

Detection of Food Adulteration using Machine Learning Tools

by

Anika Alamgir

24241161

Nashita Anjum

22201522

A. K. M. Shahadat Hossain

22201727

Debpriyo Hrishikes Deb Nath

24141221

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering

Department of Computer Science and Engineering
Brac University
January 2026.

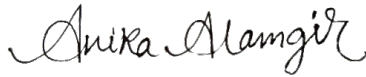
© 2024. Brac University
All rights reserved.

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:



Anika Alamgir

24241161



Nashita Anjum

22201522



A. K. M. Shahadat Hossain

22201727



Debpriyo Hrishikes Deb Nath

24141221

Approval

The thesis/project titled “Detection of Food Adulteration using Machine Learning Tools” submitted by

1. Anika Alamgir (24241161)
2. Nashita Anjum (22201522)
3. A. K. M. Shahadat Hossain (22201727)
4. Debpriyo Hrishikes Deb Nath (24141221)

of [Fall], [2025] has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science in 2025.

Examining Committee:

Supervisor:
(Member)



Dr. Chowdhury Mofizur Rahman

Professor
Department of Computer Science & Engineering
BRAC University

Co-Supervisor:
(Member)

Rafeed Rahman

Senior Lecturer
Department of Computer Science & Engineering
BRAC University

Thesis Coordinator:
(Member)

Dr. Md. Golam Rabiul Alam

Professor
Department of Computer Science & Engineering
Brac University

Head of Department:
(Chair)

Dr. Sadia Hamid Kazi

Associate Professor & Chairperson
Department of Computer Science and Engineering
Brac University

Abstract

Adulteration of food has been an urgent concern for the health and food safety of the people, and the affected area is particularly the developing countries due to the inefficient infrastructure and limited resources available to test the food items. Some traditional forms of detection, like chemical analysis and chromatographic analysis, are usually slow, costly, and require competent manpower. This paper explores an affordable and scalable Machine Learning (ML) system to identify food adulteration in sugar and poppy seed, and also another criteria is to identify the freshness of meat. The methodology of the study uses a modular approach to using self created RGB image dataset of sugar mixed with Magnesium Sulfate as adulterant which is one of its first kind in food adulteration detection of sugar and another one is of poppy seed dataset which is mixed with semolina as adulterant, collected through online to predict different fixed adulteration levels. Additionally, the dataset of meat is collected from an existing paper for further working on meat freshness detection. A self proposed ensembled model using pretrained models including: Densenet121, Conv-Next Tiny and Resnet50 with soft voting is used for detecting adulterants of sugar with Grad-CAM visualization, also showed some success in detecting other unknown adulterants. The ensembled model for sugar has shown about 88.96% accuracy. Whereas for poppyseed, Convolutional Neural Networks (CNN), and its pretrained model- EfficientNetB0 shows very high accuracy - 98.25% itself to measure the level of adulterants in class-wise manner without requiring any ensemble learning. As per previous works, there are few models that have satisfactory accuracy rate in terms of adulterant detection but there still exist issues, especially in covering diverse food items like sugar and poppyseed and as of for meat freshness, the Misclassification Cost based on some shown pretrained models was much higher in a previous work, which has been addressed by this study. For meat dataset, a Swin Transformer integrated with Explainable AI succesfully identified spots of spoilage, solving the black-box problem, showed increased accuracy upto 98% which is higher compared to that of the models shown in the existing paper from where the meat data set was collected and along with this, the study addressed the misclassification cost, an evalutation metric proposed by the original paper, which was lowered by many folds compared to the exisiting misclassification cost of the original paper. This research highlights that image data by mobile or obtained on accessible platform with AI can facilitate fast, dependable food safety testing and aid food quality inspection in supply chains.

Keywords: Food adulteration, Spices, Machine Learning (ML), Artificial Intelligence (AI), Convolutional Neural Network (CNN), Low-resource settings, Easily accessible device, RGB image analysis, Scalable detection system.

Acknowledgement

The authors would like to state that they are much indebted to Dr. Chowdhury Mofizur Rahman, Professor, BRAC University, Dhaka who has played the most important role in terms of continuously supervising, guiding and advising them through this research. They also like to thank Rafeed Rahman, Senior Lecturer, BRAC University who was working with us as co-supervisor, offered his experience and support throughout the development and assessment of this work. Additionally, Shuchismita Anwar, Research Assistant, BRAC University helped the authors with her advices whenever they needed it.

The authors are grateful to the Department of Computer Science and Engineering of BRAC University, which has enabled us to obtain the required academic setting, research facilities, and computing resources. The contributors of the publicly available datasets that were used in this study also allow us to be thankful as they made it possible to conduct the experiments and validate them.

The paper was done jointly by Anika Alamgir, Nashita Anjum, A. K. M. Shahadat Hossain, and Debpriyo Hrishikes Deb Nath, and we recognize contributions of teamwork, effort and dedication of these individuals to ensure the successful achievement of this research work.

Table of Contents

Declaration	i
Approval	ii
Abstract	iv
Table of Contents	vi
List of Figures	ix
List of Tables	x
Nomenclature	xii
1 Introduction	1
1.1 Background	1
1.2 Rational of the Study or Motivation	1
1.3 Problem Statement	2
1.4 Objective	2
1.5 Methodology in Brief	3
1.5.1 Use Cases	3
1.5.2 Data Preparation and Collection	3
1.5.3 Data Preprocessing	4
1.5.4 Model Architecture Selection	4
1.5.5 Sugar Ensemble Learning Strategy	4
1.5.6 Model Training and Optimization	4
1.5.7 Explainable AI (XAI) Integration	5
1.5.8 Generalization and Unknown Adulterant Testing	5
1.5.9 Performance Evaluation Metrics	5
1.5.10 Comparison with Existing Studies	5
1.6 Research Contributions	6
1.7 Scopes and Challenges	7
2 Literature Review	8
2.1 Preliminaries	8
2.2 Review of Existing Research	9
2.2.1 The Landscape of Food Adulteration and AI	9
2.2.2 Advanced Detection Methodologies	10
2.2.3 Applications in Oils Fruits	12
2.2.4 Meat Freshness Detection	13

2.2.5	Poppy Seed Adulteration	15
2.2.6	Visual Detection in Spices	16
2.2.7	Sugar & Honey Adulteration	17
2.3	Summary of Key Findings	18
3	Requirements, Impacts and Constraints	20
3.1	Final Specifications and Requirements	20
3.2	Societal Impact	21
3.3	Environmental Impact	22
3.4	Ethical Issues	22
3.5	Standards	23
3.6	Project Management Plan	24
3.7	Risk Management	25
3.8	Economic Analysis	25
4	Proposed Methodology	27
4.1	Design Process or Methodology Overview	27
4.2	Preliminary Design or Design (Model) Specification	28
4.2.1	Sugar Adulteration Design	28
4.2.2	Poppy Seed Adulteration Design	28
4.2.3	Meat Freshness Design	29
4.3	Data Collection	29
4.3.1	Sugar Dataset	29
4.3.2	Poppy Seed Dataset	30
4.3.3	Meat Freshness Dataset	31
4.3.4	Data Cleaning	32
4.3.5	Data Transformation	33
4.3.6	Data Integration	35
4.3.7	Data Reduction	36
4.3.8	Summary of Preprocessed Data	37
4.4	Implementation of Selected Designs	37
4.4.1	Sugar	37
4.4.2	Poppy Seed	41
4.4.3	Meat Freshness	41
5	Result Analysis	43
5.1	Performance Evaluation	43
5.1.1	Evaluation Criteria	43
5.1.2	Testing Methods	48
5.2	Analysis of Design Solutions	53
5.3	Final Design Adjustments	54
5.4	Statistical Analysis	55
5.5	Comparisons and Relationships	58
5.5.1	Sugar	58
5.5.2	Poppyseed	60
5.5.3	Meat	61
5.6	Discussions	62
5.6.1	Interpretation	62
5.6.2	Implications	62

5.6.3	Limitations	63
6	Conclusion	64
6.1	Summary of Findings	64
6.2	Contributions to the Field	65
6.3	Limitations and Future Work Plan	66
	Bibliography	71

List of Figures

1.1	Methodology Diagram	5
4.1	Sugar adulteration levels from pure to 40% adulterated	30
4.2	Poppy seed adulteration levels	31
4.3	Comparison of pure and semolina samples	31
4.4	Meat freshness levels	32
4.5	Flow Diagram of Sugar Dataset preprocessing	32
4.6	Flow Diagram of Poppyseed Dataset preprocessing	33
4.7	Flow Diagram of Meat Dataset preprocessing	33
4.8	Model Architecture of Sugar Ensemble Model	40
4.9	Model Architecture of EfficientNetB0 of Poppy Seed	41
4.10	Model Architecture of Swin Transformer for Meat Classification	42
5.1	Mahalanobis Distance Performance on Test Set	44
5.2	Grad-CAM Visualization of Sugar Dataset	44
5.3	Test ROC Curve of Sugar Dataset	46
5.4	Validation Accuracy of Poppyseed for EfficientNetb0	47
5.5	AI Attention Map using Grad-CAM	48
5.6	sugar mixed with urea samples	49
5.7	sugar mixed with saccharine samples	49
5.8	Identifying pure Sugar by Mahalanobis Distance	50
5.9	Detecting saccharine as unknown adulterant using Mahalanobis Dis- tance	51
5.10	Detecting urea as unknown adulterant using Mahalanobis Distance .	52
5.11	Confusion matrix of test data for Poppyseed Dataset	56
5.12	Confusion matrix of test data for the sugar dataset	57
5.13	Confusion matrix of test data for the Meat Dataset	58
5.14	Preliminary experimented model architectures for sugar adulteration detection.	59

List of Tables

2.1	Comparison of key findings and methodologies from the literature review.	19
3.1	Project Schedule and Deliverables (16-Week Timeline)	24
3.2	Resource and Budget Management	25
4.1	Datasets used in the experiments.	37
5.1	Performance comparison of each deep learning models used within the ensembled model based on validation and test sets for sugar adulteration	46
5.2	Overall performance of the ensembled model for sugar adulteration detection	46
5.3	Performance of Poppy seed adulteration dataset	47
5.4	Performance of Meat Freshness Classification Dataset	48
5.5	Comparative performance analysis of deep learning models on sugar adulteration datasets	60
5.6	Performance comparison of CNN models for poppy seed adulteration detection	61
5.7	Comparison of Proposed Swin Transformer Methodology against Existing State-of-the-Art Models [22]	61

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

AI Artificial Intelligence

ANN Artificial Neural Network

ARfD Acute Reference Dose

ATR – MIR Attenuated Total Reflectance - Mid-Infrared

CNN Convolutional Neural Network

DL Deep Learning

ENMs Engineered Nanomaterials

FTIR Fourier Transform Infrared Spectroscopy

GC – FID Gas Chromatography - Flame Ionization Detector

GC – MS Gas Chromatography - Mass Spectrometry

Grad – CAM Gradient-weighted Class Activation Mapping

HPLC High-Performance Liquid Chromatography

IoT Internet of Things

KNN K-Nearest Neighbors

LC – QTOF – MS Liquid Chromatography - Quadrupole Time-of-Flight Mass Spectrometry

LDA Linear Discriminant Analysis

LOD Limit of Detection

MCC Misclassification Cost

MgSO₄ Magnesium Sulphate

ML Machine Learning

NIR Near-Infrared

NMR Nuclear Magnetic Resonance
OOD Out-of-Distribution
PCA Principal Component Analysis
RF Random Forest
RGB Red, Green, Blue (Color Model)
ROC – AUC Receiver Operating Characteristic - Area Under the Curve
SVM Support Vector Machine
TTA Test-Time Augmentation
TVB – N Total Volatile Basic Nitrogen
ViT Vision Transformer
VOCs Volatile Organic Compounds
XAI Explainable Artificial Intelligence
XGBoost Extreme Gradient Boosting
YOLO You Only Look Once (Object Detection Model)

Chapter 1

Introduction

1.1 Background

The issue of food adulteration has been raising concerns about the status of food safety, civic health, as well as consumer confidence, specifically in the underdeveloped countries where foods are not regularly tested. Sugar, poppy seed, meat and spices etc, are some of the common widely adulterated food items. The intentions of adulteration agents can be either volume addition, appearance enhancement, economic gain, or cost reduction, but these adulteration agents can be very toxic, leading to many health related issues. The traditional food testing methods that are still common in the majority of developing jurisdictions involve chemical testing, chromatography, lab-based detection, etc. Besides being time-consuming and resource-reliant, these methods cannot be utilized in real-time or large-scale environments, and they also require the involvement of trained technicians, to say nothing of the fact that these methods are rather tedious and resource-dependent.

With the increased pace of AI advancement, chances are that these disadvantages can be addressed with alternatives that are smart, fast, and automated detectors. AI-based systems and other technologies, such as image processing, spectroscopy, ML, and DL, have proven to be potentially useful in the quick and efficient identification of adulterants in food items. The techniques offer non-destructive, scaling solutions that are applicable even at field-level or consumer-end applications. Some researchers have also explored the use of AI in the detection of adulteration.

Such kinds of potential of AI are furthermore illustrated in the papers in processing complex visual and spectral data that can detect the objects more fast and accurately than the traditional methods. Given the scenario, this paper will dwell on known AI-based methods that are applied in adulteration analysis of sugar and some other common foods. It will evaluate technologies, algorithms, datasets, and evaluation metrics used in the ongoing research. The objectives of the paper are the identification of strengths, gaps, and future directions, which will serve as the basis of the creation of an all-embracing, AI-based adulteration detection model that is suggested in this paper.

1.2 Rational of the Study or Motivation

The research rationale is seen in the fact that there is an urgent need in developing simple, cost effective, and readily available ways of detecting food adulteration

particularly in developing nations such as Bangladesh where food safety facilities are insufficient. A number of rural and semi-urban regions do not have certified laboratories, equipment and trained human resources, which gives the food adulteration an unknown status and exposes the consumer to severe health effects of poisoning, organ damage and cancer as well as eroding consumer confidence in the food supply chain. The contemporary laboratory-based test cannot be used by the ordinary consumer, which opens the case of constant and unnoticed exposure to the toxic material. Available literature reveals key gaps in the widely eaten foods: the sugar adulteration is well documented, but AI-based and non-destructive detection systems are uncommon; poppy seeds are often economically adulterated and natural toxicity is also common, but currently, detection depends on expensive laboratory systems, and despite the availability of meat freshness detection systems, they are not yet ready to be utilized effectively on a real-time basis. This paper will seek to fill these gaps through the application of affordable AI, computer vision and sensor-based technologies that can be used with common tools such as smartphone cameras and IoT devices. This study will help to offer the practical and real-time food safety solutions that can address the adulteration problem highlighting on sugar, poppy seeds and meat freshness in an effort of empowering consumers, maintaining fair trade, minimizing health hazards, and assisting the policymakers to enhance food safety laws.

1.3 Problem Statement

Although AI-based food adulteration detection is promising, the presence of the following issues remains:

- The majority of AI models have poor generalization to food categories and focus on individual food items.
- The available solutions are often laboratory-based and have not been optimized to be used in real time and on-site.
- The data employed in the current studies are either limited in size or restricted to a certain area or devices.
- They are missing open-source, inexpensive platforms that are suited to be used in developing countries with little infrastructure.

Hence, the question that this study will answer is:

What is the easiest and most cost-efficient way to create a scalable AI-based system of detecting food adulteration that will be applicable to various types of food and can be implemented in a real-world, low-resource setting?

1.4 Objective

The main aim of this thesis is to devise and implement a food adulteration detection system that will analyze some frequently adulterated common food substances, namely poppy seed, sugar, etc. This research attempt will:

- To design an AI system to detect adulteration in sugar and poppy seeds as part of the developing stage of research.
- To create and optimize a developed deep learning model to classify meat freshness that works well in detecting spoilage and gives great accuracy as well as reduce MCC.
- To utilize transfer learning methods to achieve satisfactory detection rates and provide a scalable framework as a substitute to lab tests.
- To use self-created smartphone datasets of sugar along with publicly available datasets to detect adulteration.
- To implement the performance of the various algorithms and feature extraction strategies on each selected food category differently instead of taking a one-size-fits-all approach.
- Create, train and analyze models for each of them that are realistic, can be easily accessible, and applicable to real-life, particularly in low-resource environments and test the performance, accuracy, and drawbacks of the individual models on publicly available datasets and self created dataset.

The study focuses on specifics rather than generalization, and each food item is treated differently to enhance the reliability of detection and relevance to the user.

1.5 Methodology in Brief

The research focuses on the problem of food adulteration and freshness detection in low-resource environments using RGB image-based Machine Learning approaches. Three primary use cases are considered: sugar adulteration detection, classification of poppy seed adulteration levels, and meat freshness detection.

1.5.1 Use Cases

The study addresses the following food safety tasks:

- Detection of sugar adulteration
- Classification of adulteration levels in poppy seeds
- Meat freshness detection

1.5.2 Data Preparation and Collection

Sugar Dataset

An RGB image dataset was self-constructed by mixing sugar with Magnesium Sulfate at predetermined adulteration levels. Images were captured using readily available imaging devices such as mobile phones. This dataset represents one of the earliest RGB image-based datasets for sugar adulteration detection.

Poppy Seed Dataset

An online-sourced RGB image dataset of poppy seeds adulterated with semolina was collected. Multiple adulteration levels were defined as class labels to enable supervised learning.

Meat Freshness Dataset

Meat freshness analysis was conducted using a publicly available dataset from prior research, allowing direct performance comparison with previously reported models.

1.5.3 Data Preprocessing

All images were resized and normalized according to the input requirements of the selected deep learning models. Data augmentation techniques such as rotation, flipping, and scaling were applied to reduce overfitting and improve generalization. Each dataset was divided into training, validation, and testing subsets.

1.5.4 Model Architecture Selection

Sugar Adulteration Detection

Three pretrained deep learning architectures—DenseNet121, ConvNeXt-Tiny, and ResNet50—were selected for feature extraction and later combined into an ensemble model.

Poppy Seed Adulteration Detection

A CNN-based approach using EfficientNetB0 was employed due to its strong performance and high computational efficiency.

Meat Freshness Detection

A Swin Transformer architecture was adopted to capture both spatial and contextual features relevant to meat spoilage.

1.5.5 Sugar Ensemble Learning Strategy

Outputs from DenseNet121, ConvNeXt-Tiny, and ResNet50 were combined using a soft-voting ensemble classifier. Mahalanobis distance-based anomaly detection was integrated to identify adulterants other than Magnesium Sulfate. This ensemble strategy aimed to improve robustness, reliability, and generalization, including the detection of unknown adulterants.

1.5.6 Model Training and Optimization

Supervised learning was performed using categorical cross-entropy loss for sugar adulteration detection and ReduceLROnPlateau scheduling for poppy seed classification. Hyperparameters such as learning rate, batch size, optimizer selection, and number of epochs were empirically tuned for each food category. Regularization

techniques were applied to prevent overfitting in sugar models, while early stopping was used for the poppy seed dataset.

1.5.7 Explainable AI (XAI) Integration

Grad-CAM was used to visualize adulterant-sensitive regions in the ensemble sugar model. For meat freshness detection, Explainable AI techniques were employed to highlight localized spoilage areas, reducing the black-box nature of deep learning models and improving trustworthiness.

1.5.8 Generalization and Unknown Adulterant Testing

The sugar ensemble model was evaluated on samples containing previously unseen adulterants to assess its generalization capability beyond the trained adulteration types.

1.5.9 Performance Evaluation Metrics

Model performance was evaluated using accuracy, confusion matrices, precision, and recall. For meat freshness detection, Misclassification Cost (MCC) analysis was conducted to assess real-world safety risks and economic impact.

1.5.10 Comparison with Existing Studies

For the meat dataset, the proposed models were directly compared with previously reported architectures using the same dataset. Quantitative improvements in accuracy and misclassification cost were analyzed relative to the baseline study. Overall, the methodology follows a systematic and unified pipeline across three distinct food categories, enabling comprehensive analysis of both adulteration detection and freshness assessment using AI-based approaches.

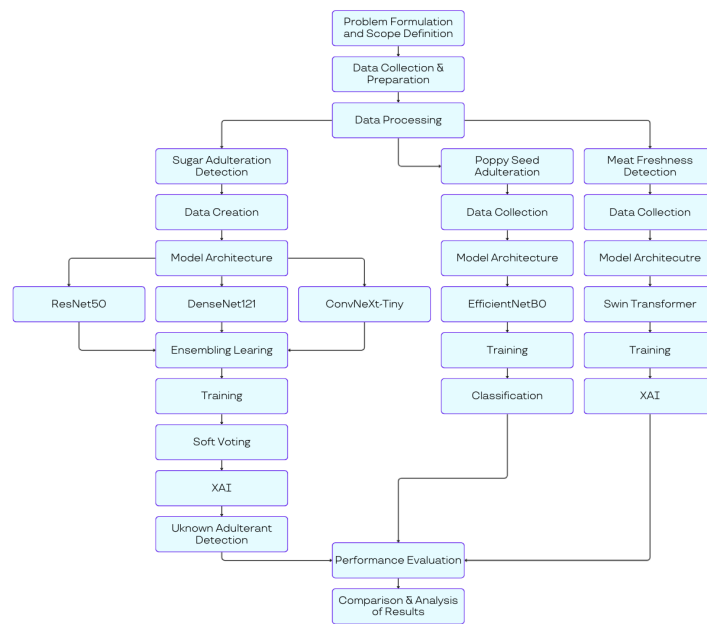


Figure 1.1: Methodology Diagram

Fig. 1.1 showcases the overall methodology followed across all the three datasets uniformly using image data and how the overall study was conducted.

1.6 Research Contributions

1. One of the first RGB image-based machine learning frameworks of sugar adulteration detection which would be using Magnesium Sulfate as an adulterant, on a food matrix that has not been well studied.
2. Presents an experimentally controlled and self-constructed image dataset of sugar adulteration at predetermined concentration levels, and recorded with the help of available imaging devices like mobile phone.
3. Creates an ensembled deep learning model based on DenseNet121, ConvNeXt-Tiny, and ResNet50 with a soft-voting classifier, which has been enhanced in terms of robustness and reliability on adulterant detection.
4. Integration of Mahalanobis distance-based anomaly detection to identify previously any unknown adulterants in sugar samples, though it the model was not trained with it.
5. Usage of Grad-CAM-explainability to visualize adulterant sensitive areas, which would work basically to minimize the black-box behavior of deep learning models in food safety.
6. Introduces a prediction system of adulteration level of poppy seeds by class based on the CNN-based approach, which only takes the EfficientNetB0 model and yielded very high accuracy (98.25%), requiring no ensemble learning. Gathers and uses an online-obtained poppy seed adulteration dataset (contaminated with semolina) to increase the variety of data in a field with limited labeled data.
7. Expands to meat freshness detection with the use of a Swin Transformer architecture and performs better on the dataset compared to the models reported before. Suggests a much less expensive misclassification cost structure on detecting freshness of meat, which is a fundamental limitation of previous research from where meat dataset was collected. Applies Explainable AI (XAI) methods to point out areas of spoilage in meat images and enhance their understandability to consumers and regulators.
8. Explores multiple food items adulteration detection using diverse models but with RGB images under the same study.
9. Focuses on a low cost, scalable and deployable system that has been specifically built to address resource limited environments, utilising image data transmitted by mobile phones or easily accessible devices.

This study offers a review with multiple food items and different techniques and models used for adulteration and freshness detection using AI-based pipeline, which is different for different food items and it is able to showcase multiple food safety tasks, adulteration detection and freshness assessment.

1.7 Scopes and Challenges

The scope of this research comprises the development of a generic food adulteration detection system that can impact community health at scale and can be implemented across a variety of common food types, like meat, poppy seeds, and sugar. by training a set of ML and Deep Learning models having the capability to facilitate fast and non-destructive testing. The study aims to explore mainly image-based data captured through mobile phone cameras or easily accessible equipments for training and evaluating models, as well as collecting the necessary features of food items through a dataset available online or from any form of accessible hands-on experiment. The developed models with the help of this research might be adapted into lightweight and user-friendly systems for future, with the potential to be integrated into mobile or desktop applications, allowing consumers, vendors, or health inspectors to perform rapid, surface-level testing with ease. In a nutshell, this research work has the potential to provide a cheaper, expandable and accessible solution to support the people in the bigger run and communities to be able to detect whether their food is contaminated or adulterated and by extension, create more transparent food systems, build greater consumer confidence, and help policymakers and health officers in their attempts to crack down food fraud more efficiently and within shorter times compared to manually chemical-testing every time to detect food-adulteration.

Now, coming to the context of challenges, one of the major challenges can be the availability of quality and labeled datasets as per requirements for this study pertaining to a wide range of food items, as it is most important to teach an effective machine learning algorithm. Information in most instances is unavailable or patchily recorded related to food adulteration, especially for food products like sugar, or even some general features can be missing from the other datasets. The other big challenge is on model generalization, where after training the models, they are expected to perform highly correctly not only on the current controlled data but also on completely different real-life data that has diverse settings. Also, hardware limitation is an issue since the system is to work effectively without addressing the need of a super high-end computing power as it is aimed to serve the public quicker and at lower cost compared to manual-testing, hence provision of mid to low-latency capabilities that would be applicable either in the mobile or desktop size implementation landscape. Another difficulty can be faced while appropriately mixing impurities with pure sugar in exact amounts using the required instruments, keeping accurate records, and later detecting the exact amount of impurities from the adulterated sugar that had been made manually for the study. If not mixed correctly in the right proportion and if it is not correctly recorded, then the model might provide wrong results, and the expected results might not match the mixed impurities amount, which will eventually result in difficulties in finding out model accuracy, and food adulterants would not be correctly detected. Lastly, the fact that the amount of adulterants and food composition and texture varies in different products adds a level of complexity, and thus, it is hard to have a single solution that can fit all products. There can be unique adulteration patterns on each of the commodities, and the detection methods should be different, too, and finding the right ML or Deep Learning models for each selected food item, giving the best accuracy, too, is a challenge itself that will require being overcome.

Chapter 2

Literature Review

2.1 Preliminaries

Food adulteration is a common global problem and it is largely motivated by economic motives whereby high-value foodstuff is deliberately contaminated with a low quality or dangerous food substances to gain higher profit. In the recent research, food adulteration is explained as one of the increasingly dangerous threats to humanity because of its severe outcomes on the population health and financial situation [1]. Traditional detection techniques, including chromatography and chemical analysis, are strong and usually destructive, laborious and time-consuming. These shortcomings have given rise to an increasing trend in the use of Non-destructive, fast, and scalable methods of assessing food quality by using Artificial Intelligence (AI) and Machine Learning (ML) solutions as well [2]. Various food commodities present different adulteration. Due to the crystalline structure and high demand of sugar in the market, sugar is very prone to adulteration. Other typical adulterants include chalk, urea, plastics and washing soda that are hard to recognize through simple physical tests, and are extremely dangerous to the human body once ingested. Poppy seeds are susceptible to both economic adulteration with less expensive look-alike seeds, and natural contamination with opium alkaloids, such as morphine and codeine. It has been pointed out by regulatory authorities like the European Food Safety Authority that there are health implications associated with such alkaloids, recent research has indicated that treatment procedures play a key role in determining the level of toxicity [3]. Growth of microbes and chemical transformations are the main factors linked to adulteration and spoilage of meat, which are frequently determined by such indicators as Total Volatile Basic Nitrogen and biogenic amines [4]. In contemporary detection systems of food adulteration, machine learning methods have been in the center stage. Convolutional Neural Networks and the use of image-based models to detect visual adulteration patterns have been highly utilized with high accuracy in a variety of food products [5]; [6]. Other algorithms like Support Vector Machines and ensemble models like Random Forest and XGBoost are also used in spectral and sensor-based data because the algorithms are robust and able to classify nontarget data as well as other data types in their classification models [7]; [8]. More often, these models are combined with state-of-the-art sensing technologies, including gas sensors, hyperspectral imaging, and electronic noses, to record chemical and physical indicators of food samples without the damage of the food sample in a non-destructive way of operation [9].

All in all, AI-based models and the latest sensing technologies should be viewed as a solid base toward the creation of efficient, accurate, and scalable food adulteration detection systems. Based on these initial ideas, the paper will seek to fill the current gaps in the literature especially those areas that are not well studied including sugar adulteration but will utilize established computational and sensing models that are documented in the recent literature.

2.2 Review of Existing Research

2.2.1 The Landscape of Food Adulteration and AI

In the paper by [2], the authors were interested in the comprehensive summary of food adulteration, its reasons, health risks, as well as various detection techniques with regard to both the traditional approaches to detect food adulteration and the new methods to detect food adulteration based on AI. The main contribution that the paper makes is in the consolidation of various detection strategies, including chromatography, spectroscopy, chemometric models, and electronic sensors that can be applied to all types of adulterants in milk, honey, spices, oils, and meat. Although it does not introduce a new dataset, the review summarises findings of numerous other studies and describes datasets obtained through laboratory experiments, hyperspectral imaging, Raman spectroscopy, and chromatographic fingerprints. Spectral signatures and chemical compositions were generally used as the set of features and the absence or presence of adulterants as labels. According to the results of referenced studies, machine learning techniques with a special focus on principal component analysis (PCA) and neural networks performed highly in terms of classification accuracy and reliability when applied in food authentication. The authors highlight a significant research gap, namely, previous research tended to target single adulterants or methods of detection, which could not be generalized. It is a unique review that combines multidisciplinary evidence and suggests AI-enhanced, cost-effective, and field-deployable solutions, thus enhancing previous disjointed methods and offering the future of scalable adulteration detection systems. A review article by the authors of the present study (Sneha et al., 2023) focuses on the transformation of traditional food testing towards the modern computer-based approach. The authors contrast the old methods, like chemical testing, sensory testing, etc., with new Machine Learning (ML) and Deep Learning (DL) methodologies. Since the article is a review one, there is no one experimental sample or accuracy value; the results of different researches are summarized in this article on different samples, such as milk, meat, honey, and spices. The review shows that methods such as spectroscopy and electronic nose, with the help of algorithms, such as Convolutional Neural Networks (CNN) and Random Forests, can effectively detect contaminants such as water in milk or pork in beef. The authors determine that these AI-based techniques, in most cases, are more efficient and quicker than manual testing, but rely on large, quality datasets to train the computers and in this regard, this is also a major concern.

[10] looked at the effectiveness of AI techniques along with sensor and ML technologies in discovering whether edible oils or other food products are adulterated. The research found that ANN, SVM, CNN, and Random Forest (RF) are very effective at classifying adulterated samples, reaching accuracy from 81.2% to 100%. Because

mustard oil is not examined in this case, comparable studies on olive, sesame, and avocado oils give important insights. The studies made use of spectral (including FTIR, NIR), dielectric, and imaging measurements that were gathered in laboratory tests, with different numbers of samples. Spectral absorptions, dielectric characteristics, and RGB images were included as features, whereas the labels mainly showed purity or the kind of adulterant present. The document refers to past studies using hyperspectral imaging and spectroscopy, proving that CNNs score better than classical analysis methods in spotting food adulteration. As standardization and real-time detection are lacking, the paper points out that using solid frameworks and sensor fusion can greatly help. By combining several analyses, the study advances existing research by offering a single approach to AI food authenticity and points out ways of using AI in detecting mustard oil adulteration.

According to [11], the study covers the history of AI applications in food authentication, giving particular importance to the detection of adulteration by various ML and DL models. In this paper, the authors assess AI approaches like SVM, Random Forest, ANN, CNN, and ensemble methods in terms of their use across many food matrices, such as edible oils. Some of the datasets mentioned are FTIR, Raman, NMR, GC-FID, thermographic, and data based on images. They come from different sources, ranging from tens to many thousands in number (for example, hundreds of thousands of oil droplet images were used in certain CNNs). The set of features checks fatty acid amounts, how light is prominently reflected, and specific textures visible in pictures, and most labels show adulteration levels or the main type of oil. Models based on CNNs performed well, reaching accuracies between 98% and 100 %, outdoing PCA and LDA. It explains that older methods did not work very well for finding unknown or light adulteration, and it points out how AI-based tools based on images and sensors have taken over. These findings are important for building mustard oil adulteration detection techniques, especially for systems that use ML with a mix of approaches and images.

[12] looked in detail at how machine learning is used to detect when food is adulterated, mainly focusing on milk. In the paper, machine learning models are sorted into supervised, unsupervised, and reinforcement learning, giving a clear picture of their differences and use in several studies. Examples of significant models are SVM, ANN, CNN, PCA, PLS, and Bayesian networks. Researchers used data from FTIR, NIR, chemical compositions, and image-based information, including experiments with more than 6,000 images. Features used for the extraction are spectral characteristics, the composition of fatty acids, and gradient images in the scene, and the samples are typically categorized as pure or adulterated, further splitting according to types of adulterants. The performance of the algorithm is measured with accuracy at up to 98%, F1-score, precision, recall, entropy, and RMSE. The study uses the findings to note that making ML work in real time and on a large scale is still a problem.

2.2.2 Advanced Detection Methodologies

[7] address the serious risk of "unknown" food adulterants by developing an AI system which finds abnormalities in raw milk without requiring the search of particular chemicals, tackles the grave threat of so-called unknown adulterants of foodstuffs such as the notorious melamine scandal. The scientists used two complex machine

learning algorithms (ExtraTrees and XGBoost) to process two different types of standard chemical data and light-spectrum measurements (FTIR) on a large industrial dataset of more than 65,000 milk samples. The system was very effective with the capability of giving a very high accuracy of known adulterants and the ability to flag blind samples with never-before-seen contaminants such as salt and caustic soda with an F1 score around 0.87. The authors conclude that such a non-targeted strategy is better than the current governmental safety standards (including the state of China GB 19,301-2010) since it could identify novel forms of fraud at extremely low rates with data that is required by dairy companies to be collected on a daily basis. In this paper, [13] describe a system that relies on machine learning to find out if Khesari dal is present in Toor dal, since its consumption is linked with serious health problems. The main tool to classify images is the CNN, whereas the YOLOv5 object detector ensures real-time detection and identification of impure grains in the system. The models were chosen because they help to find and handle complex aspects like color, shape, and texture, which are key in telling apart different grains. Images of both real and unfairly diluted dal were obtained using a digital or smartphone camera and collected by me in proper and consistent settings. Even though no number for samples was provided, there were the same number of pure and adulterated samples were used in the dataset. The system processed the images' details as inputs and assigned labels according to the DAL type. The suggested method gave an accuracy rate of approximately 92%, meaning it worked well. The work points to existing studies on food adulteration detection made with RNNs, electrochemical sensors, and spectroscopy. Since earlier methods were not scalable and had many limits, this work develops an automated system capable of detecting food quality issues at any time, so food control is more efficient and accurate.

[14] suggest using machine learning with real-time sensors alongside spectral analysis for identifying trace adulteration in edible oils (of vegetable or animal origin). There are two sets of data in the study: the first collects 160 samples from ATR-MIR spectroscopy and has 130 features, showing how much vegetable oil has been added (0%, 6.25%, 25%, 50%); and another set is obtained using MEMS and multichannel gas sensors during the oil fumigation process. Out of the sensor dataset, each class contains 280 samples, and the features are the gases' conductance values for NO_2 , $VOCs$, CO , and H_2 . Using accuracy, precision, recall, specificity, and F1-score, the models were assessed: K-NN, Random Forest, CATBOOST, and XGBOOST. All the metrics reported that Random Forest and XGBOOST attained perfect scores (100%). Although previous work in spectroscopy and chemometrics succeeded, this study looks for solutions to practical problems by designing a low-cost and portable system that detects in real-time. For use in detecting adulterated mustard oil, the research supplies an example of bringing together refractive index values, gas sensor responses, and machine learning so that the detection is reliable.

Another paper related to food adulteration detection by [15] that proposes an advanced method where fumigation techniques are combined with machine learning. The authors did not depend only on the conventional chemical tests but instead they depended on fumigation of the food samples to release the volatile compounds of the food samples which were tested to form a chemical fingerprint. The researchers trained a Support Vector Machine (SVM) and a random forest model using a large, 15,000 record Adulterant Detection Dataset, to differentiate between tampered and pure food. The research discovered that the fumigation statistics enhanced the out-

put of the two models. The model providing the best results was the Random Forest model plus fumigation with an accuracy of 98% and a precision of 0.99, which was much better than the SVM model with a 96% accuracy. The authors reach a conclusion that such an integrated method is a rapid, non-destructive, and a very high accurate method of food safety in the supply chains.

2.2.3 Applications in Oils Fruits

In their recent work, [16] suggest a combined system called SensorNet, which helps detect adulteration in fruits with an extremely high level of accuracy through computer vision and sensors. The main benefit of the study comes from using deep learning to assess images and real-time gas sensors, which provides a complete and useful approach for food safety control. The researchers created a dataset with RGB images and information from chemical sensors for apple, mango, banana, and malta, classifying them into eight types, that was fresh or adulterated. Some of the methods used in this study involve Decision Tree, Naïve Bayes, a 23-layer CNN called DurbeenNet, and a sensor fusion model called SensorNet. SensorNet turned out as the most successful model, having 97.03% accuracy and surpassing both DurbeenNet (96.71%) and classic classifiers. Earlier studies had used photographs separately from sensor data, so this research developed a strong system that processes both at once. This research uses earlier food safety and ML foundations to provide instant and large-scale adulteration detection that might be helpful for mustard oil investigation.

[17] studied using computer models to detect any adulteration in edible oils. In the study, a publicly available dataset from Kaggle was used, which included ATR-MIR spectra and chemical characteristics like viscosity, refractive index, and the fractions of fatty acids, both in pure and palm oil-adulterated mustard oil, as well as other oils. This study used four ML algorithms, namely Logistic Regression, SVM, Random Forest, and Neural Networks. When it comes to accuracy, Random Forest did the best (97.2%), then SVM (94.8%), Neural Networks (92.5%), and Logistic Regression (89.5%). Calculating fatty acids and making decisions on the important features helped the model work better. Metrics for evaluation were accuracy, precision, recall, F1-score, and confusion matrices. The authors used results from previous research on Random Forest sensors, biosensors, and spectroscopy for spotting adulteration in foods. However, they pointed out that those techniques have shortcomings related to cost, scaling up, and necessary infrastructure. Unlike research in the past, this study connects ML, spectral information, and feature engineering to ensure the system can work in real time with the help of IoT. As a result, machine learning demonstrates how easily it can help detect fraud in oil.

In this recent paper, Strati it has been analyzed that 26 plant-based edible oils, among them mustard seed oil, using different analytical methods and machine learning to check oil quality and its authenticity. While the main benefit of the research is that it tries to resolve the classification problem of oils, three different types of analyses, GC-FID, ATR-FTIR, and colorimetry, were used in combination with machine learning models to do so. There were 52 oil samples (26 different samples over two years) taken from commercial oil sources in Greece. Information was obtained for fatty acids, FTIR readings at 3008 and 1654 cm^{-1} , and color parameters, while the labels say if the sample is pulp or seed oil. In the study, the classifiers known

as LDA and Random Forest recorded the highest accuracies, at 96.0% and 88.0%, respectively. This study takes previous oil authentication research with FTIR a step further by making use of different analytical approaches as well. The gap addressed is the lack of real data and enough ways to distinguish different types of oils. For this reason, the approach offers a model for safer oil authentication and advises methods that can help spot fake mustard oil.

In this paper that is related to oil adulteration detection, [18] offer a new method for finding out if edible oil is mixed with mustard oil through the joint use of spectroscopic imaging equipment and artificial intelligence algorithms. The research team suggests a way to classify oils, pure or adulterated, through captured spectral profiles across 224 different wavelengths in the range of 400 to 1000 nm. There are 670 samples in the data, and these include 70 clean and 600 adulterated samples of mustard, olive, coconut, and almond oils mixed with castor, sunflower, and liquid paraffin on a Specim FX10 HSI camera. The main attributes are spectral reflectance and absorption, and labels indicate that there are 11 distinct groups of “oils.” A number of ML algorithms were considered, for example, SVM, Logistic Regression, Random Forest, KNN, Decision Tree, Naïve Bayes, and LDA. In both the internal and external checks, LDA proved itself to be the best performer, registering 100% accuracy, precision, recall, F1-score, and kappa. Unlike earlier techniques that were time-consuming and pricey, the study corrects issues in integrating spatial and spectral information when judging oil quality. This research performs better than earlier studies made with NIR, FTIR, or fluorescence in detecting cases of food adulteration, and could be put into industrial practice for food quality control.

In a recent paper, [19] suggested a new way to detect when mustard oil is being adulterated using low-frequency dielectric spectroscopy together with multivariate linear regression. Researchers studied mustard oil mixed with argemone oil in different percentages (2%, 4%, 6%, and 8%) and built a statistical model by studying its dielectric properties. The data set held dielectric constant and dielectric loss measurements taken with a range of frequencies (1 kHz to 10 MHz) and temperatures (30°C to 50°C). In model training, frequency, temperature, and percentage of adulteration were used as features, and dielectric properties were taken as the response. The results of the adjusted R^2 were over 0.96, and the significance of the results was also shown through paired sample t-tests ($P < 0.05$). The study takes previous research on dielectric analysis of edible oils further and seeks to fill the void in information on using dielectric spectroscopy to detect adulteration in mustard oil by ignoring such studies in the literature. Unlike before, it deals with health-advised adulteration of argemone oil and uses statistical approaches to boost the reliability of its predictions.

2.2.4 Meat Freshness Detection

In a review by [4], modern methods to determine whether meat is fresh or rotten are studied, and traditional lab tests, which are slow, are intended to be replaced. The authors examine literature in the field to analyze the implication of meat spoilage by bacteria and chemicals such as Total Volatile Basic Nitrogen (TVB-N) in 2010 to 2020. They test the high-tech devices like: Electronic Noses (which detect the gases of spoilage) and smart packaging (labels which turn color when meat spoils). Being a review, it does not give one accuracy score but it explains that these new

sensors are extremely sensitive and quick. The conclusion of the paper is that these non-destructive technologies will play a vital role in the future of food safety as they give instantaneous responses as opposed to the old ones that have taken days before detecting bacteria.

A 2022 review by [20] investigates the use of Engineered Nanomaterials (ENMs), such as gold and silver nanoparticles, to create high-tech sensors for checking the freshness of meat and seafood. Instead of relying on a single dataset, the authors analyze results from multiple studies to compare these new sensors against traditional lab tests. They evaluate performance using metrics like the Limit of Detection (LOD) and response time, focusing on the sensors' ability to detect chemical spoilage markers like Total Volatile Basic Nitrogen (TVB-N) and biogenic amines. The results show that nanomaterial sensors are extremely sensitive and fast, enabling the creation of "smart packaging" that changes color when food goes bad. However, the review concludes that before these sensors can be widely used in grocery stores, scientists must solve problems related to the safety (toxicity) and stability of the nanomaterials when they come into contact with food.

In the review by [20], examine the application of Engineered Nanomaterials (ENMs) including gold and silver nanoparticles to develop high-tech sensors to determine the freshness of meat and seafood. The authors do not use one dataset, but instead consider the outcomes of several studies, comparing new sensors to the traditional laboratory procedures. They measure performance in terms of the Limit of Detection (LOD) and response time which they consider in the ability of the sensors to respond to the chemical spoilage indicators such as the Total Volatile Basic Nitrogen (TVB-N) and biogenic amines. These findings demonstrate that nanomaterial sensors are very sensitive and quick, and they could develop a smart packaging, which would alter color upon food spoilage. The review, however, concludes that until these sensors can be extensively applied in grocery stores, scientists need to resolve issues to do with the safety (toxicity) and stability of the nanomaterials in case they come into contact with food.

An example of potential advances to bring freshness testing of chilled meat to an entirely new stage is the research by [21], examining how to use so-called nanozymes the synthetic nanomaterials that mimic biological enzymes to detect freshness. In contrast to natural enzymes, which are delicate and expensive in most cases, nanozymes are tough, stable and cost efficient to make. The review outlines the construction of these artificial enzymes into sensors capable of detecting certain chemical fingerprints of rotting meat including hydrogen peroxide, biogenic amines, and hypoxanthine. These sensors are either colorimetric (change of colour such as a litmus test) or fluorescent, and are also capable of real-time monitoring. The authors conclude that nanozyme sensors provide an effective new resource to the meat industry and smart packaging can warn consumers about spoilage, but other challenges such as food safety and influence of the food matrix should be addressed before the mass application.

The general idea of the study behind [22] is to predict the freshness in meat through an AI-based solution, basically basing it on computer vision, as well as the deep learning models. They studied in premise of image characteristics of meat in order to to evaluate the quality and degradation of meat samples which in this instance ultimately proves the capability of machine learning in food quality evaluation. Though in the case of this study, the model that was found to give the highest accuracy in

identifying the level of freshness, the Misclassification Cost (MCC) of it was found to be higher. But another model that had lower accuracy and low MCC was deemed to be a better model because it minimized the cost incurred by a business. This indicates that, although the system was generally good, it did not always work well with all the categories of freshness and hence this is where the point of additional optimization is highlighting the need to optimize the system more.

[6] had created an integrated IoT architecture to classify the meat species (beef vs. mutton) and freshness of meat quality (fresh vs. spoiled) in real time. The authors used a custom-created dataset of 4,400 images evenly distributed in four classes, which then were trained with a custom Convolutional Neural Network (CNN) to reach a validation accuracy of 99.63, which is by far superior to known architectures, such as VGG16 and ResNet50. In this system, visual analysis was enhanced with environmental data received through gas (MQ-135, MQ-4) and temperature sensors, which enabled creating a holistic evaluation of such signs of spoilage as the amount of ammonia and methane. It was concluded that the study provides an extremely accurate and non-invasive alternative to conventional manual methods of meat inspection as this multi-modal method is available through a convenient mobile application.

2.2.5 Poppy Seed Adulteration

An opinion on the scientific opinion on the safety of poppy seeds, which can be contaminated with opium alkaloids (e.g. morphine and codeine) during harvesting, is given by [3]. The paper examined data sets of different European nations and established a safety threshold, known as the Acute Reference Dose (ARfD), of 10 -g per kilogram of body mass of such chemicals. The researchers discovered that although the harmful alkaloids can be washed, ground, or baked off the seeds by between 25% and 100% of the seed content, even when the foods are eaten that contain large proportions of poppy seeds people still tend to surpass the safe limit. The report concludes that the present levels of alkaloids in selected food products have a health effect on many age groups and laments that there is no evidence on the other individual alkaloids such as thebaine that could be even more harmful.

This research paper by [23] examined higher-order methods of detecting economically-driven adulteration in poppy seeds by the use of both Liquid Chromatography-Quadrupole Time-of-Flight Mass Spectrometry (LC-QToF-MS) and microscopic imaging. It was a study of 23 botanically identified and 15 commercial seed samples of *Papaver somniferum*, both genuine *Papaver somniferum* and common adulterants like amaranth, quinoa, sesame, and *Nigella sativa*. The chemical profiling identified 47 alkaloids that included morphine, codeine, and thebaine with a concentration of 0.8-223mg/kg, 0.2-386mg/kg, and 0.1-176mg/kg, respectively. The validation method was ICH compliant, and linearity (r^2 over 0.999) and precision (less than 4 per cent) were excellent. The discriminative features were profiles of alkaloid concentrations and seed microstructure, where the classification was of the authentic and adulterated samples. This research compared morphological and high-resolution chemical fingerprinting to the previous research in which analysis was primarily done on GC-MS or LC-MS/MS, and this technique better differentiated poppy seeds at the terminal of the procedure between poppy seeds and adulterants. Furthermore, it addressed the existing lack of concomitant morphological and chemical authentica-

tion and offered a strong and reproducible model of food safety and quality control that is more comprehensive and reliable than the traditional chromatographic equipment.

A more recent experiment by the authors of the article by [24] investigates the possibility of grinding poppy seeds as a form of safety before reducing the toxic chemical content of the poppy seeds. The researchers compared the traditional high-friction grinders with the modern cooler industrial mills by using LC-MS/MS to analyze the blue poppy seeds that had high initial morphine content (approximately 20mg/kg). The findings revealed that conventional grinding that naturally raises the temperature of the seeds to above 100degC through friction, was able to reduce morphine by about 3% and thebaine by about above 60%. Nevertheless, contemporary industrial grinders which maintained the seeds at low temperatures recorded an almost 0 percent decrease in toxicity. The authors find that the heat and not the physical process of crushing is the main factor that destroys the harmful alkaloids thus passing traditional grinding as a more practical way of guaranteeing food safety.

A 2022 study by [25] developed a highly accurate method using Liquid Chromatography and Mass Spectrometry (LC-Q-TOF-MS) to measure toxic opium alkaloids in commercial blue poppy seeds. Analyzing 15 different brands available in Polish markets, the researchers achieved high accuracy (recovery rates between 87.5% and 112%) and successfully detected five distinct alkaloids, including morphine and codeine. The results were concerning: morphine levels varied widely from 2.14 to 22.84 mg/kg, and codeine ranged from 0.28 to 2.45 mg/kg. The authors concluded that in 14 out of the 15 samples, eating a standard 50g serving of poppy seeds would cause a person to exceed the European safety limit (ARfD) for morphine, highlighting a significant health risk despite food processing efforts.

In a paper by [25], researchers created a highly precise technique, based on Liquid Chromatography and Mass Spectrometry (LC-Q-TOF-MS), to quantify toxic opium alkaloids in commercial blue poppy seeds, which can also be found in its study. The researchers studied 15 brands existing on the Polish markets and the accuracy of the analysis was high (recovery rates of 87.5 to 112 percent), and the authors were able to identify 5 different alkaloids, which includes morphine and codeine. These findings were alarming: morphine and codeine were 2.14-22.84 and 0.28-2.45 mg/kg respectively. The authors found out that consuming a normal portion of poppy seeds of 50g of food was going to make an individual cross the European safety limit (ARfD) of morphine in 14 out of 15 samples, a high risk to health despite food processing attempts.

2.2.6 Visual Detection in Spices

An example of a study that tackles the problem of food fraud is by [5] who discusses a case of food fraud whereby, high-cost turmeric powder is added to the inexpensive wheat flour to maximize gains. To address it, the authors designed an artificial intelligence system based on a deep learning model named VGG-16 and applied it to digital images of the spice to identify impurities. The authors produced a vast amount of 3,000 photos of pure turmeric and samples of 10- to 30-percent concentration of wheat flour mixed with it. The system was tested on 600 previously unseen images on which the computer was supposed to be blind, and it was found that the system did exceptionally well with a 94.33 percent accuracy and was able to

identify the fake samples most of the time. The review concludes that the method of computer vision is a quick, cheap and reliable alternative to conventional laboratory analysis in assuring quality of spice in the market.

2.2.7 Sugar & Honey Adulteration

[8] introduces an effective method to find honey that has been tampered with by joining hyperspectral imaging with machine learning models. The research uses ANN, SVM, KNN, Random Forest, and Decision Tree to examine spectral signatures from hyperspectral images and judge if honey is pure or mixed with other ingredients. You can access the dataset on Figshare, and it covers 8,675 hyperspectral samples coming from 12 plants and seven different brands, all of which contain 128 spectral features. An example of a binary category would be pure milk versus adulterated milk, while there are also many levels of adulteration (e.g, 0%, 5%, 10%, 2,5%, and 50%). ANN was found to have the highest accuracy and F1-score by performance evaluation conducted using 5-fold cross-validation, which is 98.9%; Decision Tree came in second at 95.3%, and SVM was third at 94.5%. The study's method is safe and suits various types of honey, as it has been well-prepared using extensive pre-processing, feature mathematical reduction, and hyperparameter improvement before the final tests. Therefore, it shows that this method can be effectively applied to find cases of food adulteration in practice.

An example of a study which deals with the issue of fake Sidr honey, which is a high-value product usually mixed with inexpensive cane sugar syrup, is provided by the study by [26] on the topic of classification. The scholars adopted a light-based method referred to as fluorescence spectroscopy that did not cause any harm to the honey samples but verified their purity. They discovered that pure Sidr honey produces light at a certain wavelength (color) that appears to be at variance with other types of honey such as Acacia and sugar-syrup mixed honey. Through these light patterns, after analyzing them using a computer learning technique, the Bagging algorithm, they were able to categorize the samples. The machine was very accurate, as it could easily differentiate between pure honey, sugar water and all types of mixture of the two. The authors conclude that this technique is a quick environmentally friendly and dependable quality control tool. They say that it may be ultimately utilized to generate handheld devices that assist both the sellers and buyers to check instantly whether Sidr honey is genuine.

In a recent review, by [27] note a more environmentally friendly approach to the production of nanoparticles with the help of sugarcane rather than using costly and/or harmful substances. The authors explain how cane sugar juice, leaves, and even waste products (bagasse) are natural elements that assist in converting metals such as silver, copper and iron into small and useful particles. These sugarcane particles are useful in various tasks, including the removal of polluted water (dye degradation), biological change sensing, and harmful bacteria killing. The paper concludes that although this technique has succeeded in converting agricultural rubbish into riches, we have to conduct further studies to make sure that these particles are environmentally safe and on how to make them in high volumes.

The literature has extensively discussed sugar adulteration as a severe food safety challenge caused by economic stress and the increase in world demand [1]. This paper brings into focus the characteristics, health effects and detection method of

this type of food fraud. The authors list different adulterants, including chalk, urea, plastics, washing soda, etc., that are deliberately used to weight down packaging, add color, or be able to retain moisture. These chemicals pose extreme health hazards, the health hazards differ to kidney stones and lead poisoning (chalk), respiratory distress and skin irritation (urea and washing soda). The paper has explained detection techniques that are as basic as water dissolution tests to elaborate spectroscopic and chromatographic techniques. Such sophisticated techniques as HPLC, NMR, and FTIR are offered as critical to the determination of chemical compositions and the identification of organic adulterants that may escape simple physical methods. In the final analysis, the paper highlights that the measures that are being taken are inadequate. The authors draw a conclusion and suggest the creation of an all-inclusive adulteration database, better state control and the creation of affordable detection technologies to protect the health of the population against such a worldwide threat.

2.3 Summary of Key Findings

The literature review reveals a significant shift in food safety research. Traditional chemical testing methods (like chromatography) are being replaced or supported by Artificial Intelligence (AI) and Machine Learning (ML). These modern methods are faster, non-destructive, and cost-effective. Recent studies show clear trends in AI-based food adulteration detection. Computer vision methods are widely used, especially for spices and fruits, where simple digital images combined with deep learning models like CNNs have achieved high accuracy. For example, [5] reported an accuracy of 94.33% in detecting wheat flour in turmeric using only image data. Another important trend is sensor fusion, where combining different data sources improves performance. Studies using both visual data and gas sensor (fumigation) data consistently show better results; for instance, [9] increased model accuracy from 96% to 98% by adding fumigation data. Research on liquid foods such as milk and edible oils is also well developed, with models like Random Forest and LDA frequently achieving accuracy levels between 97% and 100% ([18] ; [17]). However, despite extensive research on adulteration in products like honey, meat, and oils, there is a clear research gap in AI-based studies focused specifically on sugar adulteration. Most existing work on sugar mainly discusses the problem in general rather than proposing technical AI-driven solutions [1]. The findings suggest that while AI is mature for commodities like olive oil and raw milk, it is underutilized for solid crystals like sugar. Furthermore, the success of "non-targeted" detection (finding unknown impurities) in milk [7] proves that AI can catch new types of fraud that standard tests might miss. This validates the need for a dedicated AI approach and a new dataset for sugar adulteration detection.

Table 2.1: Comparison of key findings and methodologies from the literature review.

Food Domain	Methodology/Input	Best Model	Key Result/Accuracy
Milk	FTIR Spectra + Chemical Data	ExtraTrees + XGBoost	99.24% Accuracy [7]
Pulses	Digital Images	CNN + YOLOv5	92.0% Accuracy [13]
Fruits	Images + Gas Sensors	SensorNet (Fusion)	97.03% Accuracy [16]
Edible Oils	FTIR Spectra	Random Forest	97.2% Accuracy [17]
Mustard Oil	Hyperspectral Imaging	LDA	100% Accuracy [18]
Meat	IoT Sensors + Images	Custom CNN	99.63% Accuracy [6]
General Food	Fumigation + GC-MS	Random Forest	98.0% Accuracy [15]
Spices	Digital Images	VGG-16 (Deep Learning)	94.33% Accuracy [5]
Honey	Hyperspectral Imaging	ANN	98.9% Accuracy [8]
Sugar	General Review	N/A	Identified as a global menace [1]

Chapter 3

Requirements, Impacts and Constraints

3.1 Final Specifications and Requirements

This section lists the technical, methodological and resource needs required to reproduce the experimental findings and also install the suggested food adulteration detection systems. These specifications include the hardware infrastructure, software environment, and data constraints needed to run the decoding of the DenseNet121 model in sugar adulteration quantification, the decoding of the EfficientNetB0 model in detecting poppy seeds adulteration, and the decoding of the Swin Transformer model in the classification of meat freshness. These requirements guarantee the reproducibility of the experiment, stability of the model, and consistent performance with deploying the models.

Hardware-wise, deep learning models, and especially computationally intensive models like the Swin Transformer and the DenseNet121, are expensive to train and evaluate with high-performance computing equipment. Applications requiring large scale matrix operations, backpropagation, and gradient computation will require systems with NVIDIA GPUs with CUDA acceleration (Tesla T4 or V100). To obtain the data manually, the sugar dataset had to be created manually with the use of the Pixel 7 Pro smartphone and its macro photography functionality to take photos of fine crystalline structures. Besides this, a digital balance scale of laboratory quality with accuracy of 0.01 g was necessary to measure ratios of $MgSO_4$ with accurate values and provide ground-truth.

An open-source stack was used to create the software environment so that it would be scalable and reproducible. All the experiments were written in Python (version 3.10 or later). In all datasets, PyTorch was used as the main framework of defining models, training, and evaluation. Pretrained architectures were loaded using Torchvision, and image transformations were performed using it, whereas stratified k-fold cross-validation was supported by Scikit-learn and the calculation of evaluation metrics, i.e., precision, recall, and F1-score. In freshness analysis of meat, pytorch-grad-cam library was used to provide the Generation of Grad-CAM visualization of the Swin Transformer, which allows the model to be interpreted. Visual analysis was done with Matplotlib and Seaborn with the help of confusion matrices and training convergence plots.

There were strict data requirements in order to ensure robustness and consistency

between tasks. The sugar data was a collection of 3,200 balanced images at five adulteration levels, each of which was photographed on a high contrast black background to enable analysis of granular texture. The poppy seed data set consisted of nine categories due to the adulteration degree and variation in the background thus the necessities of CutMix augmentation and weighted random sampling techniques due to an uneven distribution of classes. The dataset of meat included 2,266 RGB samples, which belonged to Fresh, Half-Fresh and Spoiled categories, and the data loaders were set to three channels to receive the information.

The methodologically, each input image was made to conform to pretrained encoders input requirements by standardizing it to 224 x 224 pixels. ImageNet standard deviation and ImageNet mean were used to normalize the input tensors so that they could optimize and converge steadily. The sugar dataset was optimized to 16 batch size and the meat and poppy seed data to 32 batch size so as to balance between memory and gradient stability. All experiments used the AdamW optimizer, with initial learning rates and ReduceLROnPlateau learning rate scheduler dataset specific to overcome convergence stagnation in training.

3.2 Societal Impact

The influence of this work on the society is immense since food adulteration and spoilage is directly connected with the risk of population health, consumer safety, and also economic losses. The proposed system can be used in mitigating circulation of unsafe or substandard food products by offering an automated, scalable and also a comparatively useful way of adulteration detection using image data from easily accessible devices.

Proper detection of adulterated sugar and poppy seeds samples aid in protecting the consumers against chronic and acute health complications brought about by the detrimental adulterants. In the same way, detection of spoiled meat can reduce the delivery of unsafe animal products to customers thus lowering the cases of foodborne diseases. There are numerous economic benefits associated with the proper detection of meat freshness in addition to the food safety concern. The model identified can minimize false positive freshness prediction, which reduces financial loss and reputational risk by businesses because it greatly lowers the MCC of freshness prediction. Misclassification Cost (MCC) is a cost-sensitive assessment measure that approximates the likely loss, in terms of finances and health-related costs, due to poor model forecasting, instead of treating all the errors equally. It finds special application in the food safety field, where false identifications that include the wrongly labeled spoiled meat as fresh can have serious health impacts and result in the loss of a lot of money. Compared to the other models can be compared with safety and hence, MCC has the ability to optimize models based on safety and minimization of real-world risk and harm by putting the errors of predictions, in terms of money. This accuracy aids in improved inventory, wastages and consumer confidence which improves the level of total profitability. On a larger scale, the system will reinforce the work of the regulator and increase the clarity of the food supply chain.

The study also brings about cultural and dietary security especially in the areas where such food products are eaten regularly. The system ensures that the product is authentic and hence consumer trust is always maintained in the local markets. Such a proposed idea could facilitate national food safety programs in the long

run and help the regulatory authorities to conduct active quality monitoring and detecting violations quickly.

3.3 Environmental Impact

This study has a vast positive implication for environmental sustainability. The system prevents unnecessary food waste through early and correct detection of adulteration and spoilage, and also helps minimize the environmental impact caused by food waste or simply adulteration of food. The proposed solution, in contrast to the traditional laboratory-based testing methods, which need chemical reagents, solvents, and a large amount of energy resources, is based on digital imaging and AI inference which have a smaller ecological footprint.

Moreover, quality control systems that run in real-time in the food processing industries can minimize the number of contaminated or wasted products and result in the better use of resources. This helps in the larger objectives of sustainable production and consumption trends in tandem with the global food security measures.

3.4 Ethical Issues

The development and deployment of this system involve important ethical considerations. The data integrity of the models is ensured by training them using clean and well-structured datasets which have been collected manually in case of sugar which did not require any outside permission and as well as collected under controlled conditions through publicly available dataset platform in case of poppyseeds which also did not have any external permissions or privacy related issue. There is an effort to prevent the introduction of bias when preprocessing data, augmenting it and in model training. Pretrained CNNs were employed and fine-tuned in order to detect sugar and poppyseed adulteration detection and in accordance with the standard ethics, transparency, and reproducibility were ensured. So, it did not have to deal with any issues of data privacy and consent, and although the model architecture is not new as they are pretrained, the application, adaptation, and enhanced performance are what make this work novel. Another major dataset involves the meat freshness dataset which was also collected through another published paper previously. By employing the use of MCC as one of the key evaluation metrics for the meat freshness detection purpose, the performance of the models can be assessed in a more balanced and reliable manner, which has been done by reducing bias towards majority classes over minority classes or edge cases. Although MCC is not novel to this paper and is adopted from prior studies, the mentioned approach in this paper also utilizes a pretrained model that demonstrates better accuracy compared to the pretrained models referenced in the original paper.

The issue of food safety is a moral requirement and the system will be helping advocate consumer rights to safe and pure products. Additionally, the study does not involve any personal or sensitive information or data and hence there is no concern of privacy breach. The data collection, labeling, training, and evaluation phases of ethical research practices are upheld and the models are developed responsibly and transparently.

3.5 Standards

The proposed study is modeled based on internationally accepted codes or standards on food quality assurance and development of artificial intelligence systems. It is founded on the food safety principles which are set by The Food and Agriculture Organization and the World Health Organization, which is primarily preemptive prevention, continuous monitoring, and effective reaction to quality and security concerns. Technically, the work is based on the principles of the International Organization of Standardization (ISO/IEC 23053) to design AI systems, which ensures the transparency, reproducibility and impartiality in the modeling pipeline. A set of widely used and accepted performance measures including test accuracy, precision, recall and F1-score which were used to provide a whole and complete analysis of the models in terms of adulteration detection of sugar and poppy seed. Whereas, it used the MCC, a highly permissible and scientifically sound measure, to the meat freshness detection problem to be more sensitive to cost-sensitive misclassification cases. Since the proposed model was also built on pretrained architectures, the application of the Matthews correlation coefficient MCC allows a fair and standard comparison with the existing methodologies available in the literature, which will assure methodological rigor and enhance the merit of the evaluation.

3.6 Project Management Plan

Table 3.1: Project Schedule and Deliverables (16-Week Timeline)

Phase & Timeline	Key Activities	Methodological Focus	Deliverables (Outputs)
Phase 1: Data Preparation (Weeks 1–4)	• Manual Sugar mixture prep (0–40%) & Macro photography. • Poppy Seed dataset cleaning & RGB conversion. • Meat dataset integration & class mapping.	• Standardization • Balancing (640 img/class for Sugar) • Normalization	• Three verified, balanced, and normalized datasets. • Preprocessing pipelines.
Phase 2: Model Implementation (Weeks 5–8)	• Sugar: Implement DenseNet121, ResNet50 pipelines. • Poppy: Setup EfficientNetB0 with custom classifier. • Meat: Coding Swin Transformer with shifted windows.	• Architecture selection • Transfer Learning setup • Custom Data Loaders	• Functional, error-free training scripts. • Baseline model architectures ready for training.
Phase 3: Training & Optimization (Weeks 9–12)	• Poppy: Execute 10-Fold Stratified Cross-Validation. • Sugar: Apply Mixup regularization & tuning. • Meat: Train Swin Transformer with XAI hooks.	• Hyperparameter Tuning (ReduceLROnPlateau) • Overfitting Control (CutMix, Mixup)	• Optimized Model Weights. • Target Accuracy: >80% (Sugar), >95% (Poppy/Meat).
Phase 4: Evaluation & Reporting (Weeks 13–16)	• Generate Confusion Matrices & F1-Scores. • Meat: Generate Grad-CAM heatmaps. • Comparative analysis & final documentation.	• Statistical Analysis • Explainability (Glass Box verification)	• Final Performance Reports. • Visual Interpretability Maps. • Project Thesis/Paper.

Table 3.2: Resource and Budget Management

Resource Category	Specifications / Requirements	Budget / Allocation Strategy
Computational Hardware	• Cloud-based GPU Instances (NVIDIA Tesla T4 or V100). • CUDA-enabled workstations.	• Allocated via Academic Cloud tiers or Pro subscriptions (Est. 200+ GPU hours).
Data Acquisition Tools	• Pixel 7 Pro (Macro Mode) for Sugar imaging. • Laboratory-grade digital balance scale (0.01g precision).	• One-time material cost for procurement. • Utilized existing laboratory infrastructure where possible.
Software Stack	• Python 3.10+, PyTorch, Torchvision. • Scikit-learn, Grad-CAM, Seaborn.	• Utilized open-source libraries to ensure reproducibility and minimize licensing fees.
Human Resources	• Researcher/Developer (Data labeling, coding, analysis).	• Conducted as primary research activity (No external labor cost).

3.7 Risk Management

Several potential risks are considered, identified and addressed through targeted mitigation strategies. The imbalance of the classes and limited dataset sizes were addressed with the help of large-scale data augmentation, stratified splitting, and also the use of MCC as a robust evaluation metric for the meat dataset. To minimize overfitting, label smoothing, drop out, early stopping and adaptive learning rate scheduling were taken into consideration. The trained model was refined end to end without any layer freezing or unfreezing techniques, and it enabled the network to adjust itself entirely to the target data. Minimizing deployment issues is achieved with modular system design, which facilitates versatile integration into different environments. In the case of the meat freshness detection module, a misclassification cost matrix is provided such that errors that can cause health risk to the population, like, misclassification of spoiled meat, are kept to a minimum, therefore minimising the losses to the health and finances of the population.

3.8 Economic Analysis

The economic feasibility of the system discussed in this paper is a key strength of the proposed approach. Conventional forms of food adulteration testing can be costly, time consuming and need special lab facilities. Conversely, the cost per-sample of the AI-based detection system is low with expenses primarily associated with imaging and model inference. The collection of datasets for poppyseeds and meat were free as collected from publicly available datasets and published paper. However, the data collection of sugar required a qualityful mobile with a macro

mode camera and buying of the sugar and its adulterant along with digital balance for accurate measurement. The cost-benefit analysis is most important in meat freshness detection. The misclassification cost matrix is integrated to measure the cost of errors in models and in fact showed that the total cost of misclassification is USD 16.00 in total in the test and zero misclassification of spoilage, which is 18 times lower than the models used in the study of the inspired paper. This points out the economic and social health worth of the model. Besides, the cost incurred to run the model is minimal when compared to the possible gains on a reduction in the spoilage cost, quality control, and consumer confidence. The system is easy to incorporate in the current quality assurance processes, which is economical both by regulatory organizations and food industries.

Chapter 4

Proposed Methodology

4.1 Design Process or Methodology Overview

This study design is based on the development of deep learning-related models [28] aimed at conducting the visual detection of food adulteration and quality evaluation, including sugar, poppy seed, and meat food samples. The algorithm was designed with the principle objectives of measuring adulteration levels and freshness in real but controlled imaging conditions and overcoming usual drawbacks associated with the scarcity of data and overfitting of models. With the sugar data, a collection of deep convolutional neural network (CNN) structures was taken into account to determine their ability to differentiate between pure and adulterated crystalline textures, i.e. to measure the fixed class concentration of $MgSO_4$. The final candidate architectures were ResNet50 [29], ConvNeXt-Tiny [30] and DenseNet121 [31], chosen to make a comparison between standard residue learning and dense connectivity plus transformer-inspired architecture. EfficientNetB0 [32] was selected as the main architecture in the case of poppy seed dataset. This was selected because it has the best balance of computation efficiency and high texture recognition ability which is essential in the separation between fine grained semolina particles and poppy seeds. Regarding the meat freshness detection task, the architecture changed to an alternative of CNNs, namely, Swin Transformer [33]. This architecture was chosen because it can be used to model local spoilage indicators, as well as, global texture variations using shifted window attention. Moreover, to overcome the issue of the black box of the traditional classification, the layout to develop the meat analysis included Explainable AI (XAI) with Grad-CAM, which meant that the decisions the model made (Fresh vs. Spoiled) were visualizable by human interactions. In the three food categories, research methodology was made reproducible, adaptive, and consistent. A standardized pre-processing and augmentation pipeline (e.g. Mixup and geometric transformations) was used to achieve similar learning conditions and metrics of fair performance evaluation. During the research, the work is still in the development phase, thus various model configurations are assessed here. The resultant model performance in comparison between these varied datasets and experimental conditions will be found out in the subsequent phases as the final optimized system.

4.2 Preliminary Design or Design (Model) Specification

This study design will involve three separate machine learning pipelines, each one being shaped to the physical properties of the particular food commodity: to measure crystal mixtures in sugar, to detect grain poppy seeds, and to detect biological food spoilage in meat.

4.2.1 Sugar Adulteration Design

In order to solve the problem of determining the concentration of $MgSO_4$ in the granulated sugar (between 0% and 40%), this research has transcended the conventional dichotomous approach. Three different architectures were chosen in order to test the trade-off of feature reuse vs. spatial hierarchy and an ensembled model with voting classifier was created using: ResNet50, ConvNeXt-Tiny, and DenseNet121. The choice of denseNet121 was made due to the dense connection structure in it, since the feature maps of all preceding levels are re-used. This property is essential in the identity of the very fine or fine textural difference between sugar and $MgSO_4$ crystals, which are visibly almost the same. ConvNeXt-Tiny was added to add modern principles of Vision Transformer (ViT)-style designs, including large kernel sizes and layer normalization, into a convolutional architecture, which may be able to capture a wider context of space. ResNet50 was the powerful base. The training model adopted a personalized ensemble with PyTorch. To avoid overfitting to the medium-sized dataset (3,200 images), Mixup regularization with a value of $\alpha=0.2$ was used, where the model was trained on linear combinations of images as a means to smooth the boundary of the decisions. Training used AdamW optimizer with differential learning rates that would work for pretrained backbone layers and again also it used newly initialized classification heads that would together with a learning rate warm-up phase which would be then followed by cosine annealing to ensure stable convergence.

4.2.2 Poppy Seed Adulteration Design

In order to analyse the poppy seed data, EfficientNetB0 was chosen as the main candidate model. The main reason why this model was taken into consideration is that it uses a compound scaling approach, which also maximises depth, width and resolution. It works especially well in relation to this data set in which the semolina grains have almost the same shape and surface texture as poppy grains and a compromise between high-resolution feature retrieval and computational performance is needed. ImageNet pretrained weights were used to initialize the architecture to make use of hierarchical feature extraction. The default classifier was substituted with a custom head with a dropout layer (probability = 0.3) and completely connected linear layer with the number of classes of interest. It was trained with AdamW and the learning rate of 0.001 and the L2 weight decay of 0.01. In order to balance the skewed distribution of classes, CutMix augmentation with a weight of 1.0 and probability of 0.5 (CutMix), plus WeightedRandomSampler and a loss weighted by the class was used. A 10-fold stratified cross-validation strategy was adopted to validate the

model to be able to have a strong generalization of cross-validation across all the levels of adulteration.

4.2.3 Meat Freshness Design

In the detection of meat freshness, the philosophy of design changed to Explainable AI (XAI) as opposed to pure classification. Spoilage of meat is biologically complex (in terms of discoloration, slime, and change of texture), so a model was needed that could model long-range dependencies over the image surface. As the basic architecture, the Swin Transformer (Hierarchical Vision Transformer using Shifted Windows) was chosen. The Swin Transformer does not use fixed grids to compute the images like standard CNNs but instead uses shifted windows to compute self-attention and enable the model to focus more on both local indicators of spoilage (such as small dark spots) and global texture degradation. Most importantly, it is designed with a Grad-CAM (Gradient-weighted Class Activation Mapping) interpretability module. This would make sure that the system does not simply give a label (e.g., "Spoiled"), but it also produces a visual layer of heatmap, which shows the particular areas that caused the decision. Such a Glass Box design is critical in the development of trust in food safety application where human operators can ensure that the AI is identifying real spoilage and not the background artifacts.

4.3 Data Collection

Three different datasets were used to ensure a comprehensive evaluation of the proposed adulteration detection model which included a manually developed self made sugar dataset and another dataset which is collected from the Mendeley Data Repository [34] which is also developed under controlled conditions. As well, the meat data used in the study was obtained following the study of the meat freshness prediction by [22] who used the Meat Freshness Image Dataset [35]. Although the original work refers to red meat in particular, it offers a professional structure on the analysis of three types of spoilage including Fresh (FS), Half-Fresh (HF), and Spoiled (SP).

4.3.1 Sugar Dataset

The sugar data set was elaborated manually in the controlled laboratory-like environment that guaranteed the accuracy of the measurements and the standardization of the imaging requirements to overcome the lack of publicly available datasets to analyze quantitative sugar adulteration. The preparation was done manually and intentionally with such a high priority to remove the occurrence of inconsistency of data to avoid any data having inaccurate labels when it comes to adulteration category. This was a controlled methodology that made every sample have the correct level of adulteration in it, which is reliable since deep learning experiments can be supervised. The adulterant of interest $MgSO_4$ was chosen because it has a great visual similarity with the sugar crystals in the form of granules, and thus can be almost impossible to distinguish using the naked eye. The data were sorted into five different categories according to the concentration of adulterant by weight pure sugar (0), and blends of 10, 20, 30, and 40%. Digital balance scale was used to carefully

weigh all samples to ensure consistent ratios as well as to avoid cross contamination during preparation hence ensuring integrity of the dataset. The acquisition of images was conducted with the help of a Pixel 7 Pro smartphone camera working in the macro mode that allowed capturing fine crystalline structures that are essential in determining subtle changes in the texture between pure sugar and $MgSO_4$. In order to reduce background noise and enhance feature extraction, the images were taken on a black cloth of non-reflective background to ensure maximum contrast between white crystals and background. The illumination was controlled during the acquisition process to minimize the shadow artifacts and to eliminate variation in illumination. Also, the photos were taken at various viewpoints to enhance diversity of the dataset and minimize the chances of overfitting when training the model. The resulting data was a total of 3,200 pictures, which were well balanced to avoid bias in the classes. In every category of adulteration, there were precisely 640 images with the five levels of concentration (0%, 10%, 20%, 30%, and 40%).

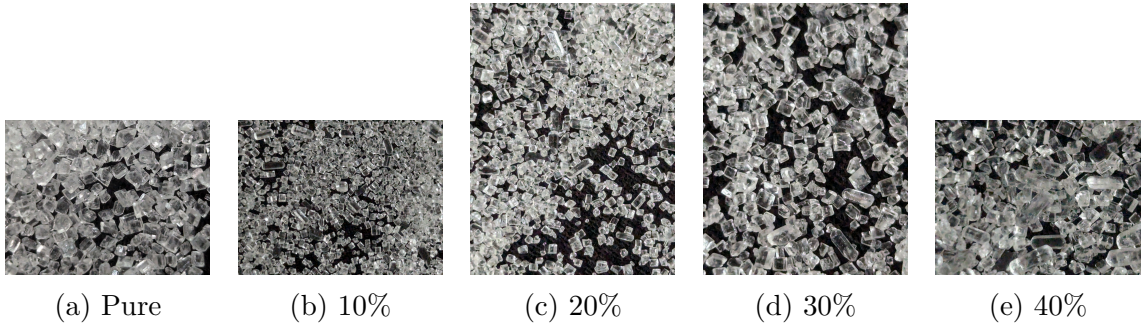


Figure 4.1: Sugar adulteration levels from pure to 40% adulterated

This balanced and diverse dataset presented a strong baseline of a deep learning model to train and test on the task of quantifying granular texture-based adulteration using DenseNet121, ConvNeXt-Tiny, and ResNet50.

4.3.2 Poppy Seed Dataset

The data on poppy seeds which was retained from the Mendeley Data Repository, consisted of pure and adulterated poppy seed samples which was stored there in folder format depending upon different levels of adulteration (10,20,30,40 and 50% adulteration). Semolina particles have the same size, almost similar shape, and surface texture as poppy seeds, but they are much smaller than the real poppy seeds under normal imaging conditions, which makes them hard to differentiate with the naked eye when mixed in limited amounts with poppy seeds. So, semolina particles were chosen to make the adulterated dataset. The images for the dataset were taken on cameras with high-resolution under controlled lighting conditions on even surfaces and with both white and black backgrounds, to make sure the potential differences in variations of different images caused by different lighting and other environmental conditions can be reduced.

This dataset had nine classes with different levels of adulteration, starting from 0% to 50% on different backgrounds, such as black and white and also it had zoomed in view and normal view of the datasets, which allowed the model to capture subtle differences between pure and adulterated seeds of different concentrations, even under with different lighting and background conditions. Its carefully controlled

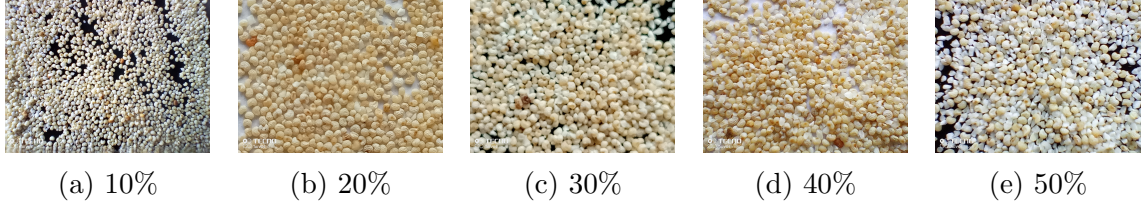


Figure 4.2: Poppy seed adulteration levels

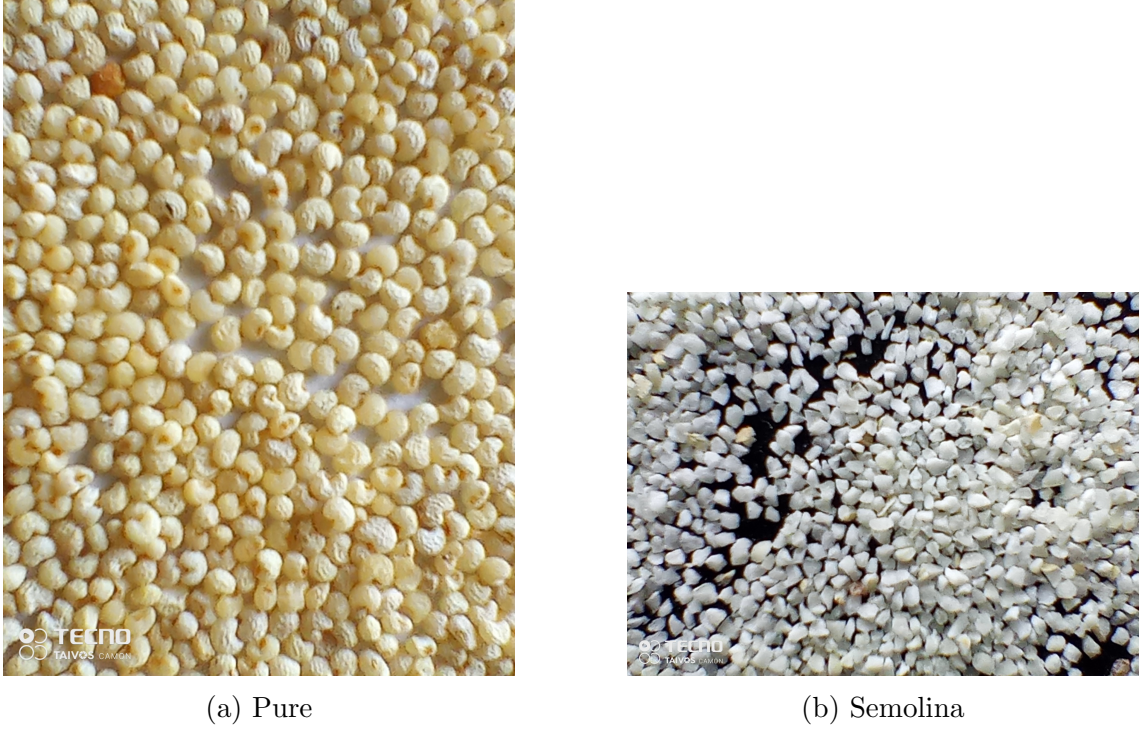


Figure 4.3: Comparison of pure and semolina samples

photo placement and clearly defined photo labelling made it useful to evaluate how well the deep learning architectures of this model could generalize to fine-grained texture classification tasks in a broader and more generalized context [32].

4.3.3 Meat Freshness Dataset

The dataset of meat used in this study was obtained in previous studies on the prediction of the freshness of meat by researchers of the article by [22] using the open-access Meat Freshness Image Dataset organized by [35]. The repository selected as the source of the analysis of the progressive stages of biological spoilage is the red meat, though it is a rather specific repository chosen due to the established and specialized framework of the analysis of the progressive stages of its spoilage. It has been extensively used in the assessment of meat quality, which is why it is a valuable method to benchmark state-of-the-art deep learning and explainability-based methods. The dataset contains a total 2266 high-resolution images that are classified under three different classes that represent different critical stages of meat degradation. The Fresh (FS) category has 889 pictures with bright red color and firm texture, which can be characterized as high-quality and unspoiled meat. The

Half-Fresh (HF) category consists of 815 images of a transition period when changes of the first stages of oxidation and slight discoloration can be found and the quality of the goods begins to diminish, without a total loss of form. The Spoiled (SP) group is composed of 665 images of images clearly spoiled with signs of surface darkening or greening, the appearance of slime, and the severe texture degradation, making the meat unsafe to eat.

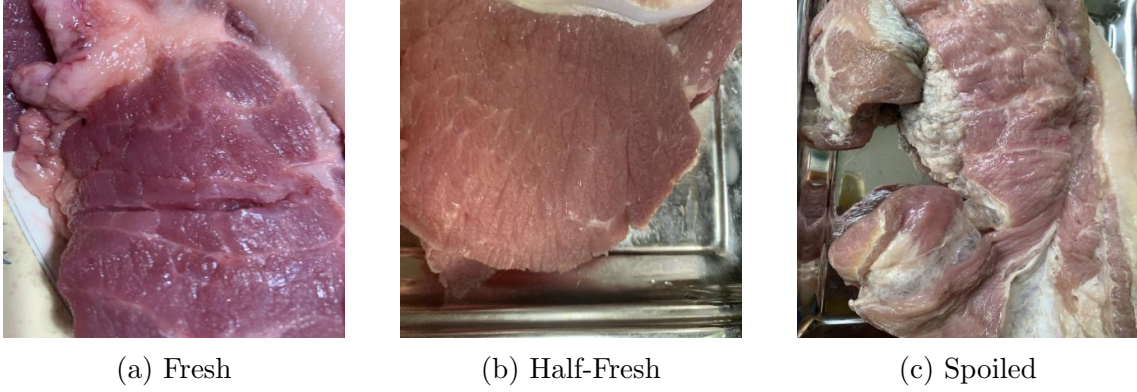


Figure 4.4: Meat freshness levels

This data was chosen to facilitate the creation and testing of the Swin Transformer and Explainable AI (XAI) models since it has distinct and gradually more difficult classes. Specifically, the visual changes that are subtle in the Half-Fresh category give rise to a complicated classification scenario and force the model to identify localized spoilage cues as well as global changes in texture. Using this validated and well-known dataset, the research will guarantee that the performance levels and interpretability results of the suggested "Glass Box" methodology are based on an existing criterion to measure meat freshness.

4.3.4 Data Cleaning

For the sugar dataset, the data cleaning process mainly focused on organizing all images into correctly labeled folders corresponding to their levels of adulteration: pure, 10%, 20%, 30%, and 40%. Images that were corrupted or misread or very blurred were discarded in the process so that consistency can be maintained with correct dataset. Only the correct image file types were saved and finally each class had 640 images at last, as some corrupted images were removed. Any other type of remaining differences between the pictures were managed by using data augmentation during pre processing. Also, any remaining variations among images, that could be such as differences in lighting, orientation, or texture were addressed during preprocessing and data augmentation stage rather than through the manual cleaning. The result of this process is a clean, uniform, and organized data set that can be further trained using the model.

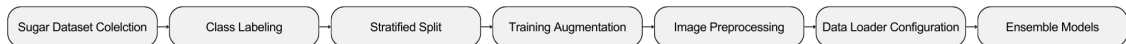


Figure 4.5: Flow Diagram of Sugar Dataset preprocessing

In the case of the poppy seed, the images were organized into nine separate class folders that had varied degrees of adulteration and types of backgrounds. A custom loader was implemented to check file integrity, automatically copy images to RGB, and replace unreadable files with blank tensors in order to prevent training disruptions. The invalid formats and non-image files were filtered, leaving only clean and usable data. Augmentation (including flips, rotation, and color jitter) was used to balance differences in lighting, background, or seed positioning to maintain consistent model input and avoid overfitting. This methodical cleaning strategy ensured that the poppy seed data was suitable to train a model because it was noise-free, corruption-free, and did not contain any inconsistencies.



Figure 4.6: Flow Diagram of Poppyseed Dataset preprocessing

In the case of the meat dataset, it is important that the data was cleaned in order to make the Explainable AI (Grad-CAM) outputs reliable. The raw dataset of 2,369 images was subjected to a stringent integrity check and an attempt was made to eliminate any metadata-heavy or corrupted files, which would slow down the hierarchical processing of the Swin Transformer. The dataset had a class imbalance, as the sample size of Fresh (889 images) was more than that of Spoiled (665 images), so the cleaning stage included a stratification of the data, so that the training, test sets, and validation sets included a representative number of all three freshness categories (Fresh, Half- Fresh, Spoiled). Moreover, each image was strongly scaled to a common resolution (usually to 224×224) and normalised. This standardization was required to eliminate the variances in the resolution that would be incorrectly recognized by the model as a texture feature and to ensure the final classification was not based on any other indicator of biological spoilage such as discoloration and surface texture.



Figure 4.7: Flow Diagram of Meat Dataset preprocessing

4.3.5 Data Transformation

In the sugar dataset, all the images were reduced to 224×224 pixels to fit the input dimensions of convolutional neural networks (CNNs) trained on ImageNet [36]. This standardization of images meant that all architectures would have the same spatial conditions (ResNet50, DenseNet121, ConvNeXt-Tiny) and feature extraction would be consistent, and a comparable assessment is possible. A fully connected classifier head with a batch normalization, nonlinear activation, and dropout are used to extract the features, and random resized cropping, horizontal flipping, rotation, color jittering, were used to augment the classifier during training to enhance resistance

to spatial and illumination variation. Mixup regularization ($\alpha = 0.2$) which basically was also again applied in order to blur decision boundaries and decrease overfitting. These types of additions or augmentations used in the models brought variability in lighting, rotation, and also to the texture image, which enhanced the robustness of the model towards unseen types of images or data, which could prevent memorizing of data and could help the models perform better without partiality. Invariance to orientation changes was further augmented with geometric augmentations (random horizontal and vertical flips and rotations) over the range of $\pm 15^\circ$. A manually designed augmentation pipeline which use torchvision transforms is applied with a view to enhancing robustness against spatial and illumination variations. Color jittering was used to deal with the differences in illumination. ImageNet statistics were then applied in order to normalize the images so as to be compatible with pretrained models. Lastly, the erasing was done randomly to imitate occlusion to minimize overfitting and enhance overall robustness.

Validation and testing subsets were only resized, center-cropped, normalized, and converted to tensors to ensure unbiased evaluation conditions. All changes were made using the PyTorch torchvision library to guarantee consistency in preprocessing and the reproducibility of all model configurations. In the case of the poppy seed dataset, a similar kind of preprocessing and augmentation pipeline was applied, such that the methodology would be uniform across the study and with a view to developing its performance. Initially, the dataset consisted of 569 images of pure and adulterated images of poppy seeds. As this amount of data is not adequate to train even a pretrained model, a data augmentation method was used to increase the number of images. The final number of pure and adulterated images during training was 1,138. Every image was resized to the size of 224x224 pixels and stretched using ImageNet statistics (mean = [0.485, 0.456, 0.406], std = [0.229, 0.224, 0.225]) to be compatible with the EfficientNet software that uses 224x224 as the default image size. The training augmentations that were implemented across the model consisted of random horizontal flips (probability, $p = 0.5$), random vertical flips (probability, $p = 0.3$), rotations ($\pm 30^\circ$), affine transformations (translate = 0.1, scale = 0.9–1.1), color jitter (brightness = 0.3, contrast = 0.3, saturation = 0.3, hue = 0.1), random perspective distortion (distortion = 0.2, $p = 0.3$), and random erasing ($p = 0.3$, scale = 0.02–0.15). Also, CutMix augmentation ($p = 0.5$) was used in training to strengthen generalisation. Connected to this, a class-weighted loss function combined with a WeightedRandomSampler was used to reduce the problem of data imbalance. Optimisation in the form of AdamW (learning rate = 0.001, weight decay = 0.01) was used along with scheduling on ReduceLROnPlateau, and dropout ($p = 0.3$) was added to the classifier to regularize. A 10-fold stratified cross-validation plan was utilized, and the plan was then evaluated on 20% heldout test set as it was required. These types of identical standardized changes, along with the optimization framework strongly encouraged- evacuated feature extraction, with reduced bias, and also assisted in efficient evaluation of transfer-learning effectivity of the developed model. In order to be compatible with the Swin Transformer architecture and to train the models in a stable and efficient fashion, a well-crafted data transformation pipeline was used to process the meat freshness data. There were three significant parts of this pipeline; data preprocessing, data augmentation and performance optimization. The transformations were made in a systematic way to ensure consistency, avoid leakage of data and ensure maximum use of computational

efficiency. Preprocessing of data was used throughout training, validation, and test data. All raw images irrespective of their original resolution or aspect ratio were down-sized to a set size of 224 x 224 pixels. This downsizing is mandatory to the architecture of Swin Transformer. After resizing, the images were all converted to a floating-point format (float32) and the pixel intensity values were scaled by converting them into the $[0, 255]$ range to the $[0, 1]$ range. Moreover, ImageNet means and standard deviation values were used to further normalize the normalized images, to align the input distribution with the pretrained Swin Transformer weights. Data augmentation was only done on the training data to enhance robustness and generalization. Orientation invariance was then provided by random horizontal and vertical flips so that classification decisions would not be based on image alignment. The maximum rotation of 0.2 radians was used to simulate natural variations in viewpoint, and the zoom is used randomly up to a scale of 0.2 to allow the model to learn scale-invariant features. The validation and test data sets were not purposely augmented to maintain the unbiased assessment setting. Caching and prefetching strategies were applied on all datasets in order to optimize performance. Caching allowed computationally intensive preprocessing and augmentation processes to only be computed once, with the benefit of much lowering the latency of loading the data in subsequent epochs. Prefetching enabled the data pipeline to issue the next batches asynchronously when the model was processing the current batch and thus optimized the use of the available GPUs and enhanced the overall training throughput. Together, this end-to-end data transformation pipeline formed a solid and highly efficient base of training and testing the meat freshness classification system based on Swin Transformer.

4.3.6 Data Integration

In this study, data integration was inclined at achieving consistency in the way the data should be arranged, how the name or the label should be given, and how to organize the sugar, poppy seed, and meat data to enable the equal and consistent deep learning experimentation with the selected different models. The data assimilation will ensure consistency and compatibility with the preprocessing, training and testing, which would happen later. This was done in the sugar data where the sample image was put together by hand, in 5 folders (adulteration levels: 0, 10, 20, 30, 40 %) as a part of data integration, so the images were in a folder of 640 images. This was an equal sample of 3,200 images. Upon augmentation, tensors were produced, and here then the images were normalized by ImageNet mean and standard deviation values, which guarantees the numerical stability and compatibility within the trained backbones. To validate and test, deterministic preprocessing (resize, center crop, tensor conversion and normalization) was used, and this means that evaluation was fair and consistent. The data was divided into 70 % training (2240 images), 15 % validation (480 images) and 15 % testing (480 images) by PyTorch functionalities with a fixed seed to ensure reproducibility.

The data of poppy seed which were pure, semolina-adulterated were gathered under controlled illumination with different backgrounds (black and white surfaces), into the Mendeley Data Repository. The initial data organization involved separation of the images into various folders. In this dataset, a stratified 10-fold cross-validation plan was incorporated into the pipeline and 20 epochs were assigned to every fold

after the balancing and cleaning processes. Also, 20 % of the total set had been set aside as a held-out test set to be finally tested with models. In the case of meat data, integration was done by placing the 2,369 images in the Meat Freshness Image Dataset into a single directory system that could be read by the Swin Transformer data loader. The raw images which were initially grouped into three folders (Fresh, Half-Fresh, Spoiled) were indexed and assigned numbers (0: Fresh, 1: Half-Fresh, 2: Spoiled) to enable the application of the tensors. To manage the intrinsic class imbalance (Fresh (889), Half-Fresh (815), and Spoiled (665)), a stratified split approach was incorporated such that the training, validation, and testing groups have a balanced representation of all the three stages of freshness. This regular integration procedure guaranteed that all three datasets sugar, poppy seeds, and meat were homogenized, balanced, and appropriately divided to offer an efficient basis to the deep learning adulteration detection activities.

4.3.7 Data Reduction

There were no specific dimensionality reduction methods like principal component analysis used directly before training a model in the case of sugar because convolutional neural networks automatically have to learn hierarchical feature representations. Rather, implicit regularization techniques have been used to regulate model complexity and overfitting. These were dropout and other preprocessing during training which gives no incentives to use a single feature but contributes to more successful generalization. Also, label smoothing basically was also again used to minimize overconfidence of the predictions, and a learning schedule was used that would include a warm-up period and then the cosine annealing to stabilize optimization was applied to promote smooth convergence. In the case of poppy seed, Dropout (rate=0.3), random erasing ($p=0.3$), and CutMix augmentation with $\alpha=1.0$ (used with probability= 0.5) served as implicit regularization schemes. Class imbalance was dealt with using weighted random sampling and using weighted loss by classes. Redundant parameter updates were reduced by adaptive learning rate scheduling (ReduceLROnPlateau, factor=0.5, patience=3) and by L2 regularization (weight decay=0.01). In the case of the meat freshness data, no explicit dimensionality reduction methods like Principal Component Analysis (PCA) were used. Rather, data reduction was naturally realized in the architectural design of Swin Transformer. The model does an effective dimensionality reduction by way of patch based representations and hierarchical feature learning. First, the input images are all resized to a standard image size of 224 x 224 pixels, which is also a form of data volume reduction since it eliminates redundant high-resolution pixel data, but maintains important spatial characteristics that can be used in the measurement of freshness. After the Swin Transformer, the input image is broken into non-overlapping 4 x 4 sized patches and the dense pixel image is encoded as a sequence of small patch embeddings. Patch Merging layers progressively decrease the spatial resolution and increase the dimensionality of the features by factors of 4, 8, and 16, decreasing the count of tokens by factors of 4, 8, and 16. This hierarchical reduction plays a similar role to pooling convolutional networks but has more semantic information. The shifted window self-attention is also used, limiting the interactions among features to local areas, which decrease the computational cost, or complexity, and maintains the spatial relationships of meaning.

4.3.8 Summary of Preprocessed Data

In the case of sugar data, the ultimate sample was 3,200 pictures that were spread into five classes denoting the various levels of adulteration (pure, 10 %, 20 %, 30 %, and 40 %). Each class has exactly 640 images. A stratified division was used to select 70%, 15%, and 15% parts of this sample as training, validation, and testing uniformly with regard to all classes.

In the poppy seed data, there were 1,138 pictures that were utilized, and these were classified into 9 classes total with adulteration levels (10%, 20%, 30%, 40%, 50%) plus background variations. The classes had between 80 and 166 images. Classes, which were to be used, were given some weights in order to have a total balance in the dataset. The 10 fold stratified split was used, and each of the nine classes was uniformly represented across the sequential training, validation, and testing phases as well. The meat data has 2,369 images obtained by the Meat Freshness Image Dataset and divided into three different classes, Fresh (889), Half-Fresh (815), and Spoiled (665). All the raw images were re-sized to a standard image size of 224×224 pixels and transformed into normalized PyTorch tensors with ImageNet statistical values in order to be compatible with the Swin Transformer architecture. The data were stratified to overcome the imbalance between the classes and separated into training, validation, and testing subsets that were processed in 32 batch groups optimally using the GPUs.

Table 4.1: Datasets used in the experiments.

Dataset	Source	Classes	Total Images	Image Size	Framework	Notes
Sugar	Self-made (controlled setup)	5 (0%, 10%, 20%, 30%, 40%)	3,200	224×224	PyTorch	Manually collected, balanced per class
Poppy Seed	Mendeley Data Repository	9	1,138	224×224	PyTorch	Used EfficientNetB0 with data augmentation to increase some images
Meat	Meat Freshness Image Dataset	3 (Fresh, Half-Fresh, Spoiled)	2,266	224×224	PyTorch	Stratified split; normalized using ImageNet statistics; batch size 32

4.4 Implementation of Selected Designs

4.4.1 Sugar

As first part of the work flow, the sugar dataset was created manually, mixed with varying %ages of adulterants of $MgSO_4$. Then models were evaluated one by one to get the models with best performance, and finally an ensembled model structure with three pretrained model was formed in order to get a more reliable, stable and accurate framework for detecting sugar adulterants. The goal of the implementation will be to automatically distinguish between a pure and adulterated sugar sample and, in the latter case, to determine the sample in one of four predefined adulteration concentration levels namely 10, 20, 30, and 40%. If outside these classes, it detects as unknown adulterant due to Mahalanobis distance. [37] is used to measure the

distance between a test sample and learned data distribution by taking into account both variance of individual features and correlation between those features.

$$D_M(\vec{x}) = \sqrt{(\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu})} \quad (4.1)$$

To begin with, deep features are obtained out of the trained CNN. Mean vector and covariance of these features are calculated using the training data. It is computed in terms of this distribution. Then a predetermined level is used- samples that are beyond the level are considered as anomalies, which suggest the possibility of unknown adulteration. It is basically a scale-invariant metric that would work by measuring distance between a sample and a feature distribution. The entire pipeline used basically involves dataset preprocessing, deep feature extraction, ensemble learning, uncertainty estimation, and also statistical distance based verification in order to guarantee both accuracy and also robustness. The system starts by making all stochastic components highly reproducible by deterministically seeding all stochastic components, such as Python random module, NumPy, and PyTorch operations [38]. This makes sure that the results of the training and evaluation will be consistent though out more than one run, which is necessary and essential to experimental reliability.

As part of dataset construction, dataset is constructed recursively searching a root folder containing sugar images, labeling being automatically determined by the name of the directory or file containing the words, such as pure, or a numerical prefix. In case the two sources do not give valid label, image is dropped in order to prevent noisy supervision. To maximize generalization, there is use of aggressive image augmentation strategy in the training pipeline. Random spatial distortions like cropping, flipping and rotation, etc as discussed in previous data preprocessing section and color jittering and grayscale conversion is basically good for explaining light and imaging differences. It contains random erasing that makes it more robust as well by promoting the reliance of the model on distributed patterns of textures rather than the local feature [39]. The transformations of validation and testing, on the other hand, are standardized and deterministic to provide performance evaluation without bias.

ResNet-50, ConvNeXt-Tiny, and DenseNet-121 are deep convolutional backbones used to extract features and the architectures are basically set to ImageNet-pretrained weights and stripped of their original classification layers, which enables them to act as a simple feature extractor [29], [30], [40]. The features obtained are run through a batch-normalized, nonlinear activation, and dropout fully connected classifier head. Regularization is basically achieved with through the working with dropout and data augmentation rather than explicit stochastic depth . The last classifier head will be a series of dropout fully connected and batch normalized layers, eventually which will guarantee stable optimization and less overfitting.

AdamW optimizer with weight decay is used in order to train the models [41]. The backbone layers which are pretrained have a lower learning rate and the classification layers which are basically newly initialized have a higher learning rate. The learning rate schedule uses a warm-up period and thus then cosine annealing, which then stabilizes initial training and also avoids the premature convergence [42]. During training, the mixup augmentation is a random combination of two image pairs and their labels that helps in producing more smooth decision boundaries [43].

In order to stabilize the training, In carrying out the validation, it is performed using

the best checkpoint selected based on validation performance. The early stopping is applied based on two criteria based on validation accuracy and validation loss, such that the training only stops when meaningful improvements do not lead to improved results. The class weights were optionally also again used to handle class imbalance, though the final dataset was then balanced with equal images per class.

Stochastic forward passes in the inference/testing stage are only carried out by the MC Dropout function that can estimate uncertainty. Stochastic Depth is stochastic only during training, not during testing [44].

The system does not use one classifier but rather it uses an ensemble learning strategy. Several backbone models that are trained separately generate the class logits on an input image. When making inferences, dropouts are active, and therefore, Monte Carlo sampling and Bayesian uncertainty estimation can be approximated [45]. Also, test-time augmentation is used by creating several spatial crops of the input image. In all models, crops, and Monte Carlo samples, results are averaged at the logit stage, which constitutes a soft-voting ensemble implicitly. This would enhance strength and decrease susceptibility to personal model errors.

The output of the ensemble is a probability distribution on the given adulteration classes. The highest probability is the predicted class and this ability is reflected in the confidence score. Notably, the ensemble will not conduct regression or interpolation, it is purely a closed-set classifier. Thus, in case an image with an adulteration level that is not one of the trained classes (such as 15% or 50) is given during inference, the system will still give the image to the closest known class. When this happens, the score of confidence is usually low indicating uncertainty and interpretation of prediction should be taken with caution. In order to overcome or address this constraint, the system has another statistical verification mechanism on Mahalanobis distance [46]. Here, covariance as required is estimated using empirical covariance estimation. In the process of inference, Mahalanobis distance between an embedding of the test sample and the pure sugar distribution is calculated. When the distance is large, the sample is not pure sugar despite the predictions of the classifier being class 0. The mechanism gives a second signal concerning the presence of an anomalous and out-of-distribution sample and helps to increase the reliability of the system to a more realistic situation. Grad-CAM visualization is applied in order to identify regions within an image that have the strongest effect on the prediction of the model, to make them interpretable [47]. This allows qualitative examination of whether the network is concerned with semantically significant structural or textural structure in the sugar images. In case Grad-CAM does not work because of architectural incompatibility or the absence of gradients, then the system gracefully recovers the error without halting prediction. Since the issue has to do with multi-classification of sugar adulteration, categorical cross-entropy loss was adopted [48]. This was thus then done in PyTorch through `nn.CrossEntropyLoss` which here basically combines `LogSoftmax` with negative log-likelihood loss to maintain numerical stability. This loss is also useful in punishing wrong predictions and such a loss is ideal in single-label classification based on CNN. A number of corner cases are clearly addressed during the implementation. In case no valid images can be located in the directory with the dataset, the training is terminated with an evident error report. In case there are no pure sugar pictures, Mahalanobis statistics cannot be calculated, and the system informs about the lack of them. Bad or invalid images boost the controlled runtime errors instead of caus-

ing silent failures. The lack of checkpoints, the asymmetry of classes, or arbitrary mapping of the classes is relieved through explicit validation and maintained meta-data. Moreover, data that is excessively noisy or has never been observed before is likely to yield low confidence predictions and large Mahalanobis distances with implicit warnings that this data should not be trusted to be classified.

When the Mahalanobis distance is large, it means that the sample is not the pure sugar distribution, which is an indication of adulteration. Mahalanobis distance is a distribution that supposes that the underlying characteristics of pure samples are distributed as a multivariate Gaussian distribution with the covariance that is estimated. The Mahalanobis distance may either be understated or overstated in case the feature distribution underlying it is very non-Gaussian, or is multimodal, or contains outliers. Also, the extractor of features itself must be robust that is they must be able to capture the relevant texture patterns because of excessive lighting conditions, image noise, or low-resolution input or the resulting distance can be inappropriate. In these situations, Mahalanobis can give false negatives (not identifying adulterated sugar) or false positives (identifying pure sugar as anomalous) results. It is hence best treated as a relative measure of deviation as opposed to an absolute measure of decision [46].

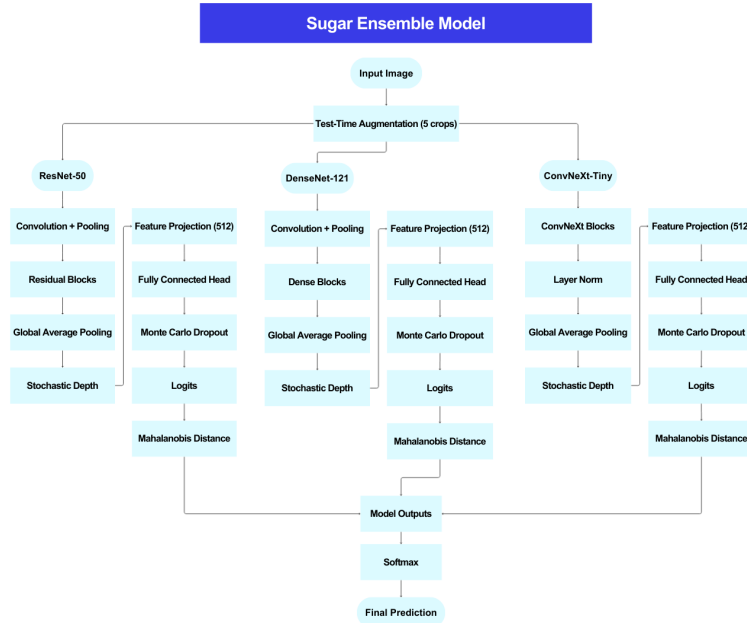


Figure 4.8: Model Architecture of Sugar Ensemble Model

In general, the implemented structure is a robust, collective-based system of sugar adulteration detection, which incorporates deep convolutional learning, probabilistic inference, statistical verification, and visual explainability. Although the classifier has a predetermined set of adulteration levels, its auxiliary uncertainty and distance-based capabilities enable it to indicate abnormal samples, which is why it can be applied in the real world of food quality inspection and academic research.

4.4.2 Poppy Seed

As for the poppy seed dataset, EfficientNetB0 was the primary and final selected model, which was also optimized for use. The model was primarily considered because of its characteristics of a balanced trade-off between computational efficiency and also it is considered to have a better high texture recognition capability. The dataset went through data pre-processing with normalization by ImageNet mean and standard deviation values, resizing all images to 224x224. Images were augmented through mainly using random horizontal and vertical flips, CutMix, color jitter, rotations up to 30 degree, affine transformation and random erasing which was applied to the dataset to create variations and enhance diversity.

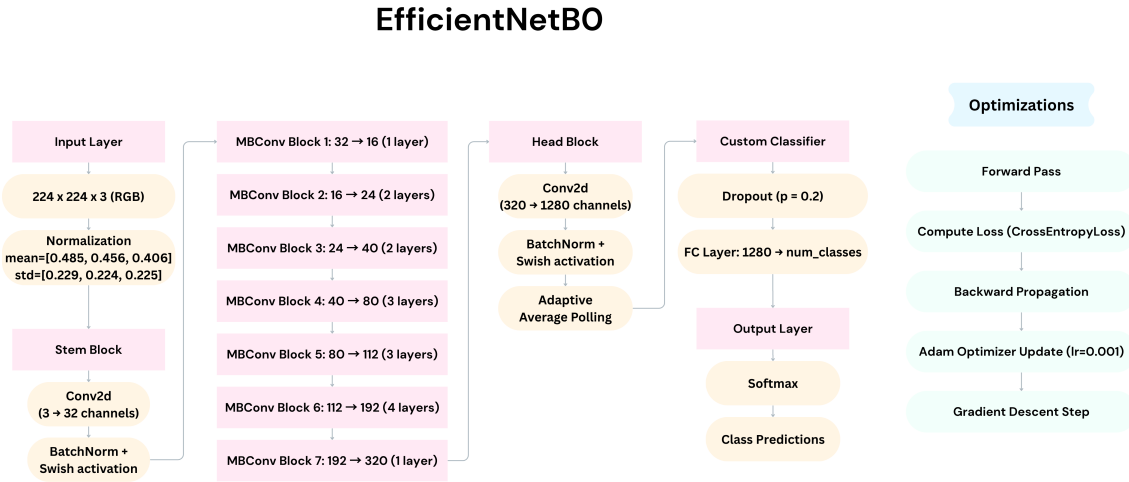


Figure 4.9: Model Architecture of EfficientNetB0 of Poppy Seed

The AdamW optimizer was used here also had similar use ($lr = 0.001$, weight decay = 0.01) and this was scheduled with the help of ReduceLROnPlateau. In order to counteract the class imbalance, the division into training, testing, and validation sets was also done here. Training was conducted under a fixed number of epochs and from the Table 5.3, it can be observed that it also followed a 10-fold stratified cross validation plan, which finally agreed on a 80-20 train test split. Cross-entropy loss function was used here as well. It had batch size of 32 for efficient computation. All these similar types of approaches helped in efficient training of the model which gave a very high test accuracy and helped in preventing the overfitting issue by maintaining consistency between validation and test accuracy per epoch. Both of the datasets were evaluated and compared from model to model on the basis of accuracy, precision, recall, F1-score, and confusion matrix to ensure reliable and mostly accurate adulteration detection.

4.4.3 Meat Freshness

The initial testing of the meat freshness detection system was done through the PyTorch framework with the Swin Transformer architecture (Swin Tiny Patch4 Window7 224) as a main feature extractor. In contrast to other CNNs, this hierarchical vision transformer uses shifted window attention, which enables it to well capture temporary indicators of local spoilage, discoloration spots, as well as global textural

degradation. The model was pre-trained with ImageNet weights to take advantage of transfer learning so that there is a strong feature extraction even with the average size of the dataset. Overfitting was reduced by fine-tuning the network on the three-class problem (Fresh, Half-Fresh, Spoiled) using the AdamW optimizer with an initial learning rate of 10^{-4} and weight decay of 0.01. Cross-Entropy Loss function was used as the main objective measure. A ReduceLROnPlateau scheduler was also included to make sure the training is stable by reducing the learning rate dynamically with the help of monitoring the validation loss. Another more specific aspect of this implementation was the engineering of an Explainable AI (XAI) module. The system was configured to produce Grad-CAM heatmaps by registering forward hooks on the final layer of normalization on the transformer. This aspect of implementation is what makes the model look like a model instead of a black-box classifier, with all the decision boundaries visible visually. The complete pipeline was compiled with NVIDIA GPU hardware (CUDA) with a batch of 32, and it was necessary to process the tensors of high resolutions efficiently.

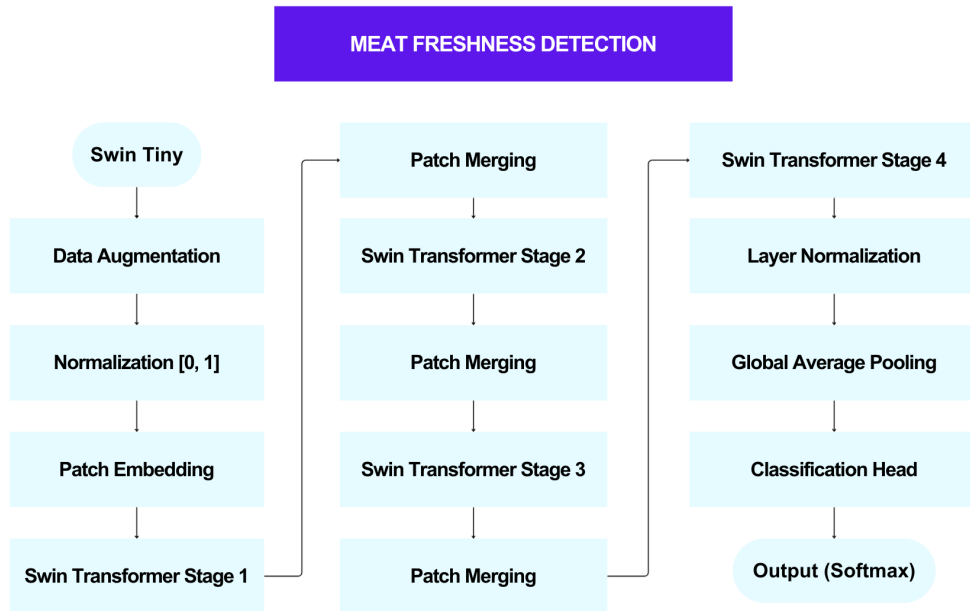


Figure 4.10: Model Architecture of Swin Transformer for Meat Classification

Chapter 5

Result Analysis

5.1 Performance Evaluation

This part explains the model of evaluation applied in determining the effectiveness, robustness, and real life usability of proposed AI-based food adulteration detection systems. The assessment was made to extend in the realm of mere accuracy measurement to include model efficiency, reliability and interpretability, which makes the models fit usefulness in a real-world application in the food safety monitoring.

5.1.1 Evaluation Criteria

The models of the sugar adulteration classification were tested with the help of a mixture of quantitative and qualitative methods to provide strong evaluation and generalization. Training and validation losses and accuracies were tracked at every epoch during training to monitor the convergence of the model and early stopping was used to stop training when the validation performances stopped improving. The then confusion matrices were produced after training in order to analyze the performance of classes, and also the class-specific misclassified adulteration levels were able to be identified. Test-time augmentation (TTA) and Monte Carlo dropout were used to enhance the stability of predictions and, in the end, the predictions were obtained by averaging the probability across several stochastic forward passes. Moreover, a combination of various CNN backbones (ResNet50, ConvNeXt-Tiny, DenseNet121) was used, and the results were combined to improve the accuracy and decrease model variance.

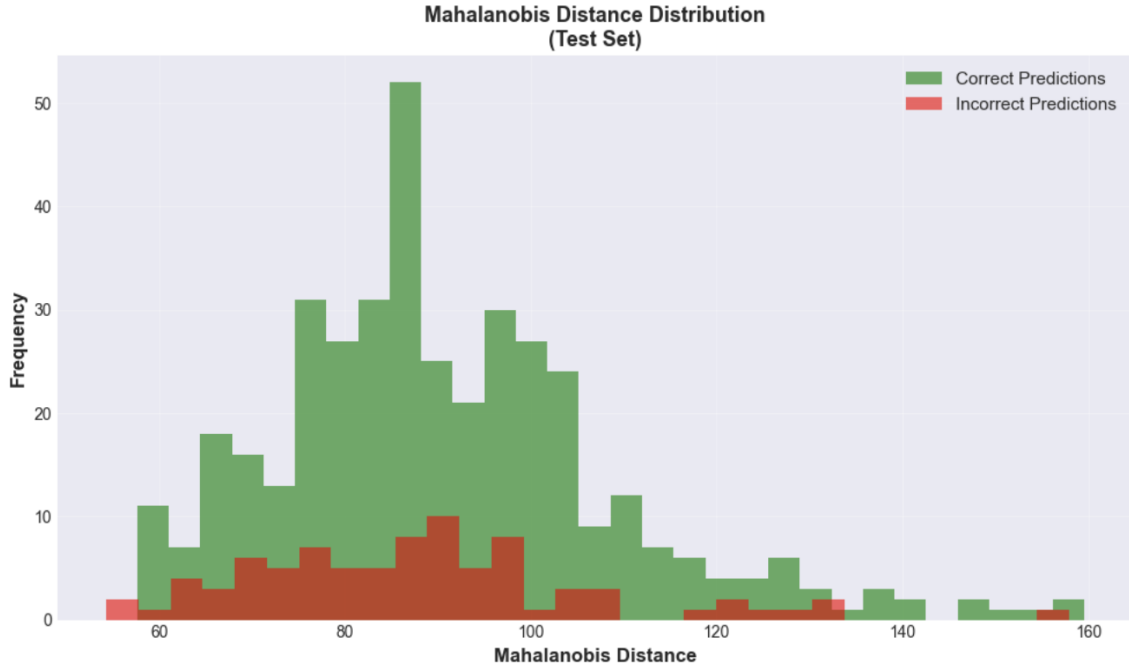


Figure 5.1: Mahalanobis Distance Performance on Test Set

The distributions of Mahalanobis distances between correct and incorrect predictions have high overlap with the distributions of correct and incorrect predictions being similar and with high peaks (80-100) and averages (90-95) and therefore, it is suited in distinguishing between the prediction correctness in samples that are in distribution where the main difference is attributed by adulteration concentration. The number of correct predictions are clearly higher than incorrect predictions. Wrong classifications between nearby levels (e.g. 10% vs. 20%) thus, still display low distances and some incorrect predictions exist. Nevertheless, Mahalanobis distance is very useful in detecting out-of-distribution, since samples that contain unknown adulterants will produce significantly greater distances ($> 200-500$), thereby being useful when used along with other uncertainty metrics like confidence, entropy and ensemble disagreement. To identify possible out-of-distribution samples, Mahalanobis distance between extracted image features and the centroid of pure sugar class was calculated giving an indication of anomaly at unseen levels of adulteration. To assess the qualitative evaluation of the results, Grad-CAM visualizations were used to demonstrate the regions of the image that have the most effect on the predictions of the model.

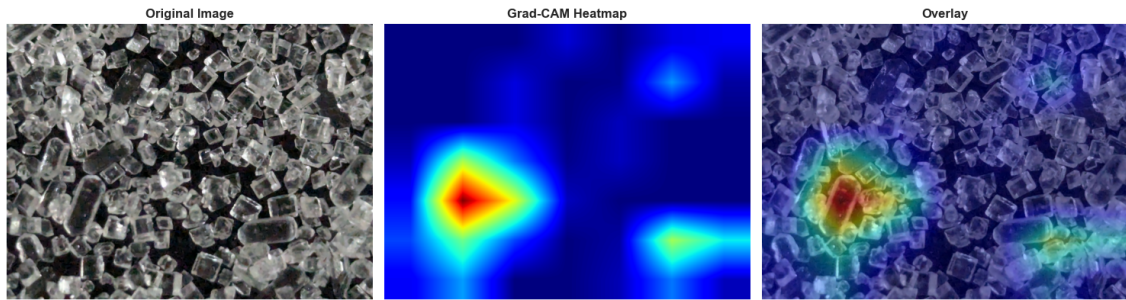


Figure 5.2: Grad-CAM Visualization of Sugar Dataset

The heatmap demonstrates that the model concentrates on crystals and surface texture patterns as the blue areas denote even pure sugar and the hotspots in red and yellow denote abnormalities and irregularities which can be used to determine the level of adulteration since the predictions are made based on discriminative features that matter and not random associations.

Soft voting ensemble is used to obtain the overall test accuracy in the `evaluate-comprehensive()` function. In each batch of tests, each of the three trained models (ResNet50, DenseNet121, and ConvNeXt-Tiny) independently produces softmax probability distributions on the five classes (0, 10, 20, 30, and 40 adulteration). The resulting probabilities are then averaged on an elementwise basis with all the models through the `torch.stack(allprobs).mean(dim=0)` which will result in a single probability distribution of each sample. The maximum averaged probability is used to arrive at the final prediction (`avgprobs.argmax(dim=1)`). To calculate the test accuracy, the prediction of the ensemble is compared with the ground truth labels using the `accuracy_score()` of scikit-learn, which counts the number of correct predictions divided by the number of the test samples in total, which gives the overall ensemble test accuracy in the results. F1 score, precision and recall scores along with test accuracy was calculated for each model and all of these were weighted since all classes have equal number of images, either weighted or macro would give same result. During training, cross-entropy loss with label smoothing was used to optimize models which could minimize the overconfidence and maximize the generalization. These techniques together gave a reliable overall scheme of testing the predictive accuracy, as well as the reliability of the models. The three models under the ensembled model's test accuracy and validation accuracy was calculated separately and then finally the overall test accuracy of the whole model was calculated. The best test accuracy was shown by Conv Next Tiny with an test accuracy of 91.04%, while the second best was Densenet121 with an accuracy of 85.83% and the third best was Resnet50 with an accuracy of 82.50%. ConvNeXt-Tiny managed to perform better on the test set reaching a precision, recall and F1-score of approximately 91 per cent, and a ROC-AUC of approximately 98.6. DenseNet121 achieved precision, recall and F1 of approximately 86 %, and ROC-AUC approximately 97.2 % and ResNet50 achieved precision, recall and F1 of approximately 82 % with ROC-AUC of approximately 96.9 %. The model that was ensembled performed well overall with a score of more than 90 on the validation set and almost 89 on the test set on the measures of accuracy, recall, precision, and F1-score. The ROC-AUC was noted to be higher than 98% which means that there is excellent discrimination between the pure and adulterated sugar samples.

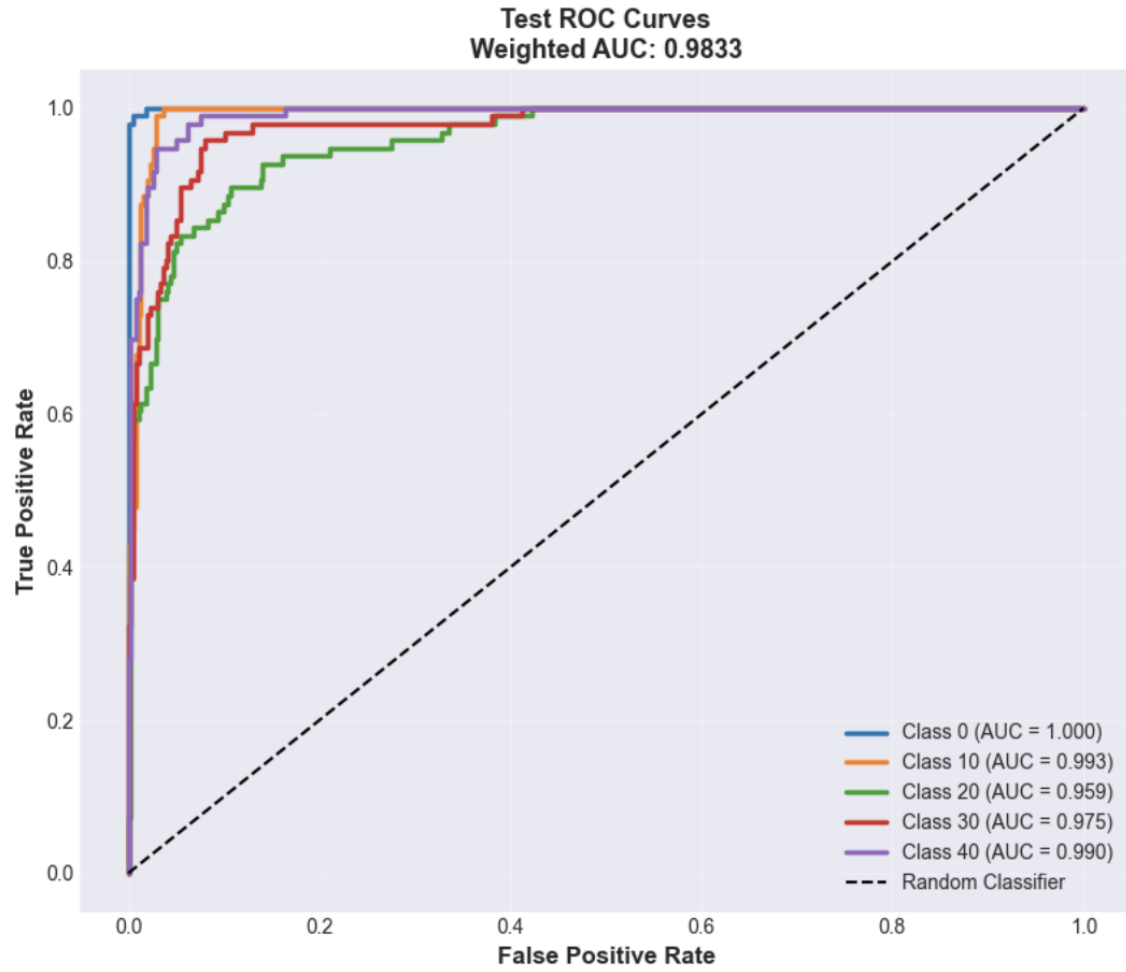


Figure 5.3: Test ROC Curve of Sugar Dataset

Table 5.1: Performance comparison of each deep learning models used within the ensemble model based on validation and test sets for sugar adulteration

Model	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
ResNet50	Test Set	82.50	82.34	82.50	82.23	96.86
DenseNet121	Test Set	85.83	86.14	85.83	85.89	97.23
ConvNeXt-Tiny	Test Set	91.04	91.09	91.04	91.05	98.64
ResNet50	Validation Set	84.17	84.05	84.17	83.76	96.90
DenseNet121	Validation Set	87.71	87.59	87.71	87.63	97.78
ConvNeXt-Tiny	Validation Set	90.62	90.66	90.62	90.59	98.85

Table 5.2: Overall performance of the ensemble model for sugar adulteration detection

Dataset	Test Accuracy (%)	Recall (%)	Precision (%)	F1-Score (%)	ROC-AUC (%)
Validation Set	90.21	90.21	90.14	90.12	98.58
Test Set	88.96	88.96	88.88	88.86	98.33

Now, to detect adulteration of poppy seeds, global accuracy was considered the main evaluation measure, which is used to determine the model to compare the mixture levels of fine grains (e.g. 10 and 20 %) and different backgrounds. Due to the visual resemblance of real seeds and semolina, class-specific measurements, Precision, Recall, and F1-score were stringently evaluated in order to make sure that the model minimized false positive, which would in effect avoid the misclassification of pure samples as adulterated. In order to provide statistical robustness and prevent the bias due to data partitioning, the assessment was based on 10-Fold Stratified Cross-Validation.

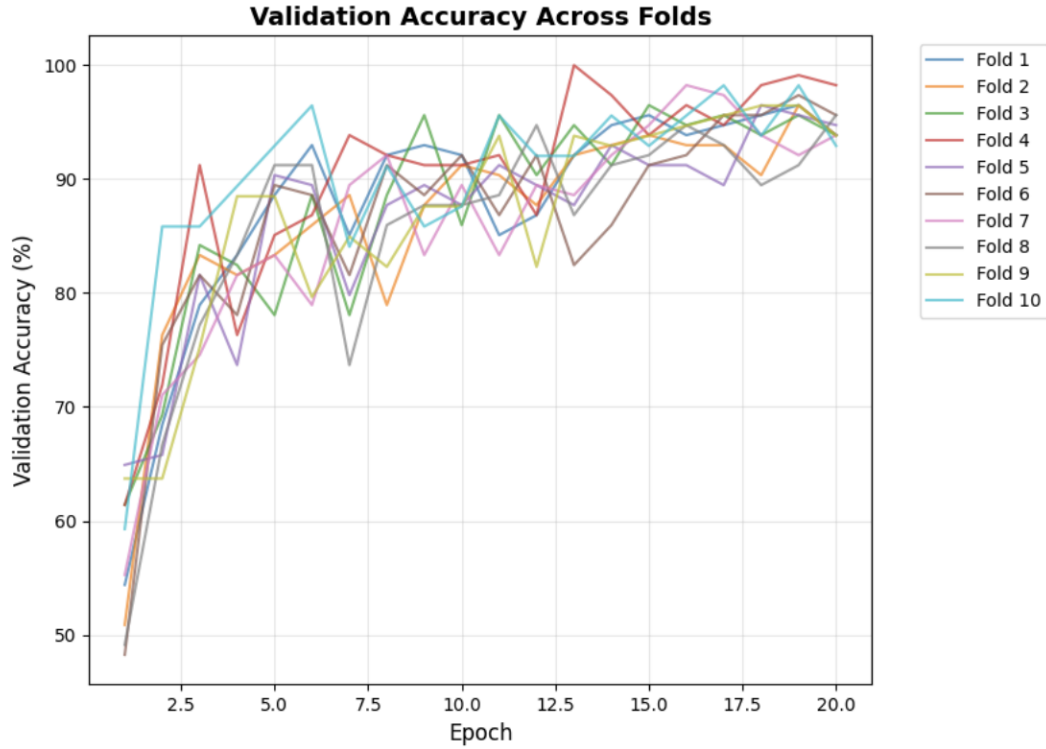


Figure 5.4: Validation Accuracy of Poppyseed for EfficientNetb0

This figure below shows the accuracy of the validation of the EfficientNetB0 model to 10-fold cross-validation of poppyseed images, with each curve relating to one of these folds and the number of training epochs (20). In general, the graph indicates a gradual growth and commonality of validation accuracy (90-98%), which represents a consistent training behavior and excellent generalization of variations in data splits. To quantify the stability of the EfficientNetB0 architecture, this method used the mean accuracy and standard deviation to ensure that the high performance was not a result of a traintest split.

Table 5.3: Performance of Poppy seed adulteration dataset

Dataset	Model	Accuracy	Precision (Macro)	Recall (Macro)	F1-Score (Macro)	F1-Score (Weighted)
Poppy seeds with different levels of adulteration	EfficientNetB0	98.25%	0.9809	0.9865	0.9832	0.9825

Where freshness classification of meat is concerned, there was a shift in the assessment system of standard performance measures towards safety and comprehensibil-

ity. Although the Accuracy, Precision, and Recall were computed to estimate the detection of the states of Half-Fresh and Spoiled, the determinant criterion was the reliability of the model with the help of Explainable AI (XAI). Grad-CAM was used to visualize attention maps in the evaluation, making sure that the Swin Transformer paid attention to biologically meaningful features - spots of surface discoloration and textual degradation, and not noise on the background.

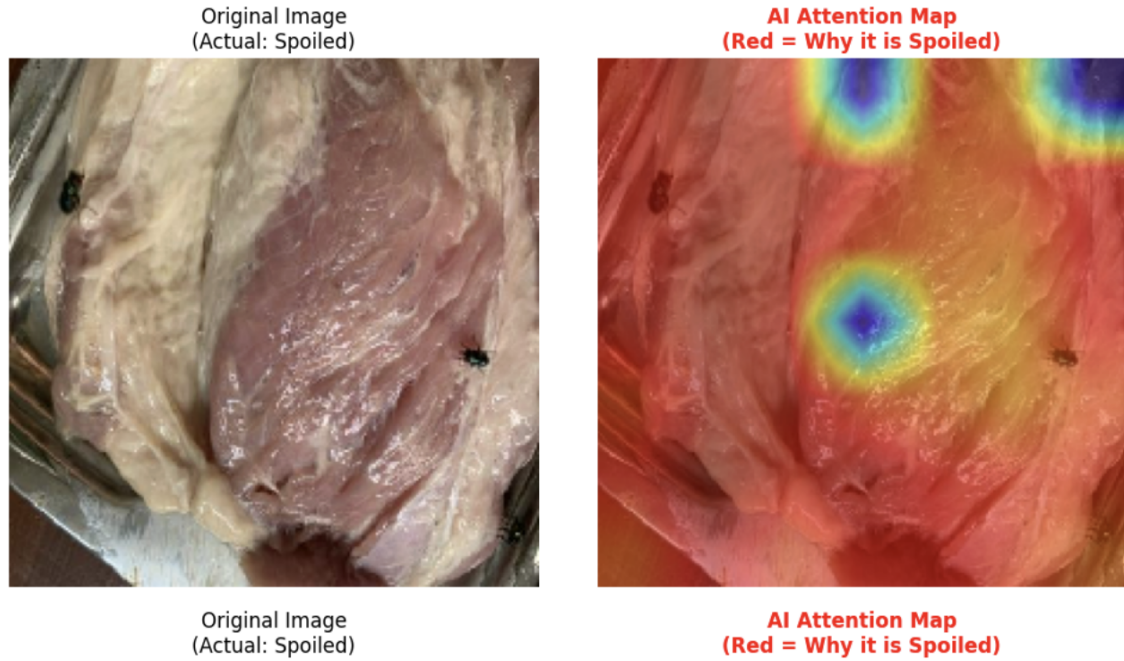


Figure 5.5: AI Attention Map using Grad-CAM

In 5.5, Grad-CAM visualization shows that the model has correctly detected the spoiled meat because it paid attention (red/yellow areas) to biofilm, slime, and discoloration of the meat surface, but not on the irrelevant background (blue areas) such as the metal tray. This, ascertains that the model relies on sound biological indicators of spoilage, namely, the glossy bacterial growth and the grey-brown tissue degradation, but not environmental artifacts, which makes it reliable as far as the actual food safety inspection is concerned. The Glass Box validation was necessary to ensure that the decisions made in the model were based on valid visual pathology, hence obtaining the confidence that is necessary when dealing with a system as important to the safety of the population.

Table 5.4: Performance of Meat Freshness Classification Dataset

Dataset	Model	Accuracy (%)	Precision (Macro)	Recall (Macro)	F1-Score (Macro)	F1-Score (Weighted)
Meat freshness with different levels of spoilage	Swin Tiny	98.88	0.9809	0.9865	0.9832	0.9825

5.1.2 Testing Methods

To detect sugar adulteration, a comparative analysis of ResNet50, ConvNeXt-Tiny, and DenseNet121 was done with a standardized 70-15-15 data split. ConvNext Tiny with test accuracy of 91.04% and a validation accuracy of 90.62% will be

the most effective and attention has been drawn to it. In ensemble experiments, the result of several CNN backbones (ResNet50, ConvNeXt-Tiny, DenseNet121) were averaged to obtain the end-result class probabilities which is the soft voting system. Accuracy was used to measure model performance and confusion matrices to visualize model performance. Moreover, a Mahalanobis distance between the features of the image and the pure sugar category was also computed to identify possible out-of-distribution samples, and Grad-CAM was also used optionally to understand model attention. To test mahalanobis distance about 30 to 35 images total mixed with urea and saccharides was clicked and they were tested one by one.

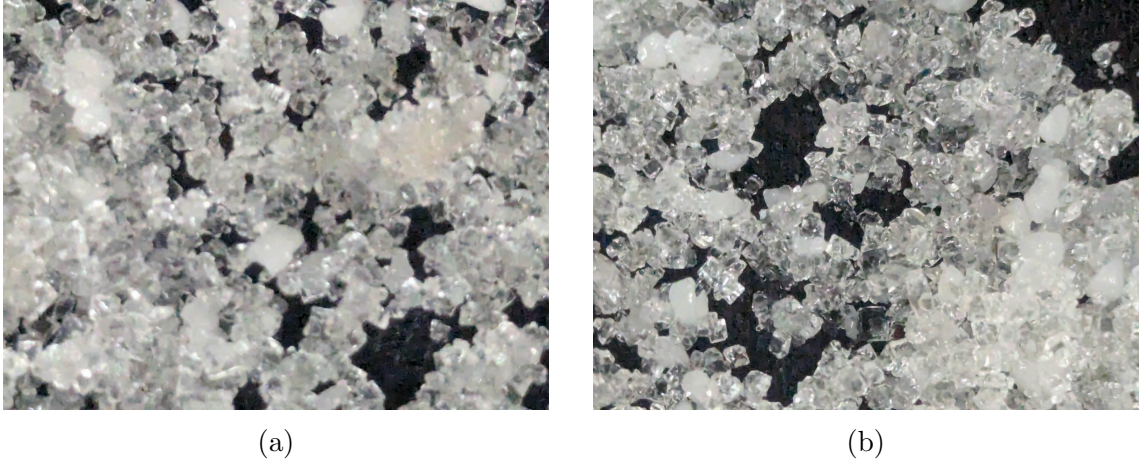


Figure 5.6: sugar mixed with urea samples

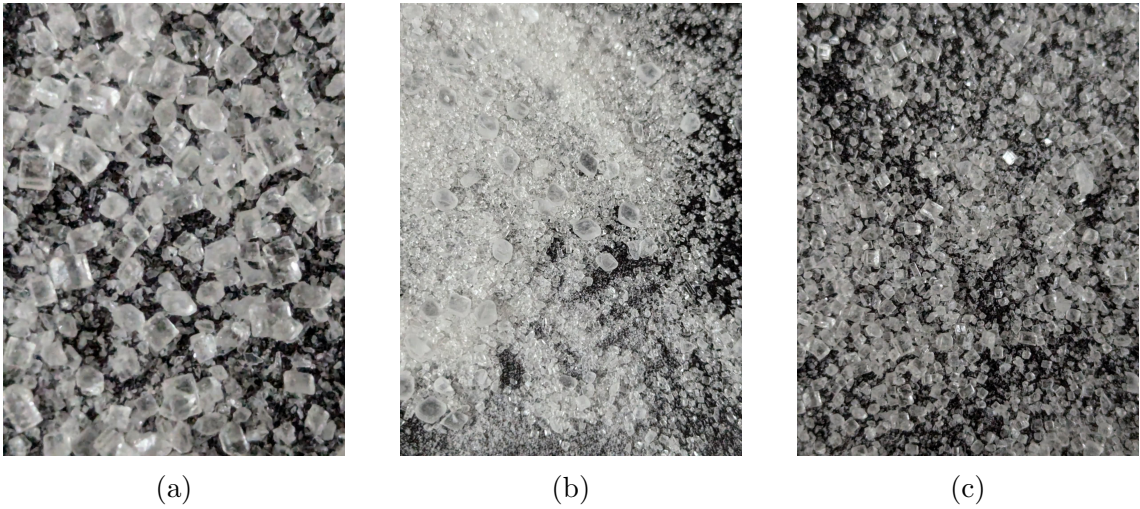


Figure 5.7: sugar mixed with saccharine samples

While some were being able to detect unknown adulterants correctly, some failed mostly urea, because of having physical granular and shape similarities with $MgSO_4$, the model presumed it to be $MgSO_4$ and predicted nearest class, this projects one of the key limitations of mahalanobis here. F1 score, recall, precision was also calculated and showcased in this paper.

Unknown Adulterant Detection

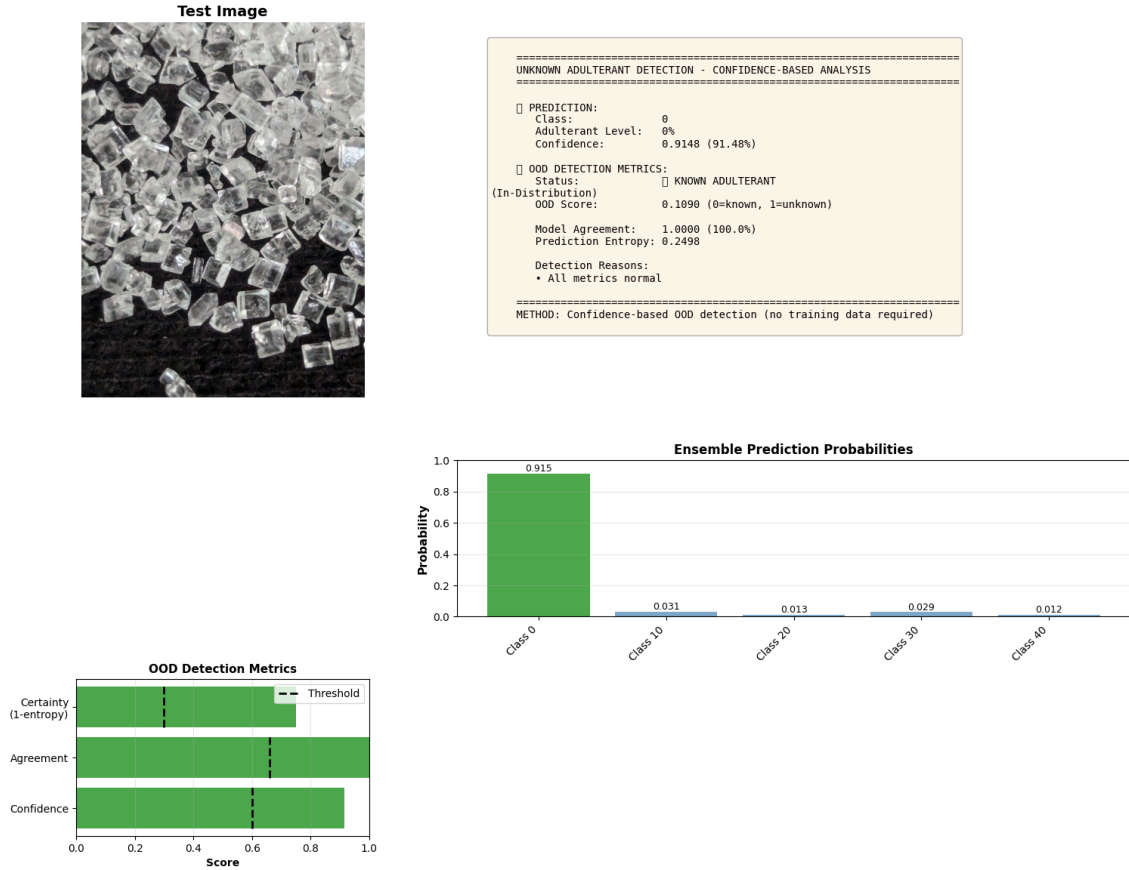


Figure 5.8: Identifying pure Sugar by Mahalanobis Distance

The ensemble model has managed to reveal the ability of OOD detection in both the test cases, indicating unique adulterations patterns. As shown in Fig. 5.8, the system identified Class 0 (pure sugar) with the confidence of 91.48% and high certainty parameters. The score of 0.1090 was less than 0.6, which is the correct number to consider this an known adulterant and thus pure sugar was correctly identified as known.

Unknown Adulterant Detection

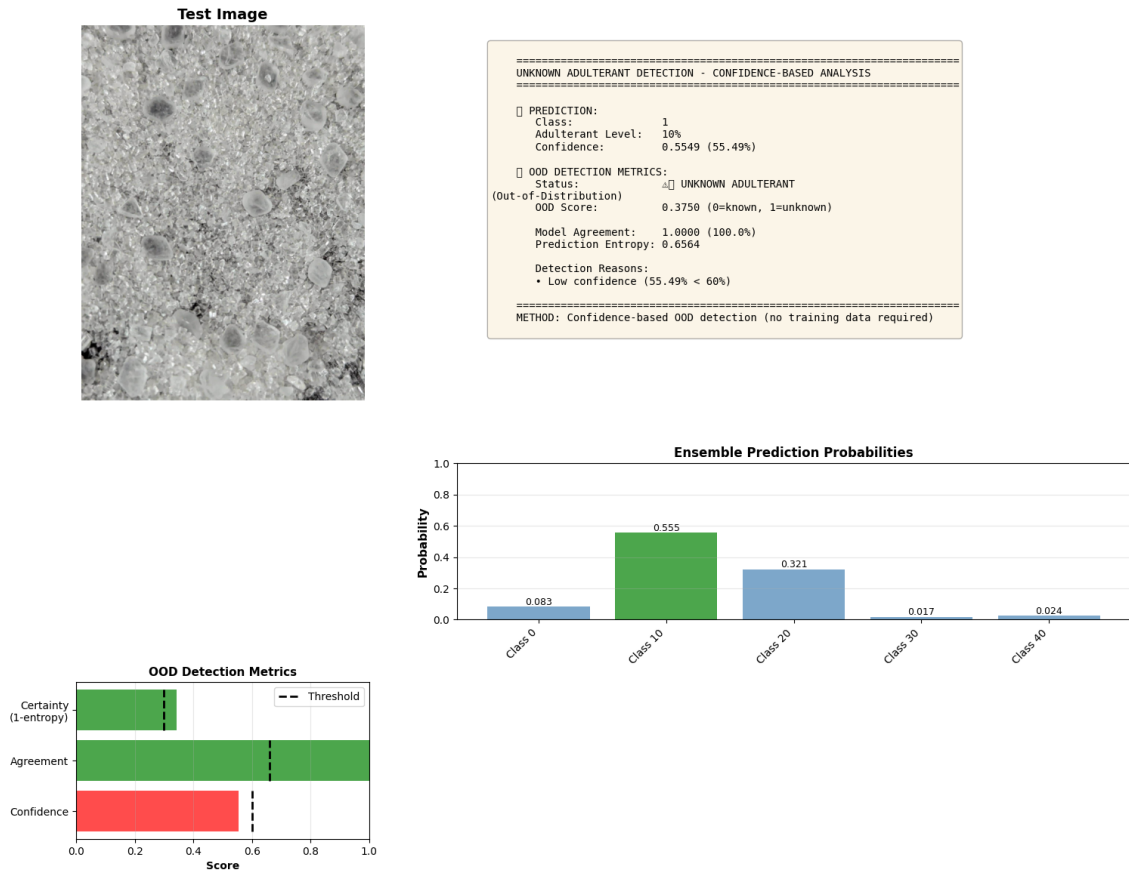


Figure 5.9: Detecting saccharine as unknown adulterant using Mahalanobis Distance

The spread of probability between the classes (Class 0: 8.3%, Class 10: 55.5%) and the high level of prediction entropy (0.6564) show that the features of the sample correspond to a previously known adulteration pattern in the training dataset, however the low confidence score of 55.49% indicates the model is not confident about the class choice even though it forcefully maps to it, further supporting the claim of an unknown adulterant being present as detected by the OOD score. It is important to note that whereas the OOD detectors make the right choice of identifying such as known, the closed-set architecture of the model requires a classification (Class 1), which has been addressed in the limitations section. In Fig. 5.9, this grayscale microscopy image of sugar crystals is predicted as Class 1 (10% adulteration level) with a confidence of 55.49% this intensive mapping is possible because of the way the model is represented. This system had indicated it as an UNKNOWN adulterant since the Out-of-Distribution (OOD) value of 0.3750 is that the patterns of features observed in the sample do most closely match a known pattern of one of the training class distributions, which was correct identification. It is worth noting that whereas the OOD detection is able to identify this as known or unknown, the closed-set architecture of the model compels a class assignment, which is a limitation, as explained in the limitations section.

Unknown Adulterant Detection

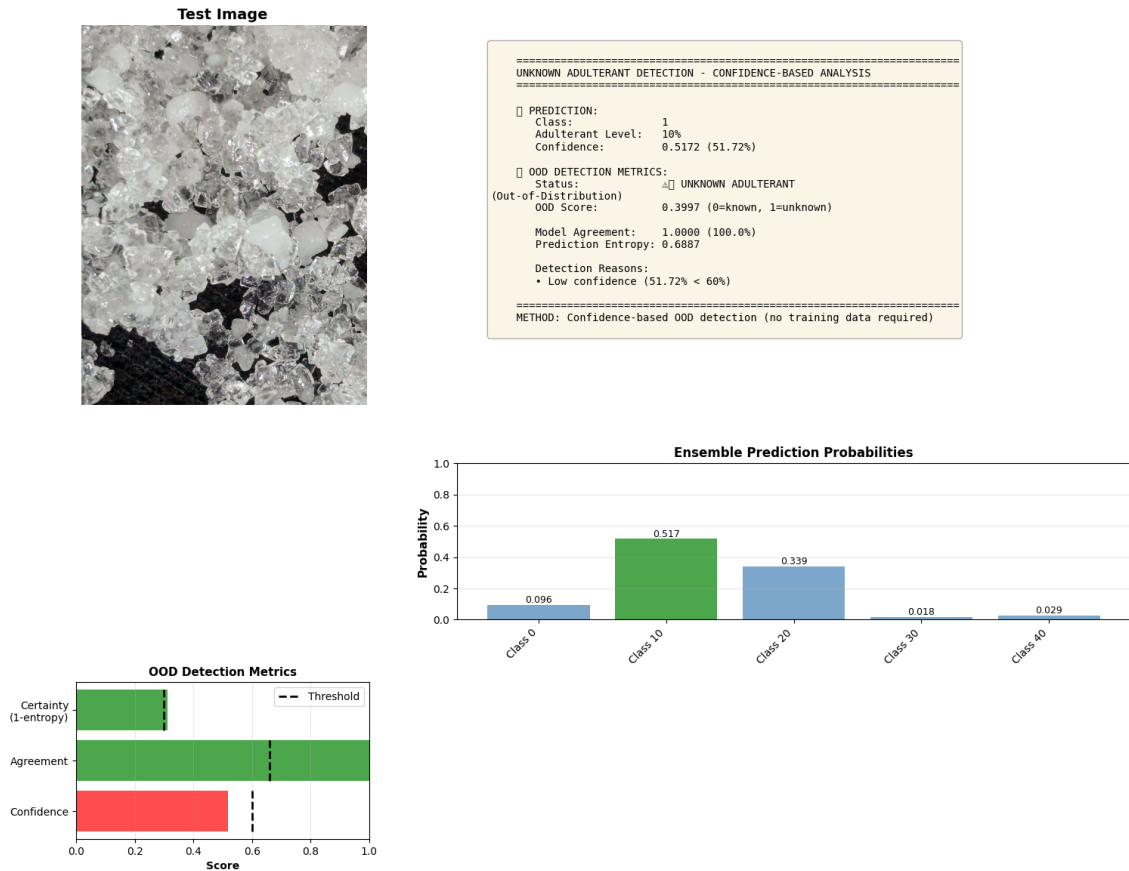


Figure 5.10: Detecting urea as unknown adulterant using Mahalanobis Distance

Equally, the uncertainty of the decision that the model makes, as shown by the spread of probability between the classes (Class 0: 9.6%, Class 10: 51.7%, Class 20: 33.9%) and the high prediction entropy (0.6887) indicates that there is a greater degree of uncertainty about the choice that the model makes, as the feature representation of the sample does not strongly correspond to a single known adulteration pattern within the training distribution. The closed-set ensemble model classifies the sample as the Class 1 (10% adulteration level), but the confidence score of the model, 51.72% is quite low, which indicates the uncertainty of the classes separation. Therefore, the confidence-based out-of-distribution detector is correct in its detection, and even though it maps to an adulteration class, its confidence is quite low and the sample identified as an UNKNOWN adulterant, as indicated by an OOD score of 0.3997, which is above the confidence threshold of in-distribution is accurate. This result indicates that high model agreement, low prediction confidence and high entropy are accurate predictors of detection of unknown adulterants. Similar to the above case, the OOD framework is effective in pointing out the sample as unknown but, due to the closed-set architecture, a forced class prediction must be made, which is also discussed in the limitations section.

These findings confirm that the scores of confidence are not detected solely. The multi-metric OOD system that is a blend of Mahalanobis distance, MC dropout uncertainty, and ensemble disagreement offers sturdy identification of samples not in the training distribution which is crucial in identifying new adulterants or unknown

adulterants than just magnesium sulfate in practical setting. So, the system is useful for detecting adulterants of sugar. It flags anything other than magnesium sulfate mixed with sugar and pure sugar as unknown using these statistical measures overall together.

In adulteration of poppy seed, Stratified K-Fold Cross-Validation was applied in order to have the statistical strength and non-biased estimation of performance. The EfficientNetB0-based model has a cross-validation average accuracy of 97.19%, which was 1.23 standard deviation, and a final accuracy of 98.25%, with average high precision in all the classes. In freshness of meat, Swin Transformer model has been tested with Grad-CAM to ensure that its predictions depended on actual spoilage signals, e.g., discoloration and texture degradation, to ensure that it is applicable in the real world in terms of inspection.

5.2 Analysis of Design Solutions

In sugar adulteration, the main aim was to determine the concentration of $MgSO_4$ adulterant in five different concentrations of 0, 10, 20, 30 and 40. This was not a task of binary detection like detecting tiny changes between images of visually similar white crystalline substances. In order to overcome this problem, three deep learning models (ResNet50, ConvNeXt-Tiny, and DenseNet121) were compared. Precision was determined using a held-out test set of images and effectiveness was determined using the stability of training and the behavior of convergence during 30 epochs. In the ensembled model, the ConvNext-Tiny recorded the highest performance at 91.04% which was higher than ResNet50 and Densenet121 (82.5 and 85.83% respectively). ConvNeXt-Tiny performed more outstanding than DenseNet121 and ResNet50 because of its contemporary framework, which contains more detailed and discriminative features with enhanced generalization concerning the minor changes in images. It has a better design such as bigger receptive fields and better normalization which enables more accurate test on the sugar dataset. ResNet50 and DenseNet121 were near to each other due to the fact that the two work on residual based architecture, basically which is effective in maintenance of information across the layers so these would be actually capable of learning similar hierarchical features. Their designs slightly differ in terms of connectivity (residual and dense blocks) but retrieve similar representations, which is why they performed equally well on the sugar dataset. For all the models, the validation accuracy graph rises slight higher than the training accuracy and it is because the training time regularization as dropout and data augmentation was on, which means that training was harder than validation, and validation data was cleaner, so model performs better. The strongest point about this solution is that it goes on to give quantitative adulteration levels that is important in determining the toxicity risks. Nevertheless, the multi-class characteristic of the task and high visual similarity of the classes make the accuracy of the classification system lower than binary classification systems. Regarding cost and viability, the method is realistic since it is based on the data of the image only, and the laboratory testing is not necessary.

In case of detecting poppy seed adulteration, the design was concerned with the vigorous multi-class classification in view of the existence of similar adulterants that are visually alike together with the authentic poppy seeds. Transfer learning strategy on EfficientNetB0 architecture was adopted to obtain the high-level

morphological features. Stratified K-Fold Cross-Validation was used to identify the reliability because it made sure that performance was consistent across various data splits, whereas precision was used to prevent the false detection of pure seeds as adulterated. The model was found to have a test accuracy of 98.25 %age and a cross-validation mean accuracy of 97.19 %age with a low variance which implies high robustness. The big advantage of such solution is that it is very accurate and reliable and can be applied in real-life within the commercial and regulatory environment immediately. Nonetheless, EfficientNetB0, while efficient, may still need optimization methods such as quantization or pruning to allow it to be used on low-end hardware, with some implications on its practicality in resource-constrained systems. In the example of detecting the freshness of meat, the design was based not only on the accuracy of classification but also on transparency and trust of the user. A Swin Transformer network was chosen due to its capability to capture long-range dependencies in images, which is significant in acquiring minor spoilage patterns. Grad-CAM was used to visualize regions that affected model decisions and explainable AI (XAI) was added. The test was on trustworthiness as the model was tested to ensure that it responded to significant spoilage indicators like discoloration and texture deterioration and not the background noise. The model was able to sort samples into fresh, half-fresh, and spoiled cases and Grad-CAM visualizations added evidence to the fact that the model made valid decisions. The primary advantage of this method is the fact that it is explainable, and anyone can visually interpret why a sample is regarded as spoiled. The use of transformer-based models, however, makes computational cost more expensive and depends on relatively modern hardware and adequate lighting conditions which would somewhat restrict feasibility in low-resource environments. In general, it can be seen that the design solutions exhibit clear tradeoffs in accuracy, efficiency, cost, and feasibility. The sugar solution is more important to fine-grained quantification, the poppy seed solution aims at robustness and high accuracy, and the meat freshness solution is more concerned with explainability and trust. Collectively, these strategies demonstrate how the design of AI models that are specific to tasks can be used successfully to balance between performance and applicability in food adulteration detection systems.

5.3 Final Design Adjustments

The final system design was modified to include a number of significant changes based on the performance analysis and experimental results in order to make the system more accurate, reliable and trustable to its users. These types of refinements have been made to overcome the limitations that have been witnessed during the testing of the baseline models with the view that the final system must be fit to serve in the real world. Initial tests with ResNet50 showed a test accuracy of 82.50 to detect sugar adulteration, although the model was not able to discriminate low levels of adulteration, with $MgSO_4$ crystals physically appearing similar to sugar. To address this limitation, the architecture was changed so that it incorporates to DenseNet121 in an ensembled version, which has dense connectivity and feature reuse mechanism enabled it to be able to capture finer textual differences. This enhanced discrimination of different adulteration levels and the test accuracy was 85.83% and training was more stable. Finally, ConvNeXt-Tiny was added since it was more successful compared to both architectures (a test accuracy of 91.04 with

higher levels of recall, precision, F1-score, and ROC-AUC).

In the initial stages of training, overfitting was found especially in the sugar dataset where the model was at risk of becoming familiar with the particular crystal patterns instead of the features that can be generalized. To solve this problem, a more sophisticated data augmentation was included in the training pipeline based on Mixup regularization. The decision boundaries were made smoother and stronger by training the model on linear combinations of images and labels. This change minimized overconfidence of ambiguous samples and enhanced the generalization response in test data that had not been observed. When it comes to the freshness of meat, it was found that a standard black-box prediction strategy was not adequately reliable in the eyes of the user. An insignificant freshness label without elaboration may cast doubt and may cause the waste of food due to its insignificance. The resulting system design incorporating an Explainable AI functionality through Grad-CAM addressed this. The system was changed to give both a classification result and a relative heatmap that visually pinpoints the areas which affect the decision. This transformation makes predictions interpretable and based on actual indicators of spoilage like discoloration and change of surface texture. The extraordinarily high initial accuracy in the poppy seed adulteration detection case was a cause of legitimate concern that the findings may have been biased by the data distribution or a good train-test split. As a way of stringently assessing statistical robustness, the validation plan was supported by substituting the simple hold-out methodology with 10-Fold Stratified Cross-Validation. Such a stringent validation process ensured the consistency and reliability of the EfficientNetB0 based model which achieved an average accuracy of 97.19% with standard deviation of 1.23% through the folds. As a result, the resulting model that is deployed is shown to be strong and resilient to changes in dataset partitioning.

5.4 Statistical Analysis

This part outlines the statistical analysis, and quantitative analysis to assess the performance of the suggested deep learning models of sugar adulteration, poppy seed adulteration and meat freshness classification. All statistical work was performed with Python, with Scikit-learn library being used to calculate performance measures and Seaborn to visualize errors patterns including confusion matrices and trend of misclassification by individual classes.

Multiple standard metrics of classification were derived to each model in order to make a rigorous and unbiased evaluation. Accuracy was taken as the main measure of the proportion of correctly classified samples in general and different model architectures. Precision was calculated to determine the %age of true positive samples and this was essential so that the false positives are kept at a minimal level and wrong labelling of pure foodstuffs as adulterated is avoided. Recall was established to estimate the ability of the model to detect the real instances of adulteration or spoilage correctly in order to minimise false negatives. The F1-score, which is the harmonic mean of precision and recall, was also claimed to give a balanced assessment particularly in multi-class problems such as levels of sugar adulteration concentration where precision and recall tend to be trade-offs.

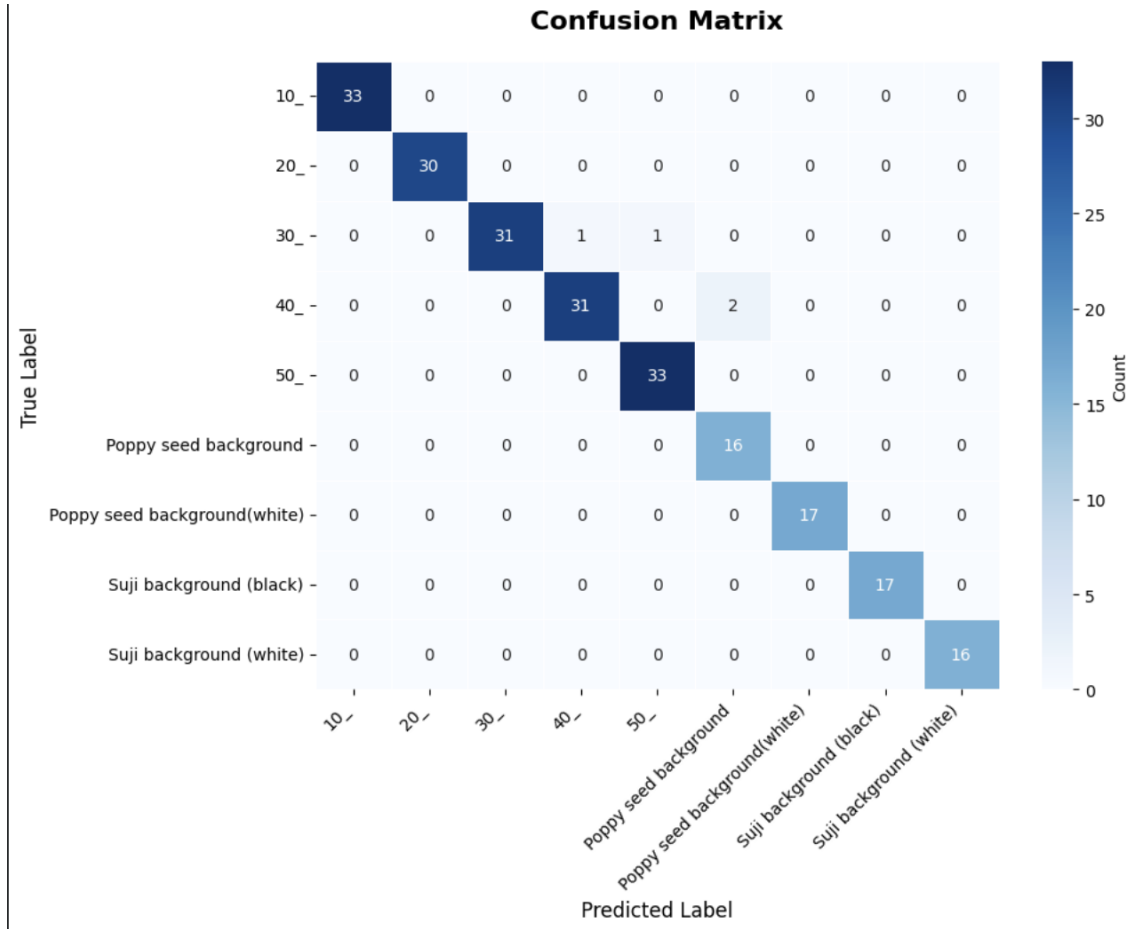


Figure 5.11: Confusion matrix of test data for Poppyseed Dataset

The application of three different statistical methods of validation was determined by the type of the data sets. In the case of poppy seed data, the Stratified K-Fold Cross-Validation has been applied to have a statistically sound outcome. The data used in this method was separated into several folds and the original distribution of classes in each fold was maintained. The model was repeatedly trained and tested and each fold was used as the test set only once. Out of these repeated experimentations, the average accuracy and the standard deviation have been calculated and the mean accuracy obtained is 97.19% and the standard deviation is 1.23. This statistical method also ensured that the high performance of the model was not due to a specific random data split.

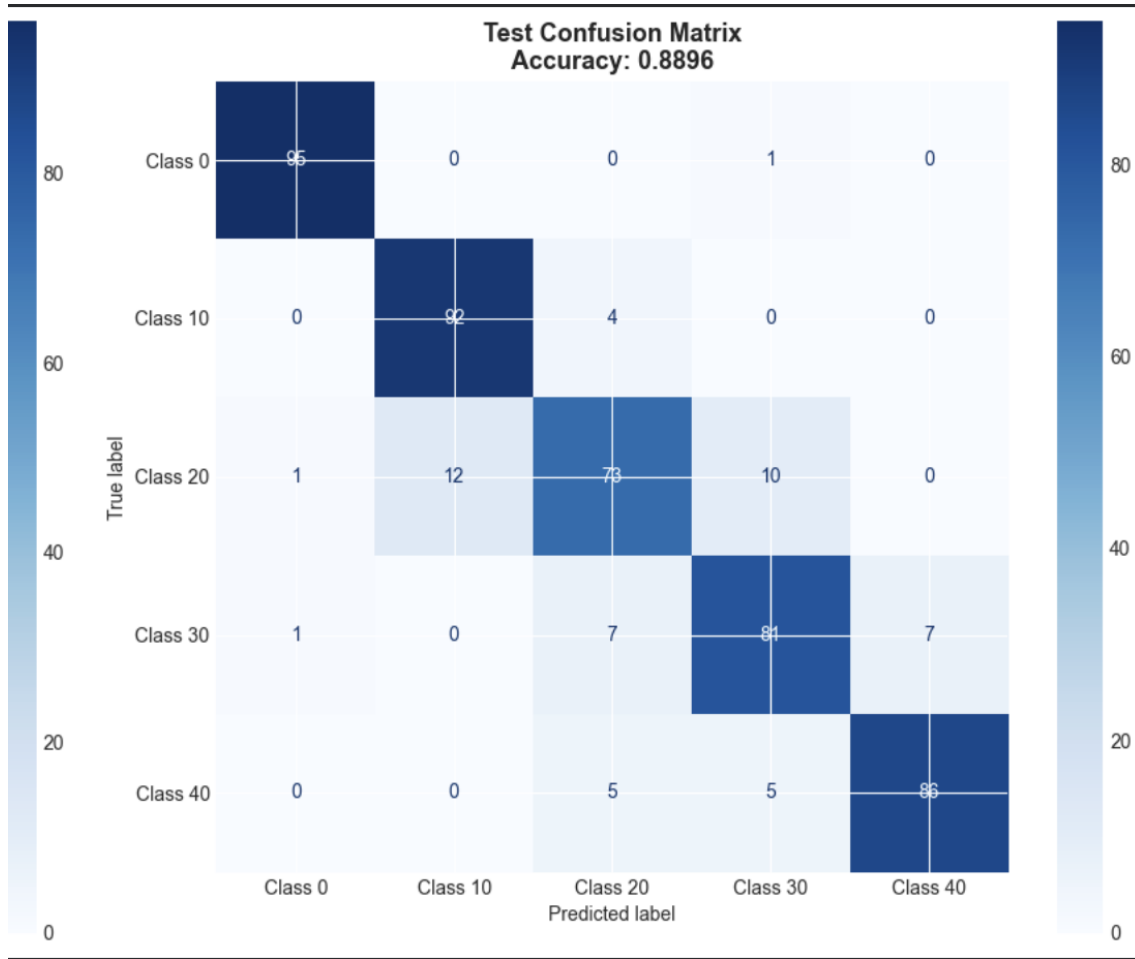


Figure 5.12: Confusion matrix of test data for the sugar dataset

In the case of sugar and meat data, a hold out validation design was adopted with a training, validation and testing split of 70, 15 and 15 respectively. The validation set was then used to measure the model performance to detect instances of overfitting and variance-related mistakes. The unseen test set was used to report final performance metrics because of an unbiased estimate of undesigned performance in the real world.

Besides metric based evaluation, the confusion matrix analysis was applied to determine the detailed error analysis on all models. Confusion matrices allowed misclassifications to be studied class-wise e.g. how often sugar samples containing 10 % adulterant were misclassified as 20 % adulterant. This discussion gave an understanding on certain confusions in classes, as well as the interpretation of the constraints of both models. In the meat freshness model case, statistical analysis also was supported by qualitative assessment with help of Grad-Cam. Grad-CAM was not a numerical statistical test, but was a method of explainability to help visually confirm high probability estimates were made on the basis of relevant spoilage characteristics, and not background artifacts which added more strength towards trusting the statistical outputs.

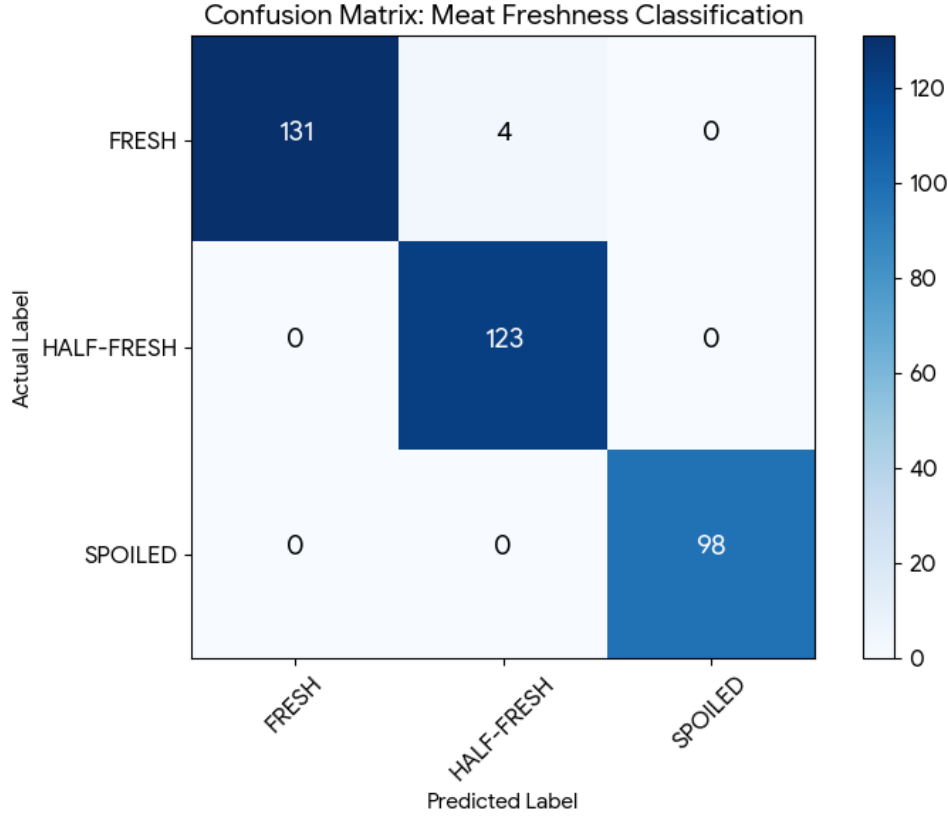


Figure 5.13: Confusion matrix of test data for the Meat Dataset

5.5 Comparisons and Relationships

As per the literature review that had been done keeping in mind to find different food adulteration techniques, it had been realised that even though sugar had been used as an adulterant for different foods commonly, there is still no adequate papers or existing works where the sugar itself has been adulterated and especially the adulteration detection part with $MgSO_4$. So, this study aimed to focus on sugar adulteration detection. Besides that, the adulteration detection of poppy seed with semolina is also something which has not many works done on, so the study tried to focus on this side of detecting semolina adulterants of poppy seed. Again, while conducting literature review a paper regarding meat freshness detection was thoroughly studied and from that paper, a new pretrained model was run on the collected meat data set, which did not only give better test accuracy but also it reduced the Misclassification cost by many folds, compared to the existing study of the models of the paper. Overall, this study was done on these three food items to focus on the part of adulteration detection of two selected food items and freshness detection of meat, with a view to using image data for adulteration or freshness detection purpose integrating with ML.

5.5.1 Sugar

At the preliminary stage, after preprocessing, selected models ResNet18, DenseNet121, and EfficientNetB0 were implemented on the sugar data. Their performances were

very average and results were not up to the expectations. ResNet18 was implemented as a separate model while DenseNet121 and EfficientNetB0 were implemented following a multiarchitectural model under the same framework. The models used a data normalization process, a dropout of 0.3, AdamW optimizer with label smoothing, augmentation, and the split into train, validation, and testing sets of 70, 15, 15 %ages respectively, as part of the implementation. The augmentation also included random rotations which would be applied to create more diversity to the dataset which could stop the model to memorize data. The model had an initial learning rate which is of 1×10^{-3} , a weight decay of 1×10^{-4} , and a batch size of 16. A warm-up learning rate strategy followed by cosine annealing and gradient clipping, were maintained throughout for optimization, to reduce overfitting and at the same time to ensure prevention of unnecessary training beyond the optimal epoch number as required was used. Models were basically trained on a categorical cross-entropy loss function along with fixed epochs. During the preliminary stage of this research, macro f1, precision and recall was calculated, however for the final selection of ensembled model the weighted f1, precision and accuracy was calculated. It was reported as weighted-average metrics as since they could represent the standard approach for multi-class classification. The dataset which was preprocessed, maintains balanced class distribution (e.g., 640 samples per class), which meaning weighted and macro averages converge to identical values. Weighted averaging was chosen finally as it maintains consistency with real-world deployment scenarios where class distributions may vary. Thus, all of these together helped in the optimization and implementation of the selected models but the results were not as expected so in the later stages, the study focused on building an ensembled model consisting of Densenet121, Convnextiny and Resnet50 with a voting classifier to build a more accurate and reliable model structure for adulteration detection of sugar. The comparison between various deep learning models to detect adulteration of sugars is presented in the table. Initial models used in the preliminary stage of research had moderate accuracy (66-72 %) which was then replaced by the final selected ensembled models as that had used higher test accuracy models like for Convnext Tiny (91.04 %) and ROC-AUC (98.64 %), and the next are DenseNet121 and ResNet50, which showed better results in detecting the level of adulteration and the comparison can be seen in the table below. The overall test accuracy of the ensembled model was 88.96% which is more reliable since it is not depending on any one model rather soft voting from three models to give final output.

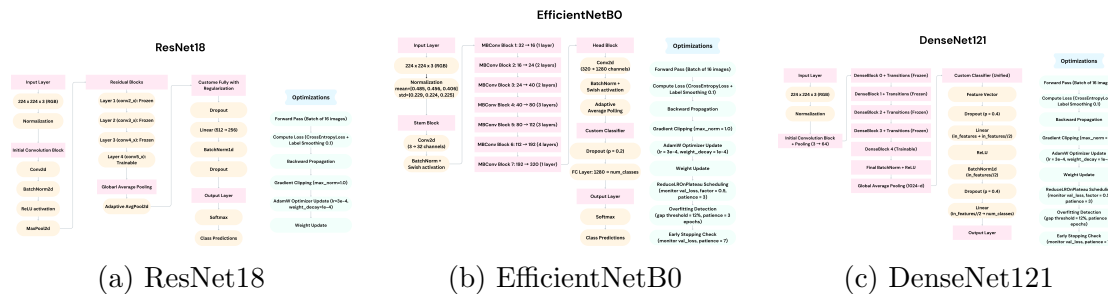


Figure 5.14: Preliminary experimented model architectures for sugar adulteration detection.

Table 5.5: Comparative performance analysis of deep learning models on sugar adulteration datasets

Dataset	Model	Test Accuracy (%)	Precision (Macro)	Recall (Macro)	F1-Score (Macro)	ROC-AUC (%)
Sugar with different levels of adulteration	ResNet18 (Preliminary)	71.04	71.37	71.76	70.43	–
Sugar with different levels of adulteration	EfficientNetB0 (Preliminary)	66.04	65.92	66.56	65.98	–
Sugar with different levels of adulteration	DenseNet121 (Preliminary)	72.08	71.96	72.75	72.11	–
Sugar Image Classification Dataset	ResNet50 (in the ensemble model)	82.50	82.34	82.50	82.23	96.86
Sugar Image Classification Dataset	DenseNet121 (in the ensemble model)	85.83	86.14	85.83	85.89	97.23
Sugar Image Classification Dataset	ConvNeXt-Tiny (in the ensemble model)	91.04	91.09	91.04	91.05	98.64

5.5.2 Poppyseed

To detect adulterated poppy seeds, 2 CNNs models were used such as ResNet50, and EfficientNetB0. At first what was tried is EfficientNetB0. EfficientNetB0 was optimized with Adam and StepLR scheduler and achieved an accuracy of 96.05% when tested with unseen data images. To ensure it was generalizing even better with a better result, further was added 10-fold stratified cross-validation, balanced classes, used AdamW optimizer with L2 weight decay = 0.1, ReduceLROnPlateau scheduler with factor = 0.5 and patience = 3, dropout layer = 0.3 and used advanced augmentation techniques. This optimization helped to achieve a mean test accuracy of $96.22\% \pm 1.45\%$ across 10 folds, an F1 score of 0.9614 and the best fold accuracy of 98.25%. Then to test if any other models would give similar accuracy other two models were tested though they did not give desired output. For ResNet50, at first, what used was ImageNet1Kv1 pre-trained weights and data augmentation which helped to get an test accuracy of 72.17% and an F1 score of 0.7278 when tested on unseen data. Later when further optimized it with 10-fold cross-validation and class weights as the dataset was small and the classes had an imbalanced number of data, it still did not help as much as the accuracy and F1 score was still lower than a satisfactory result.

The reason further more models were not explored for poppy seed neither an ensemble model was used like the one that was used for sugar dataset is because, the pretrained EfficientNetB0 was good enough on its own to detect adulterated and pure poppy seeds with very high accuracy, whereas the sugar ensemble models had to have a soft voting system of 3 CNN models to confidently and accurately find the pure and adulterated sugar. The optimized EfficientNetB0 has achieved a satisfying accuracy with good generalization with not so major overfitting/underfitting.

Table 5.6: Performance comparison of CNN models for poppy seed adulteration detection

Model	Optimization Strategy	Test Accuracy (%)	F1-Score	Remarks
EfficientNetB0	Adam / AdamW, StepLR + ReduceLROnPlateau, 10-fold Stratified CV, Class Balancing, Advanced Augmentation, Dropout (0.3)	98.25	0.9614	Best-performing model; strong generalization
ResNet50	ImageNet pre-trained, Data Augmentation, Adam optimizer, 10-fold CV, Class Weights	72.17	0.7278	Underperformed on small and imbalanced dataset

5.5.3 Meat

In this part, the proposed Swin Transformer-based solution is compared with the previous state-of-the-art models trained on the same set of meat freshness data in terms of architectural effectiveness, performance trade-offs, as well as business risk factors. Previous research mostly used CNN-based based on the ResNet-18, ResNet-50, and a hybrid UNet-DenseNet model where performance was limited to the small size of the dataset. Despite the reported accuracy of the fine-tuned ResNet-18 (93.13%) being the highest, the deep models such as the ResNet-50 did not yield any better results, suggesting that further depth increases in the network were not enough to capture the visual cues of spoilage (which were complex and non-uniform in meat) at a higher level. Conversely, the Swin Transformer builds upon the hierarchical self-attention and shifted windows based on which local indicators of the spoilage state and global changes in texture are modeled simultaneously, eliminating the representational limitations experienced by CNNs. In addition to accuracy, a misclassification cost metric was used in previous works to reduce business risk, which prefers less accurate models like ResNet-50 feature extraction based on fewer false negatives. Nevertheless, it involved no interpretability and reduced predictive reliability. The proposed solution is not compromised by such a trade since it provides high-classification results with an added Grad-CAM-based explainability to make the model a glass box rather than a black box. The system minimizes operational risk by facilitating human-checkable decisions (as opposed to the statistical minimization of costs), by giving visual justification of spoilage predictions to their reduction, and employing a more balanced and reliable paradigm than current approaches.

Table 5.7: Comparison of Proposed Swin Transformer Methodology against Existing State-of-the-Art Models [22]

Model	Accuracy	Precision	Recall	Cost
Swin Tiny (proposed model)	98.88%	98.91%	98.88%	\$16
ResNet 18-FE	82.71%	84.66%	83.71%	\$886
ResNet 18-FT	93.13%	93.97%	92.85%	\$5,076
ResNet 50-FE	88.03%	88.35%	88.86%	\$242
ResNet 50-FT	84.70%	84.50%	86.14%	\$316

5.6 Discussions

5.6.1 Interpretation

The article shows that deep learning is a potential, non-destructive, alternative to the conventional chemical testing to detect food adulteration. Nonetheless, the performance depends on the commodity and the complexity of tasks, which suggests that the strategy is more efficient with products than with other ones. In adulteration of sugar, the model tackled a more difficult problem compared with single classification: the concentration of $MgSO_4$ of 0% to 40%. The fact that ConvNext Tiny has reached 91.04% accuracy shows that AI can identify finer details of the texture of white crystals, but it is not perfect. There is confusion between the two adjacent levels of concentration i.e. 10 vs. 20 implying that at low concentration the visual difference between pure sugar and adulterated samples is harder to detect. This means that even an advanced feature extractor might not have enough discriminatory information even in RGB images. The fact that DenseNet slightly outperforms ResNet also indicates that the high frequency details or the fine edges and textures are important to the correct granular analysis. Conversely, the poppy seed model scored an almost perfect accuracy of 98.25 on the test, which means that the morphology of the fake poppy seeds and the usual adulterants are sufficiently different that convolutional neural networks can learn to make accurate predictions. The consistency within the cross-validation folds is indicative of high generalization, which can be implied that this model is robust and can be potentially implemented to the real world. This strong performance indicates that in the case of some commodities, AI-based detection can be not only fast but also very accurate without any specific equipment. The meat freshness model also underlines the role of interpretability in the AI-based food safety tools. Although the statistic accuracy was very high, the Grad-CAM images did not only give a good context of the areas of the image that affected the classification but also gave a necessary context. These patterns of discoloration observed in the model, as well as the fact that it learned relevant biological indicators of spoilage as opposed to irrelevant background artifacts, have been used to support the conclusion. This interpretability is essential to establishing confidence in consumers, regulators, and other stakeholders in the industry since it assists in vindicating that the algorithm is determining its choices according to scientifically relevant features.

5.6.2 Implications

The results of the current research have a large impact on the health of the population, consumer behaviour, and the viability of AI-based food safety applications. This quantification of adulteration as illustrated by the sugar model is especially valuable since it becomes possible to evaluate risk more accurately. The concentration of $MgSO_4$ or other adulterants is also available, which enables the stakeholders to determine the possible toxicity rather than using a mere scale of safe/unsafe. This helps to make better regulating decisions and health advisories. Equally, the high response in accuracy in the detection of poppy seeds indicates that AI applications can give small-scale traders and consumers the authority to check the quality of the products in real-time, without the need to conduct laboratory tests that are costly. That kind of democratization of testing can help decrease the market fraud, increase

the consumer trust, as well as, the reputation of the local and export markets. Moreover, the fact that these models use regular camera photos and still work proves that specialized devices, e.g. hyperspectral cameras, are not necessarily required in this case. This reduces the hurdle of creating mobile based solutions applicable in the low-resource setting, and food safety monitoring is more accessible and scalable.

5.6.3 Limitations

Although these are encouraging results, the research has a number of limitations that must be remembered when analyzing the findings. The inaccuracy or limitations of the sugar model (like test accuracy is not fully 100) means that there is a physical constraint to computer vision in the sense that an adulterant observed by an RGB image may not have sufficient information to allow perfect separation and the visually similar crystalline textures makes it difficult to detect adulteration absolutely perfectly. This visual ambiguity is especially difficult at low adulterant concentrations where crystals are virtually similar. The other weakness is associated with environmental reliance. Similar to any other computer vision system, the models are light sensitive. An image trained on bright laboratory images would perform poorly under dim conditions (as it is common in rural markets or storage premises) and this may affect the accuracy in the actual market use. Also, heavy architectures like DenseNet121 and Swin Transformers are highly accurate and need a large amount of computation. Implementing these models directly on models without optimization (e.g. quantization or pruning) could result in inferences taking long and could make them impractical in the field.

Chapter 6

Conclusion

6.1 Summary of Findings

This paper was able to show that the use of domain-specific deep learning architectures can be effective in automating the detection of food adulteration and spoilage across the variety of commodities. The experimental findings confirm the inadequacy of a one-size-fits-all strategy but instead, specific architectural strategies such as dense convolutional networks to hierarchical transformers are needed to respond to the visual peculiarities of granular, morphological, and biological data.

The findings of the sugar data show that deep learning can be successfully used to measure the level of contamination, despite the fact that the contaminant of interest is the same in visual characteristics with the final product, namely, MgSO_4 . Unlike other studies that have been conducted in the past and only examined mostly binary classification (pure vs. adulterated), this research study confirmed a multi-class level of detection that was capable of granulating the level of adulteration into five categories (0%-40%) for sugar. The highest validation accuracy of convnext tiny architecture reached 90.62% that could be explained by the combination of sophisticated regularization measures: Mixup, label smoothing, and stochastic depth. These methods were essential in training the model to identify tiny micro-textual differences in crystalline structures and not noise. To guarantee reliability in working, the framework used Ensemble Learning and Test-Time Augmentation, and an additional Mahalanobis distance-based Out-of-Distribution (OOD) detector was used to identify unknown contaminants. Conclusively, effective implementation of Grad-CAM established that ordinary smartphone macro photography, when run through dense connectivity networks, has enough resolution to be subject to intense and interpretable quantitative analysis. The overall performance of the ensembled model is also near to the test accuracy of the best performing model which is the convnext tiny, and the test accuracy is around 90%, which makes the ensembled model very reliable, compared to dependency of any single model.

The poppyseed experiments of the poppy were the most statistically sound results of the research. With an ultimate test precision of 98.25% and a 10-fold cross-validation mean precision of 97.19% (1.23%), the EfficientNetB0 framework created the final test accuracy. The low variance is exceptional which implies that the model itself is very stable and does not overfit. The results show that lightweight and compound-scaled architectures are ideally suited to be able to distinguish between fine-grained mixtures (semolina and poppy seeds) and can actually be extrapolated across various

environmental conditions (different background color black versus white).

The study has established that Swin Transformer presented a better approach to the study of the freshness of meat than a traditional CNNs (ResNet-18/50). Although past research has been concentrated on statistically reducing the business costs at the price of accuracy, Swin Transformer has been able to use shifted window attention to achieve both local discoloration and global texture degradation. Above all, a Glass Box validation mechanism was offered by the incorporation of Grad-CAM. The results indicate that the model decision making method incorporates the human biological heuristics i.e. decision-making based on particular spoilage areas and not the background artifacts thus meeting the unresolved critical issues of trust and safety of other previous solutions in the Black box.

Collectively, these results indicate that standard CNNs (EfficientNet, DenseNet) are the best models to use in static and texture-based adulteration (Sugar, Poppy Seeds), whereas Vision Transformers should be used in complex and non-uniform biological decay (Meat). The study hypothesizes that the quality and balanced data accompanied by explainable AI architecture could place food safety tests beyond the subjective and manual food inspection to automated and verifiable digitized infrastructure.

6.2 Contributions to the Field

The study contributes to the area of computer vision-based food safety analysis in several aspects directly related to the problem of quantifying granular adulteration, fine-grained texture, and biological spoilage.

The main contribution of this work is confirmation of available consumer safety in order to quantitatively analyze chemicals. This article dispels the widely held belief that crystalline analysis would need a costly hyperspectral or microscopic camera by showing that the commonly available smartphone camera with macro, paired with an ensembled architecture, could be used to quantify the concentration of $MgSO_4$ in sugar (0-40%) with satisfactory accuracy and it is one of the first adulteration detection strategy based study using image data for sugar. The ensemble comprises ConvNeXt, DenseNet and ResNet in their combination to execute more stable and robust predictions. The ensemble does not require the use of one model and prevents errors between samples. This trade-off works by enhancing the reliability and generalization that would be basically very useful in real-world sugar adulteration detection. In case of food safety, reliability is more important than higher accuracy, so that is why the ensemble model was developed and to cover any types of bias cases or corner cases. The sugar adulteration cases are hard to detect often for its similar grain structures and textures and where use of any single model may create biasness and for such cases ensembled model is the best, which also gives a very higher test accuracy nearing to almost 90%. It is one of the first studies that prove to work with adulterants detection of sugar with images. Such a discovery is a good way to reduce the barriers of entry to digital food safety testing, implying that mobile devices that are presently widespread may be used as credible edge-detection devices in supply chains and its image data can be used for adulteration detection. Again, it is one of the crucial studies that prove to apply hierarchical Vision Transformers (Swin Transformer) in the context of meat freshness classification. This work proves that Shifted Window Attention structures are better configured to capture

the non-uniform, multi-scale properties of biological decay (e.g. localized bacterial colonies vs. global textural breakdown) by going beyond the inert convolutional kernels of traditional CNNs (such as ResNet). This also defines a new architectural standard in the representation of intricate and organic surface erosion.

More importantly, this study will fill the gap between the high-performance Black Box algorithms and the reality of human verification in food safety. The study provides a way of visualizing model choices by engineering a Glass Box framework with Grad-CAM. This contribution in particular overcomes the limitations of previous works that were only optimizing the cost in the statistical sense. In its turn, our solution is more focused on safety and interpretability that enable the operators to see signs of spoilage (discoloration spots) and make disposal-related decisions with their own eyes and increase confidence in automated solutions.

Now, focusing on the poppy seed the study makes a rigorous validation of lightweight architectures of fine-grained mixture analysis through the poppy seed experiments. The experiment demonstrates that when combined with the state-of-the-art regularization methods (such as CutMix and Weighted Random Sampling), compound scaling can be trained to discriminate (near) perfectly (98.25) in visual similarity between granular textures (semolina vs. poppy seeds) in varying environmental backgrounds. For poppy seed, the one CNN model performed really well to detect its adulteration classes playing a great role in its detection and had satisfactory results. This discovery provides a blueprint that can be used to create low-resource, high-accuracy models that can be used in uncontrolled industrial conditions to detect food adulteration using image data. So, multiple food items with diverse models have been explored under the same study on the basis of detection using simple RGB images, proving that these images with AI and ML can play an important role for food adulteration detection as well.

6.3 Limitations and Future Work Plan

Despite the high levels of classification accuracy, the research suffers some constraints in the scope of the dataset and models generalization. Only MgSO_4 was modeled on four adulterant concentrations (10, 20, 30 and 40%), except urea, chalk powder and plastic grains, which are deemed common adulterants. The inherent weakness of the model is in the fact that the softmax architecture essentially compels the internal classification to fall into one of five fixed concentration classes, even though when the unknown adulteration could be correctly identified. Uncertainty quantification measures (Monte Carlo Dropout, Mahalanobis distance) are effective at identifying unknown adulterants as out-of-distribution but the model itself gives forceful results such as, "Unknown adulterant detected, like 10%." This is because the softmax layer has to place samples into the closest class of concentration of MgSO_4 on similarity of features, although the percentage concentrations of these classes do not matter to unknown adulterants. Also, substances visually similar to MgSO_4 can be mistaken as MgSO_4 instead of being identified as unknown.

Also, the sugar and poppy seed data were recorded in the laboratory under controlled conditions of a homogenous background and an equal amount of light. Although this guaranteed quality data and minimized noise, it creates a possible domain gap, that is, when used in the real world, the model will perform poorly because of variable illumination, cluttered backgrounds, or sensor noise. Both the datasets has shown

some very mild to none overfitting which was very difficult to prevent since the size of poppyseed dataset was small, and as of for sugar, the dataset size was medium but it has similar texture or granular appearances which makes it tough for the model to detect as per real world conditions, but it had no major overfitting or underfitting problem. For sugar, Grad-CAM was used to make predictions of the model visually interpretable. In other instances, however, the emphasized areas were not very obvious to the human eye. This is because this technique would be gradient-based, and the gradient-based nature of Grad-CAM is indicative of model sensitivity and does not necessarily signify physical localization of features. It is worth noting that one model would have made accurate predictions, and Grad-CAM can be deceptive and this is basically especially when classification is mainly based on distributed texture information but Grad-CAM implements a localized explanation. Grad-CAM was used on the sugar and meat adulteration data to visualize the model regions of the images that made the most impact on the model. In the case of the poppy seed data, there was no use of Grad-CAM since the dataset is small, and the model has a near-perfect performance (98% + accuracy) on all metrics, and thus, the interpretability through visualization was felt to be less urgent. In addition to that, the visual patterns of images of poppy seeds were also straightforward and homogenous, and hence no significant added insight would be provided by Grad-CAM. However, there are future scopes to work on with GRAD-CAM on this dataset as well, if more image data can be found or made or under conditions where visualization is a must. In the context of the freshness of meat study, the research was based on a freely-available dataset about red meat samples only. Consequently, it implies that the applicability of Swin Transformer model to alternative protein sources including poultry or seafood is limited. Besides, the half-fresh category is an intermediate and subjective position of spoilage. Even though this model was able to learn to differentiate this class visually, the lack of biochemical ground-truth measurements, e.g. microbial load or pH values, limits the biological accuracy and interpretability of this classification.

It is proposed that future work would be done to develop the proposed framework to real-life applicability and deployment. An ensemble model similar to that of sugar would be made and tested upon the poppyseed dataset and along with addition of GRADCAM. The next research direction that can be taken is to increase the sugar dataset by adding a variety of adulterants and mixed-contamination conditions, which allows building multi-class and multi-label detection models. In the context of the meat freshness, the future research will construct a proprietary multi-modal dataset by image data and biochemical sensor data (eighth future electronic nose measurement or pH) to objectively confirm transitional freshness conditions. In systems, future studies will focus on how to optimize computationally intensive architectures, such as DenseNet121, Swin Transformer, to make them operate on edges by pruning and quantizing models. Such optimization will facilitate the establishment of lightweight on-situ solutions that can detect adulteration in real time with a regular smart phone hardware. Lastly, future upgrading of the explainable AI module will go a step further as a visual heatmap to a text-readable explanation of the results, therefore, enhancing transparency, user confidence, and the usefulness of automated food safety systems.

Bibliography

- [1] K. O. Iwuozor, V. U. Anyanwu, B. O. Olaniyi, P. S. Mbamalu, and A. G. Adeniyi, “Adulteration of sugar: A growing global menace,” *Sugar Tech*, vol. 24, no. 3, pp. 914–919, May 2022. DOI: 10.1007/s12355-022-01122-6.
- [2] Y. K. Anagaw et al., “Food adulteration: Causes, risks, and detection techniques—review,” *SAGE Open Medicine*, vol. 12, pp. 1–10, 2024. DOI: 10.1177/20503121241250184.
- [3] EFSA Panel on Contaminants in the Food Chain (CONTAM), “Update of the scientific opinion on opium alkaloids in poppy seeds,” *EFSA Journal*, vol. 16, no. 5, e05243, May 2018. DOI: 10.2903/j.efsa.2018.5243.
- [4] K. H. Erna, K. Rovina, and S. Mantihal, “Current detection techniques for monitoring the freshness of meat-based products: A review,” *Journal of Packaging Technology and Research*, vol. 5, pp. 1–12, Aug. 2021. DOI: 10.1007/s41783-021-00120-5.
- [5] D. Meena, N. Tiwari, S. Mishra, and R. P. Pandey, “Digital image-based detection of wheat flour adulteration in turmeric powder: A deep learning approach,” *The Pharma Innovation Journal*, vol. 12, no. 11, pp. 1550–1557, Sep. 2023, ISSN: 2277-7695.
- [6] Z. W. Bhuiyan, S. A. R. Haider, A. Haque, M. R. Uddin, and M. Hasan, “Iot based meat freshness classification using deep learning,” *IEEE Access*, vol. 12, pp. 196 069–196 083, 2024. DOI: 10.1109/ACCESS.2024.3520029.
- [7] T. Chung et al., “Non-targeted detection of food adulteration using an ensemble machine-learning model,” *Scientific Reports*, vol. 12, no. 1, p. 20 956, Dec. 2022. DOI: 10.1038/s41598-022-25452-3.
- [8] E. Ahmed, “Detection of honey adulteration using machine learning,” *PLOS Digital Health*, vol. 3, no. 6, pp. 1–25, 2024. DOI: 10.1371/journal.pdig.0000536.
- [9] P. Lavatre and N. Bawane, “Identifying adulteration in food product using fumigation method and machine learning,” in *Proceedings of the 2024 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Application (ICAIQSA)*, 2024, pp. 1–8. DOI: 10.1109/ICAIQSA64000.2024.10882313.
- [10] S. Othman, N. R. Mavani, M. A. Hussain, N. A. Rahman, and J. M. Ali, “Artificial intelligence-based techniques for adulteration and defect detections in food and agricultural industry: A review,” *Journal of Agriculture and Food Research*, vol. 12, p. 100 590, 2023. DOI: 10.1016/j.jafr.2023.100590.

- [11] 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. DOI: 10.3390/foods14101808.
- [12] V. Surya and A. Senthilselvi, "A qualitative analysis of the machine learning methods in food adulteration: A focus on milk adulteration detection," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 12, no. 7, pp. 543–551, 2020. DOI: 10.5373/JARDCS/V12I7/20202037.
- [13] S. Z. Abidin, V. Kambali, U. Kumar, and L. K., "Food adulteration detection & sorting solution through machine learning," *International Advanced Research Journal in Science, Engineering and Technology*, vol. 10, no. 4, pp. 179–185, 2023. DOI: 10.17148/IARJSET.2023.10429.
- [14] U. Agrawal, N. Bawane, N. Alsubaie, M. S. Alqahtani, M. Abbas, and B. O. Soufiene, "Design & development of adulteration detection system by fumigation method & machine learning techniques," *Scientific Reports*, vol. 14, no. 1, p. 25366, 2024. DOI: 10.1038/s41598-024-64025-4.
- [15] P. Lavatre and N. Bawane, "Identifying adulteration in food product using fumigation method and machine learning," in *2024 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Application (ICAIQSA)*, IEEE, May 2024, pp. 1–6. DOI: 10.1109/ICAIQSA64000.2024.10882313.
- [16] A. Sattar, M. A. M. Ridoy, A. K. Saha, H. M. H. Babu, and M. N. Huda, "A comprehensive approach to detecting chemical adulteration in fruits using computer vision, deep learning, and chemical sensors," *IEEE Access*, vol. 12, pp. 46032–46047, 2024. DOI: 10.1109/ACCESS.2024.3384256.
- [17] P. P. Shinde, "Enhancing food quality assurance: Machine learning models for adulteration detection in edible oils," *International Journal of Food and Nutritional Sciences*, vol. 11, no. 8, pp. 5759–5763, 2022.
- [18] M. Aqeel, A. Sohaib, M. Iqbal, H. U. Rehman, and F. Rustam, "Hyperspectral identification of oil adulteration using machine learning techniques," *Current Research in Food Science*, vol. 8, p. 100773, 2024. DOI: 10.1016/j.crfs.2024.100773.
- [19] V. Mishra, S. P. Singh, S. K. Bhatia, V. S. Chandel, and R. Manohar, "Detecting adulteration in mustard oil using low-frequency dielectric spectroscopy," *Research in Agricultural Engineering*, vol. 70, no. 3, pp. 134–142, 2024. DOI: 10.17221/108/2023-RAE.
- [20] X. Duan, Z. Li, L. Wang, H. Lin, and K. Wang, "Engineered nanomaterials-based sensing systems for assessing the freshness of meat and aquatic products: A state-of-the-art review," *Comprehensive Reviews in Food Science and Food Safety*, vol. 21, no. 6, pp. 5136–5163, Nov. 2022. DOI: 10.1111/1541-4337.13074.
- [21] G. Song et al., "Research progress of chilled meat freshness detection based on nanozyme sensing systems," *Food Chemistry: X*, vol. 22, p. 101364, Jun. 2024. DOI: 10.1016/j.fochx.2024.101364.

- [22] B. Sagiraju, N. Casanova, I. C. C. Lam, M. Lohia, and T. Yoshiyasu, “Meat freshness prediction,” *arXiv preprint arXiv:2305.00986*, 2023. DOI: 10.48550/arXiv.2305.00986. [Online]. Available: <https://doi.org/10.48550/arXiv.2305.00986>.
- [23] B. Avula, K. Katragunta, S. J. Adams, Y.-H. Wang, A. G. Chittiboyina, and I. A. Khan, “Applicability of lc-qtof and microscopical tools in combating the sophisticated, economically motivated adulteration of poppy seeds,” *Foods*, vol. 12, no. 7, p. 1510, 2023. DOI: 10.3390/foods12071510. [Online]. Available: <https://doi.org/10.3390/foods12071510>.
- [24] Z. K. Nagy, G. Miklós, G. Tölmery, B. Csehi, and Z. Kormosné Bugyi, “Investigating the effect of different grinding conditions and methods on the concentration of opium alkaloids in poppy seeds as a good reduction practice,” *LWT*, vol. 191, p. 115 664, Jan. 2024. DOI: 10.1016/j.lwt.2023.115664.
- [25] O. Bińczak and J. Zembrzuska, “Determination of opium alkaloid content in poppy seeds using liquid chromatography coupled with a mass spectrometer with a time-of-flight analyzer,” *Foods*, vol. 11, no. 22, p. 3646, Nov. 2022. DOI: 10.3390/foods11223646.
- [26] H. Ali, K. Rafique, R. Ullah, M. Saleem, and I. Ahmad, “Classification of sidr honey and detection of sugar adulteration using right angle fluorescence spectroscopy and chemometrics,” *European Food Research and Technology*, vol. 248, pp. 1823–1829, Apr. 2022. DOI: 10.1007/s00217-022-04008-9.
- [27] K. O. Iwuozor, L. A. Ogunfowora, and I. P. Oyekunle, “Review on sugarcane-mediated nanoparticle synthesis: A green approach,” *Sugar Tech*, vol. 24, no. 4, pp. 1186–1197, Aug. 2022. DOI: 10.1007/s12355-021-01038-7.
- [28] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. DOI: 10.1038/nature14539.
- [29] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 770–778.
- [30] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, “Convnext: A convnet for the 2020s,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2022, pp. 11 976–11 986.
- [31] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2017, pp. 4700–4708. DOI: 10.1109/CVPR.2017.243.
- [32] M. Tan and Q. V. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” in *Proceedings of the International Conference on Machine Learning (ICML)*, PMLR, 2019, pp. 6105–6114.
- [33] Z. Liu et al., “Swin transformer: Hierarchical vision transformer using shifted windows,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, 2021, pp. 10 012–10 022.
- [34] Mendeley Data, *Image dataset of pure and adulterated poppy seeds using semolina*, <https://data.mendeley.com/datasets/z7tr4b6tw/1>, [Data set], 2023.

- [35] V. Vinayakshanawad, *Meat freshness image dataset*, Accessed: 2025, Kaggle, 2020. [Online]. Available: <https://www.kaggle.com/datasets/vinayakshanawad/meat-freshness-image-dataset>.
- [36] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2009, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.
- [37] G. J. McLachlan, “Mahalanobis distance,” *Resonance*, vol. 4, no. 6, pp. 20–26, Jun. 1999.
- [38] A. Paszke et al., “Pytorch: An imperative style, high-performance deep learning library,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2019, pp. 8026–8037.
- [39] Z. Zhong et al., “Random erasing data augmentation,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, 2020, pp. 13 001–13 008.
- [40] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 4700–4708.
- [41] I. Loshchilov and F. Hutter, “Decoupled weight decay regularization,” in *International Conference on Learning Representations (ICLR)*, 2019.
- [42] I. Loshchilov and F. Hutter, “Sgdr: Stochastic gradient descent with warm restarts,” in *International Conference on Learning Representations (ICLR)*, 2017.
- [43] H. Zhang et al., “Mixup: Beyond empirical risk minimization,” in *International Conference on Learning Representations (ICLR)*, 2018.
- [44] G. Huang, Y. Sun, Z. Liu, D. Sedra, and K. Q. Weinberger, “Deep networks with stochastic depth,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, Springer, Oct. 2016, pp. 646–661.
- [45] Y. Gal and Z. Ghahramani, “Dropout as a bayesian approximation: Representing model uncertainty in deep learning,” in *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, Jun. 2016, pp. 1050–1059.
- [46] K. Lee, H. Lee, K. Lee, and J. Shin, “A simple unified framework for detecting out-of-distribution samples and adversarial attacks,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2018, pp. 7167–7177.
- [47] R. R. Selvaraju et al., “Grad-cam: Visual explanations from deep networks via gradient-based localization,” in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 618–626.
- [48] C. Szegedy et al., “Rethinking the inception architecture for computer vision,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818–2826.