Custom CNN. The model has a very volatile initial training phase, with a sharp peak of validation loss of about 14.0 at epoch 1 and then decreases very quickly. This instability can be explained by random weight initialisation and the fine-grainedness of the classification problem. Although this is fluctuating, both training and validation loss stabilize after epoch 5 and go to zero after epoch 20. Likewise, training accuracy reaches about 99 -100 percent by the last epoch, whereas validation accuracy, which was very variable at the first ten epochs, reaches a stabilization point above 99 percent after epoch

20. There is a close correspondence between training and validation accuracy in later epochs indicates strong generalization and minimal overfitting, confirming the effectiveness of dropout regularization and the early stopping strategy employed in the CCNN architecture.

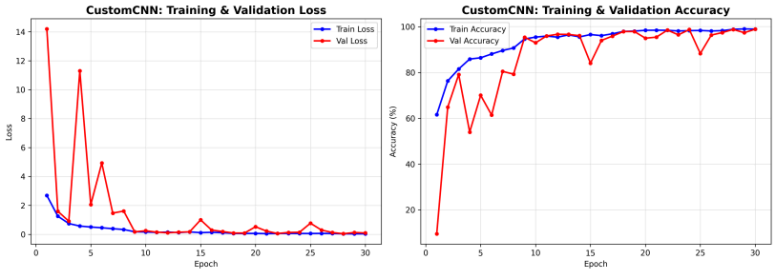


Figure 5.4: Training and validation loss (left) and accuracy (right) curves of the Custom CNN over 30 epochs.

ShuffleNetV2:

As illustrated in Figure 5.5, the training dynamics of ShuffleNetV2 contrast notably with those of the Custom CNN. ShuffleNetV2 exhibits a very smooth and stable convergence trajectory from as early as the first epoch. The training and validation loss decrease near-simultaneously, falling from approximately 2.55 to 1.50 within the first 30 epochs, with no apparent divergence between the two curves. Similarly, accuracy increases sharply to 41% at epoch 1, reaches 90% by epoch 5, and stabilizes between 97% and 98% throughout the final epochs. The high consistency between training and validation accuracy across all 30 epochs demonstrates that ShuffleNetV2 generalizes effectively to this dataset. Nonetheless, its ultimate accuracy ceiling of approximately 98.12% remains lower than that of the Custom CNN, suggesting that the pretrained ImageNet feature representations, although computationally efficient, may not be fully optimized for the fine-grained visual characteristics of GI-certified Bangladeshi desserts.

EfficientNet-B3:

The convergence behaviour of EfficientNet-B3 can be seen in Figure 5.6. The model achieves the lowest initial accuracy between the three models, initially at around 13% at the first epoch, and converges slowly compared to the Custom CNN and ShuffleNetV2. The loss is slowly reduced between 2.52 and 1.80 during the training process and accuracy increases between 13 percent and about 91-92 percent at the end of the training epoch. Even though the training and validation curves are well-correlated during the training process, which is a sign of stable generalization, the lower convergence rate, and ultimate accuracy of 93.17% point to the

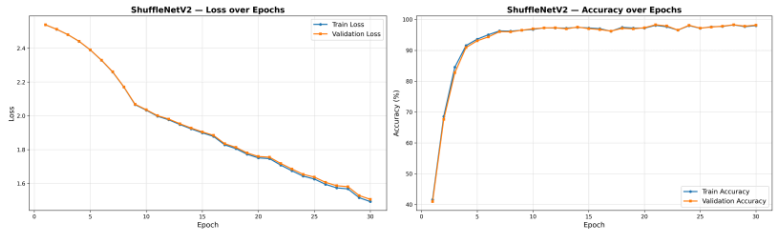


Figure 5.5: Training and validation loss (left) and accuracy (right) curves of ShuffleNetV2 over 30 epochs.

possibility that more epochs or a more fined learning rate schedule is needed to ensure that EfficientNet-B3 adapts more to the visual features of GI dessert classes specific to the domain.

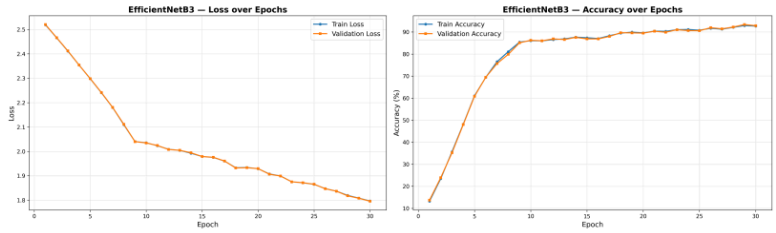


Figure 5.6: Training and validation loss (left) and accuracy (right) curves of EfficientNet-B3 over 30 epochs.

In all three models, there is no significant overfitting (evidenced by the similarity of training and validation curves). The Custom CNN is volatile at the initial epochs, but in the end, it attains the best accuracy and best convergence at the terminal stage. ShuffleNetV2 is the most stable and smooth in terms of training. The slowest converging and lowest final accuracy of the three models indicate that longer training or additional hyperparameter optimization can enhance its future performance of EfficientNet-B3.

5.2.4 Grad-CAM Interpretability

In the Figure 5.7 shows Gradient-weighted Class Activation Mapping (Grad-CAM) visualizations of the Custom CNN. These activation maps are overlaid on the original dessert pictures with the red and orange areas representing the maximal model attention and the blue and green ones represent low activation. The visualization results are good indications that the Custom CNN has acquired semantically-meaningful and domain-relevant features as opposed to background artifacts or spatially irrelevant areas.

Correctly Classified Classes.

For most correctly predicted classes marked in green in Figure 5.7, the model demonstrates highly focused and class-relevant attention. For *Bagura_doi*, activation is strongly concentrated on the central creamy yogurt surface inside the clay pot, particularly the smooth caramel-coloured top layer that is visually distinctive of this GI dessert. For *Brammanbaria_chana_mukhi_misti*, the heat map highlights the sharp angular edges and crystalline sugar-coated boundaries of the quadrilateral pieces. In *Cumilla_rashmalai*, activation broadly covers the spongy chhena dumpling body soaked in thick malai, indicating that the model relies on the characteristic soft and milk-saturated surface texture. For *Gopalgonj_rosogolla*, attention is focused on the rounded spongy ball surfaces, especially the porous outer crust and syrup-reflective edges that define this class morphologically.

For *Muktagacha_monda_brown*, Grad-CAM emphasizes the browned curved upper surface and caramelized jaggery-coloured boundary, which serve as the main visual discriminators from the white variant. In contrast, for *Muktagacha_monda_white*, attention shifts toward the pale flat surface center and lighter edge profile, confirming that the model successfully learns the subtle intra-class colour distinction between these visually similar subclass variants. *Natore_kacagolla* shows dense activation across the granular and crumbly chhena surface, consistent with the loosely formed texture that makes this dessert unique. *Netrokona_balish_mishti* produces strong activation over the pillow-shaped body contour and the cream layer on its upper surface, confirming focus on the defining morphological structure of this dessert. For *Kushtia_tiler_khaza*, the heat map is evenly distributed across the flat rectangular sesame-embedded surface, reflecting the distinctive tiled and layer-pressed form of this class. The activation map of *Tangail_sondesh_brown* concentrates on clustered rounded brown mound forms and caramelized surface boundaries, whereas *Tangail_sondesh_white* focuses on the pale rounded tops, confirming reliable discrimination between the Sondesh subclass pair through colour contrast and surface tone.

Misclassified Instance.

One misclassification is visible in Figure 5.7, where *Sherpur_chanapayesh* is incorrectly predicted as *Cumilla_rashmalai*, as indicated by the red label. The corresponding activation map reveals diffuse and unfocused heat spread across the image rather than the tight, structurally targeted attention observed in correctly classified samples. This diffusion suggests that the model is unable to identify a sufficiently discriminative local feature to separate Sherpur_chanapayesh from Cumilla_rashmalai. This confusion is visually understandable, as both desserts are milk-based, soft-textured, chhena-centered sweets immersed in thick dairy liquid, producing highly similar chromatic and textural appearances under certain imaging conditions. This single misclassification is consistent with the confusion matrix results in Figure 5.1, where the Sherpur_chanapayesh–Cumilla_rashmalai pair was identified as one of the few error sources of the Custom CNN.

The Grad-CAM analysis confirms that the Custom CNN has successfully learned class-

discriminative and origin-aware visual representations for BD-GI-Desserts13. In eleven out of twelve visualized classes, the model's attention is tightly concentrated on the most visually distinctive features of each GI dessert, including surface texture, morphological boundary, colour tone, and structural form, rather than on background containers, props, or spatially irrelevant regions. This attention pattern validates the suitability of the model for real-world GI dessert authentication, where reliable focus on product-specific visual characteristics is essential for accurate provenance verification.

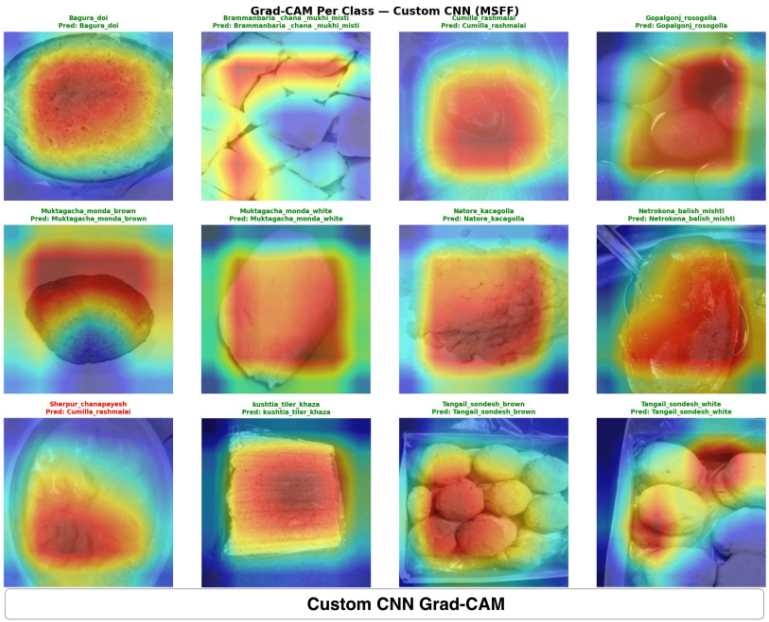


Figure 5.7: Grad-CAM activation maps for Custom CNN predictions across all GI dessert categories. Green labels indicate correct predictions; red labels indicate misclassified instances. Hot colors (red/orange) represent high activation regions; cool colors (blue/green) represent low activation.

5.3 Traditional Machine Learning Baseline Performance

The three traditional ML classifiers were evaluated on deep feature embeddings extracted from the Custom CNN's penultimate layer. Their class-wise results are summarized in Table 5.2. The k-NN classifier achieved the highest traditional ML accuracy at 91.18%, followed by SVM at 91.01% and Decision Tree at 86.74%.

The three classic ML classifiers were tested on deep feature embeddings of the penultimate layer of the Custom CNN. Table 5.2 summarizes their results according to their classes. The k-NN classifier had the highest traditional ML accuracy of 91.18%, then SVM and Decision Tree with 91.01% and 86.74% respectively.

Table 5.2: Class-wise Performance of Traditional ML Models on BD-GI-Desserts13

Class	SVM			k-NN			Decision Tree		
	Prec.	Rec.	F_1	Prec.	Rec.	F_1	Prec.	Rec.	F_1
Bogura Doi	0.9085	0.9556	0.9314	0.8313	0.9852	0.9017	0.8359	0.7926	0.8137
Brahmanbaria Chhanamukhi Mishti	0.9032	0.8296	0.8649	1.0000	0.7111	0.8312	0.7535	0.7926	0.7726
Cumilla Rashmalai	0.8477	0.9481	0.8951	0.9903	0.7556	0.8571	0.9179	0.9111	0.9145
Gopalgonj Rosogolla	0.8110	0.7630	0.7863	1.0000	0.7556	0.8608	0.7552	0.8000	0.7770
Muktagacha Monda Brown	0.9852	0.9779	0.9815	0.9433	0.9779	0.9603	0.8636	0.8382	0.8507
Muktagacha Monda White	1.0000	0.9259	0.9615	1.0000	0.9407	0.9695	0.8803	0.9259	0.9025
Natore Kachagolla	0.8467	0.8593	0.8529	0.7258	1.0000	0.8411	0.7812	0.7407	0.7605
Netrokona Balish Mishti	1.0000	0.9926	0.9963	1.0000	0.9852	0.9925	0.9565	0.9778	0.9670
Sherpur Chanapayesh	0.9344	0.8382	0.8837	0.7545	0.9265	0.8317	0.8824	0.8824	0.8824
Tangail Chomchom	0.9220	0.9630	0.9420	0.9643	1.0000	0.9818	0.9478	0.9407	0.9442
Tangail Sondesh Brown	0.9000	0.9333	0.9164	0.9841	0.9185	0.9502	0.8955	0.8889	0.8922
Tangail Sondesh White	0.9058	0.9259	0.9158	0.9919	0.9111	0.9498	0.8777	0.9037	0.8905
Kushtia Tiler Khaza	0.8794	0.9185	0.8986	0.8750	0.9852	0.9268	0.9370	0.8815	0.9084
Overall Accuracy	**91.01%**			**91.18%**			**86.74%**		
Macro Avg F_1	**0.9097**			**0.9119**			**0.8674**		

Although discriminative CNN embeddings were used, all three classical classifiers were significantly worse than the end-to-end deep learning models, achieving high F1 scores on well-clustered categories (e.g., Netrokona Balish Mishti at $F_1 = 0.9925$), but had a devastating Recall loss in Cumilla Rashmalai (Recall = 0.7556) and SVM showed the most homogenous distribution of errors among all 13 classes, and Decision Tree had the most scattered pattern of misclassification errors, which proved that axis-aligned recursive splits are not able to represent the intricate non-linear textural relations that GI dessert subclasses bear.

5.4 Cross-Model Comparative Analysis

Table 5.3 summarizes the total accuracy and macro-means of F_1-scores of all six models. The Custom CNN is the obviously best-performing between the two metrics, with a macro F_1 of 0.9937 — 12.6% more than the highest traditional classifier and 1.25 points better than the closest deep learning baseline.

One of the major findings is that it is observed that there is a consistent performance hierarchy in the deep learning and traditional ML groups. The domain-specific Custom CNN, in the deep learning category, achieves better performance than the transfer learning baselines, and the alignment of the architecture to the target visual domain is more beneficial than the size of pretrained features. Embedding-based classifiers within the traditional ML group affirm that

Table 5.3: Overall Performance Summary of All Six Models

Model	**Accuracy**	**Macro F_1**	**Primary Limitation**
Custom CNN	99.37%	0.9937	Minor errors on extreme visual similarity
ShuffleNetV2	98.12%	0.9812	Fine textural gradients in similar subclasses
EfficientNet-B3	93.17%	0.9315	Generic ImageNet features; Recall gaps
k-NN	91.18%	0.9119	Euclidean distance insufficient for overlap classes
SVM	91.01%	0.9097	No probabilistic output; bounded by embedding
Decision Tree	86.74%	0.8674	Axis-aligned splits; poor non-linear separation

CNN representations are discriminative to the extent that they can be used with simpler models, although end-to-end gradient-based training is necessary to solve the subtleties of intra-class differences.

The difference in the performance of EfficientNet-B3 (93.17%) and the Custom CNN (99.37%) is especially high, when considering the trade-off of Precision and Recall in GI authentication. EfficientNet-B3 failed to find 15–18% of real cases in Netrokona Balish Mishti and Muktagacha Monda White - a regulatory burden in any real-life certification process. Instead, the Custom CNN scored Precision and Recall of over 0.93 in all 13 classes with the highest commercially sensitive GI categories such as Cumilla Rashmalai, Gopalgonj Rosogolla, and Tangail Chomchom scoring perfect (1.0000) scores.

The experimental findings prove the proposed Custom CNN to be the most accurate, robust and deployable in this study. It scored better than all five baseline models in all the reported metrics - the overall accuracy, macro F_1-score, class-wise Precision, and class-wise Recall. Its better diagonal concentration is verified by the confusion matrices and the discriminative dessert-specific localization of its feature localization is confirmed by Grad-CAM activa- tions.

Traditional ML classifiers, while benefiting from rich CNN embeddings, confirmed the irreplaceable advantage of end-to-end gradient-based optimization for fine-grained GI dessert recognition. These findings provide a solid empirical foundation for deploying the Custom CNN as the core inference engine in real-time Bangladeshi GI dessert authentication systems.

6 Web-Based GI Dessert Recognition Framework

To demonstrate the practical feasibility of the proposed recognition system, the trained Custom CNN model was deployed into a responsive full-stack web application capable of performing real-time inference of Bangladeshi GI-certified desserts. This deployment bridges the gap between offline model training and end-user accessibility. The platform functions as an automated authentication system that can be accessed through a standard web browser without requiring specialized hardware or software, thereby supporting GI dessert verification and digital heritage preservation.

6.1 System Architecture and Model Integration

The application follows a modern decoupled three-tier microservice architecture to ensure scalability, maintainability, and separation of concerns. As illustrated in Figure 6.1, the framework is divided into three independent layers: the Client Layer, the Server Layer, and the ML Inference Layer. Each layer operates as a separate service and communicates through clearly defined interfaces.

The frontend was developed as a Single-Page Application (SPA) using React.js and built with Vite for optimized asset delivery. Vite supports hot module replacement during development and tree-shaking for production builds, resulting in a lightweight and efficient deployment bundle. The React client follows a component-based design where reusable user interface components manage application state, routing, and user interactions.

The backend architecture is separated into two specialized services: an API Gateway developed using Node.js and Express.js, and a dedicated machine learning inference server implemented in Python using Flask. This separation allows application-level concerns such as authentication, request validation, and routing to remain isolated from computationally intensive inference tasks. As a result, each layer can be scaled and maintained independently.

The complete request lifecycle, illustrated in Figure 6.1 of the Appendix, proceeds as follows:

1. The user uploads an image of a dessert through the React client interface, which sends a multipart HTTP POST request to the Node.js API endpoint.
2. The Node.js API Gateway acts as a secure middleware layer that decodes the binary image buffer, validates the request, and forwards the payload to the Python inference engine through an internal HTTP connection.

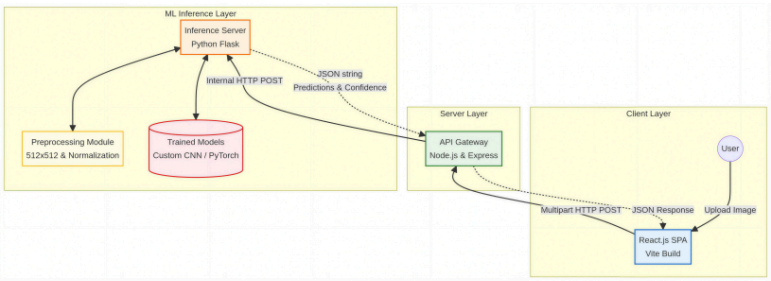


Figure 6.1: Three-tier microservice architecture of the web-based GI dessert recognition system, illustrating the Client Layer (React.js SPA), API Gateway Layer (Node.js/Express), and ML Inference Layer (Python/Flask).

3. The Python Flask inference server receives the image, loads the appropriate trained model, and executes the inference pipeline. Keras is used for the Custom CNN model, while PyTorch is used for transfer learning models such as EfficientNet-B3 and ShuffleNetV2.
4. The inference server returns a structured JSON response containing the predicted class label, district metadata, and confidence score, which is then displayed in real time by the React frontend.

This architectural isolation ensures strong modularity. The computational inference engine can be upgraded, replaced, or scaled independently without affecting the frontend client or API gateway. For example, multiple inference server instances may be deployed behind a load balancer to support higher request volumes without modifying the user interface.

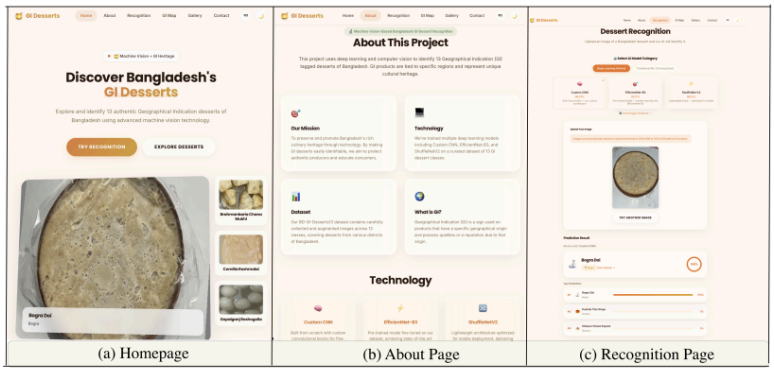
During inference, each uploaded image undergoes the same preprocessing pipeline used during training to maintain consistency. The image is resized to 512×512 pixels and normalized channel-wise so that the input distribution matches the statistical properties of the training dataset. The processed image is then passed through the Custom CNN model, followed by a softmax activation function at the output layer to generate a normalized probability distribution across the 13 GI dessert classes:

where $K = 13$ represents the total number of GI dessert categories, z_i denotes the raw logit score for class i, and $\hat{y}_i$ is the predicted probability of class i. The class with the highest probability is selected as the final prediction, and the corresponding confidence score is returned to the user interface. This provides both the classification result and a numerical measure of model certainty.

6.2 User Interface and Functional Modules

The web interface is divided into a series of functional modules which aim at facilitating automated GI dessert recognition and at the same time lead to awareness of the GI-certified

culinary heritage of Bangladesh. The platform will have six large pages, as seen in Figure 6.2 and Figure 6.3. The interface has a light mode default theme and a clean and warm visual language. The development was based on access and inclusivity, made accessible via internationalization (i18n) where users can toggle between English and Bengali by clicking on the toggle bar, which is labeled as "Bn". The application also comes entirely with responsive layouts on both desktops and mobile devices, as well as Light/Dark mode switching using the crescent icon in the header.



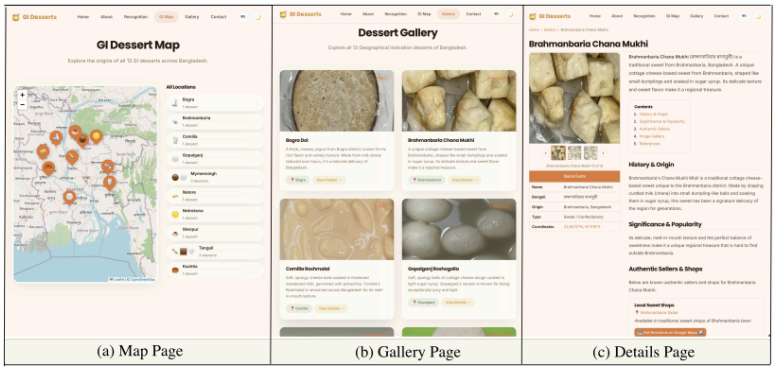
(a) Homepage (b) About Page (c) Recognition Page

Figure 6.2: Web application interface (Tablet View): (a) Homepage with an overview of the GI dessert recognition system and featured dessert thumbnails; (b) About Page presenting the project mission, BD-GI-Desserts13 dataset, and deployed deep learning models; (c) Recognition Page showing inference of Bogura Doi predicted with 100% confidence using the Custom CNN model.

6.2.1 Homepage

The Homepage Figure 6.2 a serves as the main entry point of the platform. It begins with a prominent hero section containing the headline *Discover Bangladesh GI Desserts*, followed by the tagline: *Explore and identify 13 authentic Geographical Indication desserts of Bangladesh using advanced machine vision technology.* Two large call-to-action buttons are positioned beneath the tagline: *Try Recognition* (primary orange button) and *Explore Desserts* (secondary outlined button), allowing users to navigate either to the AI recognition module or to the dessert gallery.

Below the hero section, a featured dessert panel displays visually rich card-based previews of selected desserts. The main highlighted card features Bogura Doi, while additional thumbnail cards present Brahmanbaria Chana Mukhi, Cumilla Roshmalai, and Gopalganj Roshogolla. Each card includes both the dessert name and district label, providing direct visual cues for class identity and regional origin.



(a) Map Page (b) Gallery Page (c) Details Page

Figure 6.3: Web application interface (Tablet View): (a) GI Dessert Map Page showing an interactive Leaflet-based geographical visualization of all 13 GI dessert origins in Bangladesh; (b) Dessert Gallery Page displaying all GI dessert categories; (c) Dessert Details Page presenting encyclopedic cultural, historical, and geographical information.

6.2.2 About Page

The About Page Figure 6.2 b presents the purpose and technical background of the project. The page begins with a project badge titled *Machine Vision-Based Bangladeshi GI Dessert Recognition*, followed by the heading *About This Project* and a brief summary of the platform.

The main content is organized into four informational quadrants. The first section, Our Mission, focuses on preserving and promoting the culinary heritage of Bangladesh through digital technology. The second section, Technology, introduces the three developed deep learning models: Custom CNN, EfficientNet-B3, and ShuffleNetV2. The third section, Dataset, highlights the BD-GI-Desserts13 dataset and its 13 dessert classes with augmented image samples. The fourth section, Impact, emphasizes the role of artificial intelligence in cultural heritage documentation, recognition, and digital preservation. The section below the quadrant section has a specific technology section, where the three models deployed are shown in a card grid format. The card with the highest orange accent is Custom CNN (99.37%), followed by EfficientNet-B3 with blue (93.17%), and ShuffleNetV2 with a teal accent (98.12%). Each card has a brief architectural overview and a link to an external reference.

6.2.3 Recognition Page

The Recognition Page Figure 6.2 c is the main automated authentication component of the platform. The heading of the page is Dessert Recognition, subheading: Upload an image of a Bangladeshi dessert and our AI will identify it. There is a model selection panel on the top of the page, where the users can select either Deep Learning (active mode) or Traditional ML (coming soon). Under the Deep Learning mode, users can choose one of three possible models:

Custom CNN (99.37%), EfficientNet-B3 (98.17%), or ShuffleNetV2 (98.52%).

There is a drag-and-drop image upload section below the model selection. A system note reminds the users that posted pictures are automatically downsampled to the ideal input resolution prior to inference. Predictions are then shown in the results panel in real time. In the example provided, a Custom CNN model identifies an uploaded image of Bogura Doi with 100% accuracy. The outcome is also confirmed by a Top-3 chart of prediction: Bogura Doi 100%, Kushtia Tiler Khaza 0% and Sherpur Chana Payesh 0%. This proves the high discriminative power of the trained model.

6.2.4 GI Dessert Map Page

The GI Dessert Map Page Figure 6.3a in the Appendix) is an interactive geospatial visualization layer created with Leaflet.js and OpenStreetMap tiles. The map shows the geographical region of Bangladesh, with custom dessert-icon markers showing the home districts of each GI-certified dessert category. An information panel on the right side displays all the represented districts and the number of desserts they have: Bogura (1), Brahmanbaria (1), Cumilla (1), Gopalganj (1), Mymensingh (2), Natore (1), Netrokona (1), Sherpur (1), Tangail (3), and Kushtia (1). The dataset assigns the greatest GI density of desserts to Tangail with three unique desserts. An interaction with any map marker shows the localized category of dessert, and the page is an excellent digital geographical database of the intangible Bangladesh cuisine.

6.2.5 Dessert Gallery Page

The Dessert Gallery Page Figure 6.3b is a fully navigated and richly visualized page of all 13 GI-certified dessert categories. The page is structured into a two column grid of cards that are responsive and show high resolution images of the dessert alongside the name, brief cultural and culinary explanation, a district-of-origin badge, and a View Details navigation button. Such examples can be found in the interface, such as Bogura Doi, a thick and creamy yogurt dessert of Bogura district with a strong flavor and smooth texture; Brahmanbaria Chana Mukhi, a cottage cheese-based sweet, popular in Brahmanbaria; and Cumilla Roshmalai, dumplings of soft cottage cheese, served in thickened sweetened milk and commonly topped with nuts. The card-like uniform design allows the users to compare visually the intra-class and inter-class morphological diversity of desserts in various districts of Bangladesh. Due to this, the gallery is not only a browsing interface, but also an educational visual repository of the local culinary identity.

6.2.6 Dessert Details Page

The Dessert Details Page Figure 6.3c, illustrated using *Brahmanbaria Chana Mukhi* as an example, provides encyclopedic and structured information for each GI dessert category. The page is accessible through breadcrumb navigation (`Home > Gallery > Brahmanbaria Chana Mukhi`) and is divided into clearly defined informational blocks covering History and Origin, Significance and Popularity, Authentic Sellers and Shops, Image Gallery, and References. A

dedicated *Quick Facts* panel on the right side summarizes key metadata, location (Brahmanbaria, Bangladesh), type (Sweet / Confectionery), and exact geographic coordinates. An integrated multi-image carousel allows users to browse several authentic photographs of the selected dessert. The *Authentic Sellers & Shops* section documents traditional local vendors, particularly established sweet shops in Brahmanbaria Sadar, and a *Get Directions on Google Maps* button is provided to encourage physical cultural tourism and facilitate sourcing of authentic GI-certified products. This page directly supports one of the primary objectives of geographical indication certification: consumer awareness, origin authenticity, and protection of traditional food heritage.

6.3 Real-Time Inference Capability

The proposed Custom CNN is exceptionally lightweight compared with deep and computationally intensive architectures, making it highly suitable for low-latency, real-time server-side applications. As shown in the Recognition Page Figure 6.2.c, the system successfully recognizes real-world examples–such as Bogura Doi with 100% confidence–under unconstrained conditions without reliance on curated image uploads. This demonstrates the strong discriminative power and generalization capability of the trained model on uncontrolled real-world data. When directly compared with computationally expensive pretrained models such as EfficientNet-B3, the Custom CNN achieves an almost identical test accuracy of 99.37%, while requiring significantly fewer parameters and substantially lower inference cost. This optimization enables rapid JSON response delivery through the REST API and ensures a responsive user experience even when deployed on standard commodity server hardware. All inference requests are processed through a stateless HTTP workflow consisting of file upload, asynchronous preprocessing, neural network forward pass, and standardized JSON response generation, with no session state retained between requests. This stateless design is also architecturally advantageous, as it supports natural horizontal scalability. Additional inference server instances can be deployed behind a load balancer, allowing multiple requests to be processed concurrently without structural modification to the system. Furthermore, the three-tier microservice architecture remains modular, enabling the inference engine, API gateway, and frontend client to be independently deployed, monitored, and scaled according to operational demand. The computational efficiency of the proposed system ultimately validates the central thesis of this research: a purpose-built, domain-specific lightweight CNN can be effectively transformed from a controlled experimental model into a robust, scalable, and publicly accessible authentication engine for recognizing Bangladesh's traditional GI-certified desserts in digital form. Therefore, the developed web application is not merely a proof-of-concept deployment, but a functional cultural documentation platform that applies machine vision technology in service of national heritage preservation.

7 Benchmarking and Comparative Analysis of GI Dessert Recognition Models

The experimental findings of the Custom CNN should be placed in the wider context of food and dessert recognition, which is carried out by machine vision to understand their importance. The chapter puts this proposed architecture in context compared to previously published literature in terms of classification accuracy, scale of dataset, and deployment suitability, and explores the performance benefits that the Custom CNN specifically offers over the internal baselines as well as the published systems.

7.1 Comparative Evaluation Against Published Studies

The development of food image recognition has been advanced with numerous works being done on general food classification, dessert recognition and culturally relevant tasks. To benchmark the contribution of This Work, a comparative analysis was done in detail against 16 related published works. The comparison is based on the research scope, model architecture, the size of the dataset, the number of classes, the reported accuracy/performance, and significant contributions. These studies are tabulated in Table 7.1 in the order of accuracy/performance.

The literature reviewed shows a continuous improvement in the classification of dessert and food images. Previous research like Ciocca et al. [16] obtained an accuracy of 79.00% on a small dataset with hybrid CNN + SVM. Subsequent literature added in more complex networks and attention models. The results of Xu et al. [10] showed that 87.33% Top-1 accuracy is achieved by using a combination of different CNN backbones with CBAM and Mixup augmentation. Nijhawan et al. [40] achieved 94.63% with a hybrid Vision Transformer hand-crafted features. InceptionV3 was reported to achieve 95.89% accuracy on a large binary dataset, VGG-16 was reported to achieve 96.75% accuracy, and MobileNet was reported to achieve up to 98% accuracy on Bangladeshi desserts in studies that were specifically focused on desserts. In This Work, the dataset BD-GI-Desserts13 was presented, comprised of 11,692 high-resolution RGB images, in 13 officially recognized Bangladeshi GI-certified dessert categories. The proposed Custom CNN was fairly compared to two powerful transfer learning baselines namely EfficientNet-B3 and ShuffleNetV2 by being trained in the same preprocessing and stratified data splits with the same training conditions. The highest accuracy was 93.17% with EfficientNet-B3 that uses compound scaling to balance depth, width, and resolution, however, with more confusion between visually similar subclasses, in particular between brown and white variants of monda and sondesh. The computationally efficient (via channel split and shuffle) ShuffleNetV2 achieved a 98.12% accuracy and showed good generalization with fewer resources, but with some perfor-

Table 7.1: Comparative Analysis of Existing Food Image Recognition Approaches

Study	**Scope**	**Model / Architecture**	**Images**	**Classes**	**Accuracy / Performance**	**Key Notes / Contributions**
Mukhiddinov et al. [31]	Fruits & Vegetables Freshness	Improved YOLOv4 (Mish activation)	Custom	Fresh vs Rotten	50.4% AP	Freshness detection (higher than YOLOv4 & YOLOv3)
Ciocca et al. [16]	Food	CNN + SVM	1,027	73	79.00%	Hybrid CNN-SVM approach
Xu et al. [10]	Asian Food	MobileNetV2 + VGG16 + ResNet50 + CBAM + Mixup	Asian food dataset	-	87.33%	Attention mechanism + mixed data enhancement
Nijhawan et al. [40]	Food Cuisines	Hybrid Vision Transformer + Hand-crafted features	Custom Indian food	13	94.63%	Hybrid VTNet approach
Chen et al. [36]	Food Image Recognition	MAF-Net (Multi-level Attention Feature Fusion)	CETH Food-101, Vireo Food-172, UEC Food-100	-	Top-1: 90.22%, 89.86%, 90.61%	Multi-level attention fusion + self-attention
Polin et al. [46]	Tomato Pest Recognition (Bangladesh)	CNN features + RF/SVM/KNN	Custom	5 pest types	95.49%	Bangladesh-specific tomato pest recognition
Exarchou et al. [14]	Dessert	InceptionV3	38,000	2	95.89%	Dessert vs non-dessert for dietary assessment
Hossain et al. [15]	Dessert	VGG-16	2,633	15	96.75%	Dessert classification
Ramesh et al. [6]	Food	SSD + InceptionV2	4,200	60	97.60%	Object detection + classification
Rahman & Hasan [2]	Bangladeshi Dessert	MobileNet	4,118	27	98.00%	Bangladeshi dessert classification
This Work	**GI Dessert**	**Custom CNN**	**11,692**	**13**	**99.37%**	**Proposed model for Bangladeshi GI Dessert recognition**

mance decline on texture-similar dessert pairs. To identify the subtle chromatic, textural, and structural differences in GI desserts, the proposed Custom CNN, consisting of five convolutional blocks with filter progressively growing to 32, 512, etc. was designed accordingly. This Work had a test accuracy of 99.37 on the BD-GI-Desserts13 dataset. Direct comparison using the same dataset indicates that the Custom CNN was more effective than both EfficientNet-B3 (93.17%) and ShuffleNetV2 (98.12) in overall accuracy and lowering the computational complexity. The confusion matrix analysis showed a high degree of diagonal dominance with low inter-class confusion, and Grad-CAM visualizations demonstrated that the model always paid attention to discriminative features like surface texture and structural patterns rather than the background noise. As well, deep features derived out of the Custom CNN backbone were used to apply traditional machine learning classifiers (Random Forest, Support Vector Machine, and K-Nearest Neighbors). These classic ML models offered decent performance and valuable information on feature discriminability. Yet, in general, they could be less accurate than the end-to-end deep learning models (Custom CNN 99.37%, ShuffleNetV2 at 98.12 and EfficientNet-B3 at 93.17), particularly in the task of comparing the fine-grained visual similarities between the 13 GI dessert classes. This also confirmed the power of the end-to-end Custom CNN solution to the GI dessert recognition problem. This Work shows very competitive performance when compared to the rest of the literature on the topic in Table 7.1 The obtained accuracy of 99.37 is near or slightly higher than the highest reported in dessert-specific classification (98% by Rahman and Hasan [1], [2]).

The emphasis on origin-conscious, fine-grained GI dessert recognition with a specific dataset fulfills a gap in the literature which mainly focuses on general or binary classification of desserts. This relative analysis points to the conclusion that a task-specific Custom CNN with appropriate training on a challenging, culturally-relevant dataset can match the performance of powerful transfer learning baselines like EfficientNet-B3 and ShuffleNetV2, and has the added benefits of computational efficiency and interpretability.

7.2 Performance Analysis and Superiority of the Proposed Architecture

The proposed Custom CNN overcomes all the limitations mentioned above through a combination of domain-specific architecture design, purpose-built dataset construction, and end-to-end optimization based on gradients. Using BD-GI-Desserts13, the largest and most geographically representative GI dessert dataset published to date, with 11,692 high-resolution RGB images captured in various certified districts of Bangladesh, the Custom CNN achieves a 99.37% test accuracy, which is 1.37% points higher than the nearest published result.

The difference of 6.20% points between EfficientNet-B3 (93.17%) and the Custom CNN (99.37%), given the same preprocessing and training conditions, indicates that the richness of ImageNet pretraining fails to compensate for domain-specific architecture design when the classification task requires fine-grained subcategory differentiation in a culturally specialized visual domain. The operationally unacceptable Recall deficits of EfficientNet-B3 on Netrokona

Balish Mishti (0.8444) and Muktagacha Monda White (0.8222) represent false negatives that are penalized in a regulatory GI certification setting. ShuffleNetV2 performed slightly lower with 98.12%, only 1.25% points below the Custom CNN, indicating that lightweight architectures can still be used for GI authentication deployment, but the Custom CNN remains superior, especially for visually ambiguous subclasses.

The fact that the Custom CNN outperforms traditional ML classifiers also confirms the proposed approach. Although these classifiers were trained on discriminative CNN embeddings rather than raw pixels, k-NN achieved only 91.18% accuracy and the Decision Tree 86.74% — 8.19 and 12.63% points lower than the Custom CNN, respectively.This difference in performance between end-to-end CNN and conventional ML classifiers is in agreement with the literature on food authentication across the world. Independently found that CNN-based models outperformed SVM and k-NN in multiple food matrices on the task of adulteration and origin detection definitively and confirmed that hierarchical feature learning has a structural advantage over margin-based and distance-based decision strategy [53]. Feng et al. [49] also determined the deep CNN designs, which automatically extract spatial and spectral features, as being more accurate on the origin discriminative than hand-crafted feature classifiers on food variety identification tasks. The consistency of these results with other independent studies around the world supports the validity and generalizability of the Custom CNN as the best model to use in fine-grained GI dessert recognition. These gaps substantiate that end-to-end gradient-based optimization has an inherent advantage over distance-based and margin-based decision strategies for non-linear textural classification tasks.

In addition to accuracy, the Custom CNN offers useful deployment advantages not provided by existing systems. Its lightweight architecture supports real-time web-based inference with a significantly smaller computational footprint than EfficientNet-B3. Grad-CAM visualizations reveal that classification decisions rely on discriminative dessert-specific surface characteristics such as caramelization patterns, syrup pooling boundaries, and chhena granularity, rather than background artifacts. This provides the interpretability required for regulatory audit and consumer trust. Together, the 99.37% accuracy, balanced Precision and Recall above 0.93 across all 13 classes, low computational cost, and visual interpretability make the Custom CNN the most consistent, deployable, and domain-aligned solution for Bangladeshi GI dessert authentication to date.

8 Conclusion and Future Work

8.1 Conclusion

This thesis proposed a machine vision architecture to identify and verify Geographical Indication (GI) certified Bangladeshi desserts to fill a critical gap at the convergence of computer vision, food provenance verification, and cultural heritage preservation. The work was inspired by the increasing susceptibility of GI-certified regional desserts to counterfeiting and mislabeling in commercial markets, where a lack of an automated recognition mechanism has long compromised consumer protection as well as the economic well-being of the legitimate regional manufacturers. To overcome this, a curated dataset, 11,692 high-resolution RGB images, BD-GI-Desserts13, was assembled and included 13 officially GI-certified dessert classes, and acquired by field-based sampling at certified vendors in ten geographically diverse districts of Bangladesh. The recognition task was formalized as an image classification task of a fine-grained nature, which is very challenging as the visual similarity between the classes is high, especially between color variations, e.g. the brown and white subclasses of Muktagacha Monda and Tangail Sondesh, as well as natural intra-class variation due to variation in illumination, camera angle, and batch-to-batch appearance of products. They constructed a lightweight Custom Convolutional Neural Network (CCNN) and trained and tested it against transfer learning baselines (EfficientNet-B3 and ShuffleNetV2) and three conventional machine learning classifiers (SVM, k-NN, and Decision Tree). The recommended CCNN had a test accuracy of 99.37%, highest compared to others tested. Grad-CAM visualizations revealed that the model repeatedly focused on discriminative surface textures and structural features but not background noise. An experimental web-based interface was then created to show the real-time recognition ability of the system in a real-life scenario.

8.2 Contributions and Impact

The paper introduces a number of original and substantial contributions to the research areas of fine-grained visual recognition, food provenance verification, and digital preservation of the cultural heritage via computer vision. The first contribution is the creation of the first curated benchmark dataset called BD-GI-Desserts13, which is specifically created to be used in GI-certified Bangladeshi dessert recognition. The data is 11,692 high-resolution RGB images of 13 officially recognized dessert categories, all gathered at the field at certified sellers within ten geographically diverse districts of Bangladesh. In contrast to the previous Bangladeshi

food datasets, in which the GI provenance information is either not provided or similar looking products are grouped together without considering geographical location, BD-GI-Desserts13 captures the GI certification status and the district of production of each of the classes. The visually similar subclass pairs, the brown and white of Muktagacha Monda and Tangail Sondesh, introduce a challenging fine-grained discrimination task, and the 70/15/15 experimental conditions are fully reproducible by the research community. The second contribution is the design and validation of a computationally efficient Custom CNN architecture that is designed specifically to this domain. The proposed CCNN demonstrated a test accuracy of 99.37 %, which is higher than the test accuracy of the two transfer learning baselines, ShuffleNetV2 (98.12 %) and EfficientNet-B3 (93.17 %) with the same preprocessing and training conditions. The finding shows that smaller, domain-relevant data-trained lightweight, task-specific networks can efficiently perform fine-grained regional food classification, at much lower computational complexity, compared to larger pretrained networks. The practical feasibility of the model in resource constrained systems, allows it to be integrated in retail authentication systems, mobile apps and point of sale systems where large pretrained networks would not be feasible. The third contribution is a stringent and open comparative evaluation framework that cuts across deep learning and conventional machine learning paradigms. This research offers a reproducible baseline by comparing the CCNN with ShuffleNetV2, EfficientNet-B3, SVM, k-NN, and Decision Tree on the same experimental conditions, which can be directly extended by the future researchers. The analysis of confusion matrices, Grad-CAM visualizations, and learning curve diagnostics all testify to a high discriminative power of the model, the accuracy of its feature localization, and the stability of its generalization, which confirms its applicability to real-world use as a GI authentication model.

8.3 Limitations

Even though the suggested framework has a high classification performance, there are some limitations that should be considered in order to estimate the current scope and generalizability of the framework. To begin with, although BD-GI-Desserts13 has 11,692 pictures in 13 categories, all of the pictures were obtained in a confirmed yet limited sample of certified vendors in the chosen districts. This constrained sourcing might not be fully representative of the extent of stylistic, seasonal and preparation-based visual variation among the various producers in the same GI region, which may restrain model robustness when applied to more heterogeneous real-world conditions. Second, the dataset offers image-level classification labels and district and capture environment data only. The lack of bounding box labels, pixel-based segmentation mask, or ingredient-level labels restrain the dataset usage to more focused downstream tasks like multi-object detection or portion estimation. Third, the existing model only allows closed-set classification of known GI categories. It is not able to identify non-GI imitations or to reject out-of-distribution inputs, a requirement of a complete functioning product authenticity verification system. Lastly, the photos were all taken on iPhone devices, which could bring a level of the device-specific bias and limit cross-device generalization to photos that are obtained

on different hardware settings or with controlled lighting environments when the model is used to test photos. One of the fundamental shortcomings of this visual classification system is that it identifies the appearance attributes historically and culturally attributed to each GI dessert category -it does not confirm the geographic origin of a particular sample. An objectively looking fake product can be correctly classified to belong to the right group even though it is not located within the specified GI area. This is a natural limitation of the image-based classification and is in line with the visual-recognition scope of this study. To overcome this limitation in future work, a multi-modal solution involving chemical fingerprinting, isotope ratio analysis, or blockchain-based provenance tracking and visual classification would be necessary.

8.4 Future Work

The future studies should have some directions that need to be followed. The most urgent one is the presence of BD-GI-Desserts13 in order to serve newly GI-certified dessert products as the Bangladesh registry is expanded, and the widening of the range of sources due to the increased number of locales of the vendors, time of the year, and image shooting in a broader range of equipment and under a broader range of lighting. Developing counterfeit detection based on anomaly-based learning is one of such important directions in the future. The system is presently in the shape of the recognition of known GI images but does not recognize non-GI imitations directly. Research in the future should consider the one-class learning, open-set recognition and anomaly detection systems that can distinguish between actual GI desserts and visual analogues, which can directly influence the regulatory and consumer protection objectives of GI certification. The other important direction is deployment on resource constrained and edge computing platforms. Compression of models, through the use of knowledge distillation, quantization and pruning would enable their execution on mobile devices and point-of-sale terminal devices, and would also make the framework applicable to shopkeepers, customs officers, and end consumers. Finally, the more advanced explainability methods than Grad-CAM, and the analysis of CNN-Transformer hybrids, would further enhance the discriminative capacity and institutional explainability of the recognition system, which would help to apply it to a more regulatory and compliance-heavy environment.

Publication Records

An original version of this work was prepared and submitted to the international Conference on Artificial Intelligence and Smart Data for Sustainable Engineering and Innovation (AISEI 2026), held on 27–28 April 2026 at Jadara University, Jordan. After a thorough peer-review process and subsequent revisions on the basis of what the reviewer commentated on the paper, it was accepted to be presented and published in the conference. This acceptance substantiates the scientific value of the proposed machine vision system to fine-grained detection of GI-certified Bangladeshi desserts and justifies its applicability to the larger context of artificial intelligence and smart data technologies to sustainable engineering and innovation.

A similar publication was also accepted at the 24th IEEE/ACIS International Conference on Software Engineering Research, Management and Applications (SERA 2026) to be held on 27-29 May 2026 at Towson University, Towson, Maryland, USA. The findings given in this thesis are a more detailed and extended version of the accepted conference papers, including the additional experimental findings, further testing of the models, and a full web-based recognition interface.

Bibliography

[1] I. Firozy, S. Supriya, M. S. Rahman, S. Nawshin, and M. T. Habib, "Cnn modeling for bengali sweet recognition," *SSRN Electronic Journal*, 2023. 4376884.

[2] M. Rahman and J. Hasan, "A comprehensive dataset for bangladeshi dessert classification," *Data in Brief*, vol. 57, p. 111077, 2024.

[3] B. Ahmed, M. Haque, A. Huq, M. Masud, M. Ali, and M. Naznin, "Foodbd: a polygon-annotated meal image dataset of bangladeshi cuisines with visual and nutritional labels," *BMC Research Notes*, vol. 19, 2025.

[4] A. Dhar, M. Hossain, and P. Barua, "Bangladeshi street food calorie estimation using improved yolov8 and regression model," *arXiv preprint*, 2025.

[5] S. Shekhar, P. S. Minz, and K. Prasad, "Image analysis as non-destructive approach in characterization of indian sweet meat spongy rosogulla," *Journal of Postharvest Technology*, vol. 8, no. 3, pp. 50–60, 2020.

[6] A. Ramesh, A. Sivakumar, and S. S. Angel, "Real-time food-object detection and localization for indian cuisines using deep neural networks," in *IEEE ICMLANT*, pp. 1–6, 2020.

[7] S. Goyal and G. K. Goyal, "Machine learning time-delay model for analyzing shelf life of burfi," *International Journal of Scientific & Engineering Research*, 2012.

[8] N. Theera-Ampornpunt, P. Treepong, N. Phokabut, and S. Phlaichana, "Kedinet: A hybrid deep learning architecture for thai dessert recognition," *IEEE Access*, 2025.

[9] D. Trisnawarman, A. Supriyanton, V. Mawardi, and U. Okengwu, "Comparative study of cnn and vision transformers on indonesian traditional cakes classification," *International Journal of Advances in Artificial Intelligence and Machine Learning*, 2025.

[10] B. Xu, X. He, and Z. Qu, "Asian food image classification based on deep learning," *Journal of Computer and Communications*, 2021.

[11] S. Feng, Y. Wang, J. Gong, X. Li, and S. Li, "A fine-grained recognition technique for identifying chinese food images," *Heliyon*, vol. 9, 2023.

[12] A. Karabay, H. Varol, and M. Chan, "Improved food image recognition by leveraging deep learning and data-driven methods with an application to central asian food scene," *Scientific Reports*, vol. 15, 2025.

[13] R. Nijhawan, G. Sinha, A. Batra, M. Kumar, and H. Sharma, "Vtnet+handcrafted based approach for food cuisines classification," *Multimedia Tools and Applications*, vol. 83, pp. 10695–10715, 2023.

[14] D.-M. Exarchou, A. Alexiadis, A. Triantafyllidis, D. Ioannidis, K. Votis, and D. Tzovaras, "A deep learning-based dessert recognition system for automated dietary assessment," in *International Conference on Wireless Mobile Communication and Healthcare*, pp. 60–70, Springer, 2021.

[15] M. S. Hossain, M. Al-Hammadi, and G. Muhammad, "Automatic fruit classification using deep learning for industrial applications," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 1027–1034, 2018.

[16] G. Ciocca, P. Napoletano, and R. Schettini, "Food recognition: A new dataset, experiments, and results," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 588–598, 2016.

[17] G. Ciocca, P. Napoletano, and R. Schettini, "Cnn-based features for retrieval and classification of food images," *Computer Vision and Image Understanding*, vol. 176–177, pp. 70–77, 2018.

[18] G. Ciocca, G. Micali, and P. Napoletano, "State recognition of food images using deep features," *IEEE Access*, vol. 8, pp. 32003–32017, 2020.

[19] M. M. Anthimopoulos, L. Gianola, L. Scarnato, P. Diem, and S. G. Mougiakakou, "A food recognition system for diabetic patients based on an optimized bag-of-features model," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, pp. 1261–1271, July 2014.

[20] L. Boyd, N. Nnamoko, and R. Lopes, "Fine-grained food image recognition: A study on optimising convolutional neural networks for improved performance," *Journal of Imaging*, vol. 10, 2024.

[21] A. Chaitanya, J. Shetty, and P. Chiplunkar, "Food image classification and data extraction using convolutional neural network and web crawlers," *Procedia Computer Science*, 2023.

[22] C.-S. Chen, G.-Y. Chen, D. Zhou, D. Jiang, and D.-S. Chen, "Res-vmamba: Fine-grained food category visual classification using selective state space models with deep residual learning," *arXiv preprint arXiv:2402.15761*, 2024.

[23] Z. Chen, J. Wang, and Y. Wang, "Enhancing food image recognition by multi-level fusion and the attention mechanism," *Foods*, vol. 14, 2025.

[24] V. G, P. Vutkur, and V. P, "Food classification using transfer learning technique," *Global Transitions Proceedings*, 2022.

[25] T. Ghosh and E. Sazonov, "Improving food image recognition with noisy vision transformer," in *IEEE EMBC*, 2025.

[26] B. N. Jagadesh, S. V. Mantena, A. P. Sathe, T. Prabhakara Rao, K. K. Lella, S. S. Pabboju, and R. Vatambeti, "Enhancing food recognition accuracy using hybrid transformer models and image preprocessing techniques," *Scientific Reports*, vol. 15, p. 5591, Feb. 2025.

[27] P. McAllister, H. Zheng, R. Bond, and A. Moorhead, "Combining deep residual neural network features with supervised machine learning algorithms to classify diverse food image datasets," *Computers in Biology and Medicine*, vol. 95, pp. 217–233, 2018.

[28] W. Min, Z. Wang, Y. Liu, M. Luo, L. Kang, X. Wei, X. Wei, and S. Jiang, "Large scale visual food recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, p. 9932–9949, Aug. 2023.

[29] K. A. Nfor, T. P. Theodore Armand, K. P. Ismaylovna, M.-I. Joo, and H.-C. Kim, "An explainable cnn and vision transformer-based approach for real-time food recognition," *Nutrients*, vol. 17, no. 2, 2025.

[30] S. Rokhva, B. Teimourpour, and A. H. Soltani, "Computer vision in the food industry: Accurate, real-time, and automatic food recognition with pretrained mobilenetv2," *Food and Humanity*, vol. 3, p. 100378, 2024.

[31] M. Mukhiddinov, A. Muminov, and J. Cho, "Improved classification approach for fruits and vegetables freshness based on deep learning," *Sensors*, vol. 22, no. 21, 2022.

[32] M. Mansouri, S. B. Chaouni, S. J. Andaloussi, and O. Ouchetto, "Deep learning for food image recognition and nutrition analysis towards chronic diseases monitoring: A systematic review," *SN Computer Science*, vol. 4, no. 5, p. 513, 2023.

[33] L. M. Amugongo, A. Kriebitz, A. Boch, and C. Lütge, "Mobile computer vision-based applications for food recognition and volume and calorific estimation: A systematic review," *Healthcare*, vol. 11, no. 1, 2023.

[34] K. V. Dalakleidi, M. Papadelli, I. Kapolos, and K. Papadimitriou, "Applying image-based food-recognition systems on dietary assessment: A systematic review," *Advances in Nutrition*, vol. 13, no. 6, pp. 2590–2619, 2022.

[35] F. S. Konstantakopoulos, E. I. Georga, and D. I. Fotiadis, "A review of image-based food recognition and volume estimation artificial intelligence systems," *IEEE Reviews in Biomedical Engineering*, vol. 17, pp. 136–152, 2024.

[36] F. P. W. Lo, Y. Sun, J. Qiu, and B. Lo, "Image-based food classification and volume estimation for dietary assessment: A review," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 7, pp. 1926–1939, 2020.

[37] W. Min, S. Jiang, L. Liu, Y. Rui, and R. Jain, "A survey on food computing," *Acm Computing Surveys (CSUR)*, vol. 52, no. 5, pp. 1–36, 2019.

[38] H. O. Velesaca, P. L. Suárez, R. Mira, and A. D. Sappa, "Computer vision based food grain classification: A comprehensive survey," *Computers and Electronics in Agriculture*, vol. 187, p. 106287, 2021.

[39] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: analysis, applications, and prospects," *IEEE transactions on neural networks and learning systems*, vol. 33, no. 12, pp. 6999–7019, 2021.

[40] A. Babenko and V. Lempitsky, "Aggregating deep convolutional features for image retrieval," *arXiv preprint arXiv:1510.07493*, 2015.

[41] W. Zhang, B. Zhao, S. Gao, Y. Ji, L. Zhou, K. Niu, Z. Qiu, and X. Jin, "Online recognition of small vegetable seed sowing based on machine vision," *IEEE Access*, vol. 11, pp. 134331–134339, 2023.

[42] M. T. Habib, A. Majumder, A. Jakaria, M. Akter, M. S. Uddin, and F. Ahmed, "Machine vision based papaya disease recognition," *Journal of King Saud University-Computer and Information Sciences*, vol. 32, no. 3, pp. 300–309, 2020.

[43] I. Sharmin, N. F. Islam, I. Jahan, T. Ahmed Joye, M. R. Rahman, and M. T. Habib, "Machine vision based local fish recognition," *SN Applied Sciences*, vol. 1, no. 12, p. 1529, 2019.

[44] M. A. A. Siddique, J. Ferdouse, M. T. Habib, M. J. Mia, and M. S. Uddin, "Convolutional neural network modeling for eye disease recognition.," *International Journal of Online & Biomedical Engineering*, vol. 18, no. 9, 2022.

[45] M. Nuruzzaman, M. S. Hossain, M. M. Rahman, A. S. H. C. Shoumik, M. A. A. Khan, and M. T. Habib, "Machine vision based potato species recognition," in *2021 5th international conference on intelligent computing and control systems (ICICCS)*, pp. 1–8, IEEE, 2021.

[46] M. A. Rahman, A. A. Khan, M. M. Hasan, M. S. Rahman, and M. T. Habib, "Deep learning modeling for potato breed recognition," *IEEE Transactions on AgriFood Electronics*, vol. 2, no. 2, pp. 419–427, 2024.

[47] J. A. Polin, N. Hasan, M. T. Habib, A. Rahman, Z. N. Vasha, and B. Sharma, "Tomato pest recognition using convolutional neural network in bangladesh," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 1, pp. 619–627, 2024.

[48] M. M. Islam, M. Akter, M. S. Uddin, M. A. Rahman, and M. T. Habib, "Machine vision–based insect recognition in agriculture: A comprehensive review," *Engineering Reports*, vol. 8, no. 4, p. e70729, 2026.

[49] L. Feng, B. Wu, S. Zhu, Y. He, and C. Zhang, "Application of visible/infrared spectroscopy and hyperspectral imaging with machine learning techniques for identifying food varieties and geographical origins," *Frontiers in Nutrition*, vol. 8, p. 680357, 2021.

[50] M. Ariful Hassan, M. Sydul Islam, M. Mehedi Hasan, S. B. Shorif, M. Tarek Habib, and M. S. Uddin, *Medicinal Plant Recognition from Leaf Images Using Deep Learning*, pp. 137–154. Singapore: Springer Singapore, 2022.

[51] M. T. Habib, M. A. I. Arif, S. B. Shorif, M. S. Uddin, and F. Ahmed, *Machine Vision-Based Fruit and Vegetable Disease Recognition: A Review*, pp. 143–157. Singapore: Springer Singapore, 2021.

[52] Z. Deng, T. Wang, Y. Zheng, W. Zhang, and Y.-H. Yun, "Deep learning in food authenticity: Recent advances and future trends," *Trends in Food Science & Technology*, vol. 144, p. 104344, 2024.

[53] D. A. Magdas, A. R. Hategan, M. David, and C. Berghian-Grosan, "The journey of artificial intelligence in food authentication: From label attribute to fraud detection," *Foods*, vol. 14, no. 10, p. 1808, 2025.

[54] D. Kundu, S. R. Bhuiyan, and M. T. H. Siam, "BD-GI-Desserts13," 2026.

2026_Spring____Internship_Report_Template.pdf

ORIGINALITY REPORT

15%	11%	11%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	arxiv.org Internet Source	1%
2	www.caedes.net Internet Source	1%
3	ar.iub.edu.bd Internet Source	1%
4	www.mdpi.com Internet Source	<1%
5	dspace.daffodilvarsity.edu.bd:8080 Internet Source	<1%
6	journal.uad.ac.id Internet Source	<1%
7	Submitted to Universitas Bina Sarana Informatika Student Paper	<1%
8	Daniel Ponte, Eduardo Aguilar, Mireia Ribera, Petia Radeva. "Multi-task visual food recognition by integrating an ontology supported with LLM", Journal of Visual Communication and Image Representation, 2025 Publication	<1%
9	data.mendeley.com Internet Source	<1%
10	www.ijert.org Internet Source	<1%

11	export.arxiv.org Internet Source	<1%
12	faculty.daffodilvarsity.edu.bd Internet Source	<1%
13	ikee.lib.auth.gr Internet Source	<1%
14	www.irjmets.com Internet Source	<1%
15	Paolo Ferro, Harinadh Vemanaboina, Chander Prakash. "Computational Techniques and Smart Manufacturing", CRC Press, 2026 Publication	<1%
16	Rafeed Rahman, Mehfuz A Rahman, Md. Saiful Islam, Mahady Hasan. "DeepGrip: Cricket Bowling Delivery Detection with Superior CNN Architectures", 2021 6th International Conference on Inventive Computation Technologies (ICICT), 2021 Publication	<1%
17	Submitted to Kennesaw State University Student Paper	<1%
18	ousar.lib.okayama-u.ac.jp Internet Source	<1%
19	thesai.org Internet Source	<1%
20	ualberta.scholaris.ca Internet Source	<1%
21	ijeecs.iaescore.com Internet Source	<1%
22	pubs.thesciencein.org Internet Source	<1%
	www.iti.gr	

23	Internet Source	<1%
24	R. Selvaraj, M. S. Geetha Devasena. "A novel attention based vision transformer optimized with hybrid optimization algorithm for turmeric leaf disease detection", Scientific Reports, 2025 Publication	<1%
25	Submitted to Sardar Patel Institute of Technology Student Paper	<1%
26	online-journals.org Internet Source	<1%
27	repositorio.uam.es Internet Source	<1%
28	Submitted to University of Hertfordshire Student Paper	<1%
29	Submitted to University of Birmingham Student Paper	<1%
30	e-journal.stmiklombok.ac.id Internet Source	<1%
31	Panadda Kongsilp, Naratip Sangsai. "Thai Milled Rice Quality Classification Based on Deep Learning Approach", 2022 International Electrical Engineering Congress (iEECON), 2022 Publication	<1%
32	Submitted to University of Limerick Student Paper	<1%
33	Submitted to University of Sheffield Student Paper	<1%
34	insightsociety.org Internet Source	<1%

35	Pushpa Choudhary, Sambit Satpathy, Arvind Dagur, Dhirendra Kumar Shukla. "Recent Trends in Intelligent Computing and Communication", CRC Press, 2025 Publication	<1 %
36	Submitted to Swinburne University of Technology Student Paper	<1 %
37	Submitted to University of Edinburgh Student Paper	<1 %
38	www.eurchembull.com Internet Source	<1 %
39	Weilong He, Fan Wu, Lori Unruh Snyder, Jean Cheng, Evelynn Wilcox, Lirong Xiang. "Improved two-stage deep learning algorithm and lightweight YOLOv5n for classifying cottonseed damage", Computers and Electronics in Agriculture, 2025 Publication	<1 %
40	Bowne, Madeline. "Flexible Space Junk Allocation & Waste Abatement (SPACE JAWA): Flexibility Framework to Screen Strategies & Options for Sustainable On-Orbit Servicing Infrastructures in LEO.", Georgia Institute of Technology Publication	<1 %
41	d197for5662m48.cloudfront.net Internet Source	<1 %
42	Submitted to University of Leeds Student Paper	<1 %
43	ijsrem.com Internet Source	<1 %

44	B. G. Nagaraja, S. Kannadhasan. "Information and Communication Systems", CRC Press, 2026 Publication	<1 %
45	Submitted to The University of Manchester Student Paper	<1 %
46	Submitted to Vrije Universiteit Amsterdam Student Paper	<1 %
47	Xu Xiao, Yiming Jiang, Yaonan Wang. "Key Technologies for Machine Vision for Picking Robots: Review and Benchmarking", Machine Intelligence Research, 2025 Publication	<1 %
48	beei.org Internet Source	<1 %
49	"Innovations in Data Analytics", Springer Science and Business Media LLC, 2025 Publication	<1 %
50	Sourav Chakrabarty, Chandan Kumar Deb, Sudeep Marwaha, Md. Ashraful Haque et al. "Application of artificial intelligence in insect pest identification - A review", Artificial Intelligence in Agriculture, 2025 Publication	<1 %
51	acspublisher.com Internet Source	<1 %
52	Submitted to Babes-Bolyai University Student Paper	<1 %
53	Submitted to University of Wolverhampton Student Paper	<1 %
54	"Innovative AI Technologies Driving Sustainable Farming: Strategies for Improving	<1 %

Food Security", Springer Science and Business Media LLC, 2026
Publication

55	Submitted to University of Lancaster Student Paper	<1%
56	Submitted to Independent University Bangladesh Student Paper	<1%
57	researchcage.com Internet Source	<1%
58	link.springer.com Internet Source	<1%
59	Guangyi Xiao, Qi Wu, Hao Chen, Da Da, Jingzhi Guo, Zhiguo Gong. "A Deep Transfer Learning Solution for Food Material Recognition Using Electronic Scales", IEEE Transactions on Industrial Informatics, 2020 Publication	<1%
60	Submitted to Islington College,Nepal Student Paper	<1%
61	cse.buet.ac.bd Internet Source	<1%
62	Submitted to Gaborone University College of Law and Professional Studies Student Paper	<1%
63	J Martin Verkuyl, Andrew Matus. "Time-lapse imaging of dendritic spines in vitro", Nature Protocols, 2006 Publication	<1%
64	Bal S. Virdee, Tanweer Ali, Jaume Anguera, Suman Lata Tripathy. "Connecting Intelligence: Trends in Computation and Data Communication - Proceedings of First	<1%

	International Conference on Computational Intelligence and Data Communication (ICCIDC 2025)", CRC Press, 2026 Publication	
65	Submitted to GLA University Student Paper	<1 %
66	Karima Moumane, Ikram El Asri, Taoufiq Cheniguer, Soufiane Elbiki. "Food Recognition and Nutrition Estimation using MobileNetV2 CNN architecture and Transfer Learning", 2023 14th International Conference on Intelligent Systems: Theories and Applications (SITA), 2023 Publication	<1 %
67	dokumen.pub Internet Source	<1 %
68	"Computational Visual Media", Springer Science and Business Media LLC, 2025 Publication	<1 %
69	Merieme Mansouri, Samia Benabdellah Chaouni, Said Jai Andaloussi, Ouail Ouchetto. "Deep Learning for Food Image Recognition and Nutrition Analysis Towards Chronic Diseases Monitoring: A Systematic Review", SN Computer Science, 2023 Publication	<1 %
70	bibli.ec-lyon.fr Internet Source	<1 %
71	d-scholarship.pitt.edu Internet Source	<1 %
72	www.diva-portal.se Internet Source	<1 %
73	Submitted to University of Northampton Student Paper	<1 %

74	www.myhuiban.com Internet Source	<1%
75	"Computer Vision and Image Processing", Springer Science and Business Media LLC, 2024 Publication	<1%
76	10wvvw.easychair.org Internet Source	<1%
77	Submitted to Coventry University Student Paper	<1%
78	Submitted to Indian Institute Of Information Technology, Kalyani Student Paper	<1%
79	M. Dhavamani, K. Arun Prakash, P. Vadivel. "Applied Mathematics, Automation and Computing", CRC Press, 2026 Publication	<1%
80	Nyathi, Mthabisi Adriano. "Structural Health Monitoring Using Deep Learning.", University of South Wales (United Kingdom) Publication	<1%
81	Submitted to Anna University Student Paper	<1%
82	Submitted to Florida International University Student Paper	<1%
83	Submitted to University of Melbourne Student Paper	<1%
84	iksadyayinevi.com Internet Source	<1%
85	inpact-psychologyconference.org Internet Source	<1%
	Submitted to Arab Open University	

86	Student Paper	<1%
87	Md. Ashikul Aziz Siddique, Jannatul Ferdouse, Md. Tarek Habib, Md. Jueal Mia, Mohammad Shorif Uddin. "Convolutional Neural Network Modeling for Eye Disease Recognition", International Journal of Online and Biomedical Engineering (iJOE), 2022 Publication	<1%
88	Submitted to Middlesex University Student Paper	<1%
89	Shuhao Wang, Gang Xu. "Convolutional neural networks—Image classification and object detection", Elsevier BV, 2025 Publication	<1%
90	Submitted to UNICAF Student Paper	<1%
91	Submitted to University of Bath Student Paper	<1%
92	Submitted to Vellore Institute of Technology Student Paper	<1%
93	centuryscipub.com Internet Source	<1%
94	inass.org Internet Source	<1%
95	journal.abhinaya.co.id Internet Source	<1%
96	www.ijirset.com Internet Source	<1%
97	Al-Sakib Khan Pathan. "Securing Cyber-Physical Systems", CRC Press, 2019 Publication	<1%

98	ijai.iaescore.com Internet Source	<1%
99	ijarpr.com Internet Source	<1%
100	manjulaskitchen.com Internet Source	<1%
101	"Artificial Intelligence for Environmental Sustainability and Green Initiatives", Springer Science and Business Media LLC, 2024 Publication	<1%
102	Ahmed Shahzad, Bushra Zafar, Ramzan Talib, Muhammad Murad Khan. "A hybrid transfer learning approach for multi-class classification of medical deepfakes generated by GAN and diffusion models", Discover Computing, 2026 Publication	<1%
103	Jiangshun Yu, Poyu You, Jian Zhao, Xianzhe Long, Yuran Chen. "Anomaly Detection and Fault Diagnosis of Power Distribution Line Point Cloud Data Based on Deep Learning", International Journal of Advanced Computer Science and Applications, 2025 Publication	<1%
104	anjanatiha.wixsite.com Internet Source	<1%
105	archive.org Internet Source	<1%
106	drops.dagstuhl.de Internet Source	<1%
107	ebin.pub Internet Source	<1%
	pure.mpg.de	

108	Internet Source	<1%
109	www.unhabitat.org Internet Source	<1%
110	39www.easychair.org Internet Source	<1%
111	Jianing Qiu, Frank Po Wen Lo, Shuo Jiang, Charlie Tsai, Yingnan Sun, Benny Lo. "Counting Bites and Recognizing Consumed Food from Videos for Passive Dietary Monitoring", IEEE Journal of Biomedical and Health Informatics, 2020 Publication	<1%
112	Submitted to University of Westminster Student Paper	<1%
113	cuestionesdefisioterapia.com Internet Source	<1%
114	nmsreview.org Internet Source	<1%
115	public-pages-files-2025.frontiersin.org Internet Source	<1%
116	www.ijarst.in Internet Source	<1%
117	"International Conference on Technology Advances for Green Solutions and Sustainable Development", Springer Science and Business Media LLC, 2025 Publication	<1%
118	"Soft Computing for Security Applications", Springer Science and Business Media LLC, 2023 Publication	<1%

119	Diep N. Nguyen, Vu Ly, Quang Nguyen, Dinh Hoang. "Generative AI for Cybersecurity - Fundamentals, Applications, Risks, and Opportunities", CRC Press, 2025 Publication	<1 %
120	Keerthi Garisa, Ravi Kant Kumar, Priyanka Singh. "Single-item training for multi-dish recognition: a class-agnostic framework for Indian food platters", Frontiers in Computer Science, 2026 Publication	<1 %
121	Lecture Notes in Electrical Engineering, 2016. Publication	<1 %
122	Md Rakibul Hasan, Md. Mahbubur Rahman, Fahim Shahriar, Md. Saikat Islam Khan, Khandaker Mohammad Mohi Uddin, Md. Mosaddik Hasan. "Smart farming: Leveraging IoT and deep learning for sustainable tomato cultivation and pest management", Crop Design, 2024 Publication	<1 %
123	Mohammad Monirul Islam, Morium Akter, Mohammad Shorif Uddin, Shihab Mahmud Anik, Md. Ataur Rahman. "Machine-Vision-Based Paddy Pest Recognition Using Deep Learning and Classical Machine Learning Methods", The Journal of Engineering, 2026 Publication	<1 %
124	Nie, Yuhao. "Short-Term Solar Forecasting from All-Sky Images Using Deep Learning", Stanford University, 2023 Publication	<1 %
125	Riccardo Jovine. "Chapter 3 Food and Protection of Geographical Indications",	<1 %

	Springer Science and Business Media LLC, 2024 Publication	
126	Tusharkanta Samal, Ambarish Panda, Manas Ranjan Kabat, Ali Ismail Awad, Suvendra Kumar Jayasingh, Deepak K Tosh. "Intelligent Computing Techniques and Applications - A proceeding of ICETICT – 2024", CRC Press, 2025 Publication	<1 %
127	Xiangyi Zhu, Wenli Zhang, Yingnan Sheng, Congrui Lv, Guorui Sheng, Weiqing Min, Shuqiang Jiang. "DPFA-net: a lightweight hybrid neural network with dual path feature aggregation for food image recognition", Multimedia Systems, 2026 Publication	<1 %
128	aei.pitt.edu Internet Source	<1 %
129	assets-eu.researchsquare.com Internet Source	<1 %
130	downloads.hindawi.com Internet Source	<1 %
131	iieta.org Internet Source	<1 %
132	ijimai.org Internet Source	<1 %
133	repo.lib.duth.gr Internet Source	<1 %
134	stax.strath.ac.uk Internet Source	<1 %
135	www.ir.juit.ac.in:8080 Internet Source	<1 %

136	www.iris.unict.it Internet Source	<1%
137	www.planhub.ca Internet Source	<1%
138	www.scirp.org Internet Source	<1%
139	"Computer Vision and Machine Learning in Agriculture, Volume 2", Springer Science and Business Media LLC, 2022 Publication	<1%
140	A. Abirami, S. Bhuvaneswari, R Aadhithya, R Devesh, P Deexith. "An Empirical Analysis on Medicinal Plants Associated with Impact on Biodiversity, Types, Challenges, and Technologies", 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS), 2023 Publication	<1%
141	Ghalib Ahmed Tahir, Chu Kiong Loo. "Explainable deep learning ensemble for food image analysis on edge devices", Computers in Biology and Medicine, 2021 Publication	<1%
142	Gianluigi Ciocca, Giovanni Micali, Paolo Napoletano. "State Recognition of Food Images Using Deep Features", IEEE Access, 2020 Publication	<1%
143	Mohammad Monirul Islam, Morium Akter, Mohammad Shorif Uddin, Md. Ataur Rahman, Md. Tarek Habib. "Machine Vision–Based Insect Recognition in Agriculture: A Comprehensive Review", Engineering Reports, 2026	<1%

Publication

144	Ouidadi, Hasnaa. "Machine Learning Solutions for Knowledge Transfer in Smart Additive Manufacturing Processes", Arizona State University Publication	<1 %
145	W. Shi, B. M. Golam Kibria. "On some confidence intervals for estimating the mean of a skewed population", International Journal of Mathematical Education in Science and Technology, 2007 Publication	<1 %
146	"Image Processing and Vision Engineering", Springer Science and Business Media LLC, 2026 Publication	<1 %
147	Kandakji, Lynn. "Probabilistic Modelling and Deep Learning for Multi-Dimensional Keratoconus Detection.", University of London, University College London (United Kingdom) Publication	<1 %
148	Kintoh Allen Nfor, Tagne Poupi Theodore Armand, Kenesbaeva Periyzat Ismaylovna, Moon-Il Joo, Hee-Cheol Kim. "An Explainable CNN and Vision Transformer-Based Approach for Real-Time Food Recognition", Nutrients, 2025 Publication	<1 %
149	Sakib Shahriar, Maria G. Corradini, Shayan Sharif, Medhat Moussa, Rozita Dara. "The Role of Generative Artificial Intelligence in Digital Agri-Food", Journal of Agriculture and Food Research, 2025 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		