Gongcheng Kexue Xuebao

ISSN 2095-9389

Volume 11 || Issue 6 || 2026

Impact Factor: 6.7

https://kexuexuebao.org/

Editorial Board
GONGCHENG KEXUE XUEBAO

Table Of Contents

An automated approach for fruits and vegetable image enhancement and classification using GAN

Dr. Suvarna Nandyal
Prof, Department of Computer science and engineering
PDA College of engineering
Kalaburagi,India

Mohammad Masroor Arbaz Khan
Student, Department of Computer science and engineering
PDA College of engineering
Kalaburagi,India

Syed Abdur Rafey Anhar
Student, Department of Computer science and engineering
PDA College of engineering
Kalaburagi,India

Abstract - **Fruits and vegetables are essential components of human nutrition, providing vital vitamins, minerals, antioxidants, and dietary fibers that help prevent chronic diseases and maintain overall well-being. In modern society, the accurate recognition and classification of these food items have become increasingly important across multiple domains: agriculture, food processing, retail markets, health monitoring, and supply chain automation. Traditional recognition methods rely on manual inspection, which is highly subjective, labor-intensive, and prone to errors, particularly when dealing with large quantities of produce.**

With the rise of artificial intelligence (AI), particularly computer vision and deep learning, automated fruit and vegetable recognition systems have gained traction. However, a major challenge persists: image quality. In real-world environments, images of fruits and vegetables are often captured under poor lighting conditions, with occlusions, motion blur, or low resolution due to inexpensive cameras or mobile devices. Such degraded images significantly reduce classification accuracy.

To address these limitations, this work proposes an automated approach for fruits and vegetable image enhancement and classification using Generative Adversarial Networks (GANs). By enhancing image quality before classification, the system improves recognition accuracy, making it robust and practical for real-world applications. Such as grading of fruits at market. the work is helpful for farmers and online shopping sites for quality check.

Keywords-*Fruit, vegetables, GAN*

I. INTRODUCTION

1.1 Background of the Study

Fruits and vegetables form the cornerstone of human nutrition, offering essential vitamins, minerals, dietary fibers, and phytochemicals that play a pivotal role in maintaining health and preventing diseases. With the increasing emphasis on healthy diets, the demand for fruits and vegetables has been steadily rising across the globe. However, the identification, classification and quality control of fruits and vegetables remain complex and challenging tasks in modern agriculture and food industries.

Traditional methods of recognition rely heavily on manual inspection by experts or workers in markets, farms, and factories. While effective in small-scale environments, manual methods suffer from subjectivity, inconsistency and inefficiency when applied to large-scale operations. With globalization and rapid expansion of food supply chains, there is an urgent need for automated recognition systems capable of delivering accurate, consistent and real-time classification.

The advent of artificial intelligence (AI), Machine Learning (ML), and Deep learning (DL) has revolutionized image recognition and classification. Convolutional Neural Networks (CNNs), in particular, have demonstrated extraordinary performance in object detection and classification tasks. Despite these advancements, one critical challenge persists: low-quality images.

Images of fruits and vegetables are often captured in uncontrolled environments where lighting conditions, camera quality, angles and background noise significantly impact image clarity. Poor image quality directly affects the classification accuracy of CNN models. Thus, enhancing images before classification becomes essential for building a robust recognition system.

Generative Adversarial Networks (GANs), introduced by Ian Goodfellow in 2014, have emerged as one of the most powerful tools in deep learning for image enhancement. GANs can generate realistic high-resolution images from low-quality inputs by learning complex mappings between degraded and enhanced domains.

This proposed work leverages GANs for image enhancement and CNNs for classification, resulting in a dual-stage automated system designed for the recognition of fruits and vegetables.

II.LITERATURE SURVEY

2.1 Related works

The literature survey provides a comprehensive overview of prior research and developments in the fields of fruit and vegetable recognition, image enhancement, and deep learning-based classification techniques. The purpose of a literature survey is to critically evaluate existing studies, understand the strengths and limitations of various approaches, and identify research gaps that justify the proposed system.

With the rise of machine learning and deep learning, the domain of agricultural computer vision has experienced significant advances. Automated recognition of fruits and vegetables is essential in multiple applications, including:

- Agriculture: crop monitoring, yield estimation, and quality assessment
- Food Industry: sorting, grading, and automated packaging
- Retail: smart checkout systems and inventory management
- Healthcare: dietary monitoring and nutritional analysis

https://doi.org/10.5281/zenodo.20502198

Despite advancements, several challenges persist in accurately classifying fruits and vegetables, especially under real-world conditions where lighting, background clutter, occlusions and low-resolution images affect performance. Therefore, a systematic analysis of prior research is necessary to design robust recognition systems.
This literature survey is organized into the following categories for clarity:

1. Studies on image classification using machine learning and deep learning – Examines foundational techniques and their evolution from traditional machine learning to modern deep learning architectures.
2. Studies on fruit and vegetable recognition – Focuses on domain-specific research, datasets, and application-specific challenges.
3. Research on image enhancement techniques, with a focus on GANs – Discusses approaches to improve image quality, which directly affects classification accuracy.
4. Hybrid approaches integrating enhancement and classification – Reviews studies that combine preprocessing/enhancement with classification models.
5. Identification of research gaps – Highlights the limitations in existing literature, motivating the need for the proposed GAN and CNN system.

2.2 Deep Learning in Image Classification

The emergence of Convolutional Neural Networks (CNNs) marked a paradigm shift in image classification. Unlike traditional machine learning methods, which require manual feature extraction, CNNs can automatically learn hierarchical representations of images directly from raw pixel data. This capability enables CNNs to detect low-level features (edges, corners) in early layers and more complex patterns (shapes, textures) in deeper layers.
CNNs have achieved remarkable success across various domains, including handwritten digit recognition, object detection, facial recognition, medical imaging, and fruit and vegetable classification. The evolution of CNN architectures over the past two decades has progressively enhanced their accuracy, efficiency, and scalability.

2.2.1 Key CNN Architectures

1. LeNet-5 (1998)
   - Developed by Yann LeCun et al., LeNet-5 was among the first successful CNN models for digit recognition (MNIST dataset).
   - Introduced convolutional layers, pooling layers, and fully connected layers, establishing the foundation for modern CNN architectures.
   - Significance: Demonstrated that hierarchical feature extraction using convolution could significantly outperform traditional machine learning methods.
2. AlexNet (2012)
   - Won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) with a top-5 error rate of 15.3%, far better than previous approaches.
   - Innovations:
     - ReLU activation functions for faster convergence
     - Dropout layers to reduce overfitting
     - Use of GPU acceleration to handle deeper architectures efficiently
   - Significance: Sparked widespread adoption of deep learning in computer vision.
3. VGGNet (2014)
   - Proposed by Simonyan and Zisserman, VGGNet demonstrated the effectiveness of very deep networks using small (3×3) convolution kernels.
   - Advantages:
     - Simpler architecture that is easy to implement and scale
     - Deeper layers capture more abstract features
   - Popular for transfer learning and feature extraction in specialized domains like agriculture.
4. ResNet (2015)
   - Introduced residual learning, allowing very deep networks (up to 152 layers) to converge without suffering from vanishing or exploding gradients.
   - Residual blocks add shortcut connections that bypass layers, enabling smoother gradient flow.
   - Significance: Facilitates accurate classification for complex datasets and large-scale image recognition tasks.

2.2.2 Application to Fruits and Vegetables

CNNs have been widely applied in agricultural computer vision tasks, particularly for fruit and vegetable recognition. Compared to traditional methods, CNNs provide:

- Higher accuracy in distinguishing visually similar fruits
- Automatic feature extraction, reducing dependency on domain expertise
- Scalability for large and diverse datasets

Examples of CNN Applications:

1. Custom CNN for Fruit Recognition
   - A study implemented a custom CNN to classify 15 fruit categories.
   - Achieved 94% accuracy, demonstrating the effectiveness of CNNs over color/shape-based methods.
2. MobileNet-Based Models
   - Lightweight CNN architectures designed for mobile and embedded devices.
   - Effective for real-time applications such as mobile apps for fruit recognition.
   - Trade-off: Slightly lower accuracy compared to deeper networks but highly efficient.

Limitation:

- CNN performance drops when input images are low-resolution, noisy, blurred, or captured under uncontrolled lighting conditions.
- These challenges necessitate preprocessing or enhancement methods before classification to maintain high accuracy in real-world applications.

2.3 Research on Fruit and Vegetable Recognition

Several studies have specifically focused on fruit and vegetable classification using deep learning, demonstrating both progress and limitations in this field:

1. Kaur et al. (2019)
    - Implemented CNNs to classify 10 fruit classes.
    - Achieved accuracy above 90%.
    - Dataset: Controlled images captured under consistent lighting.
    - Limitation: Model accuracy may decrease under variable lighting or real-world conditions.
2. Singh et al. (2020)
    - Developed a multi-class classification system using ResNet50.
    - Achieved high accuracy due to deep residual learning.
    - Limitation: High computational cost makes real-time deployment challenging.
3. Aamir et al. (2021)
    - Implemented fruit recognition on embedded devices using MobileNet.
    - Advantages: Lightweight and suitable for real-time applications.
    - Limitation: Struggled with noisy or low-quality images, indicating a need for enhancement modules.
4. Dataset Contributions
    - Publicly available datasets, such as Fruit-360, have accelerated research by providing labeled, diverse fruit and vegetable images.
    - Benefits: Enables reproducible experiments and benchmarking.
    - Limitation: Many datasets consist of clean images captured under controlled conditions, limiting generalization to real-world scenarios.

III. OBJECTIVES

1. To implement a GAN model capable of enhancing image quality, especially for low-resolution, blurred, and noisy inputs.
2. To train and fine-tune a CNN model for accurate classification of 36 types of fruits and vegetables.
3. To integrate both modules into a unified pipeline where the output of the GAN becomes the input to the CNN.
4. To develop a Streamlit-based web interface allowing real-time interaction, image upload, and classification.
5. To evaluate and compare the performance of the classifier with and without GAN preprocessing using key performance metrics.

IV. EXISTING SYSTEM

4.1 Introduction

Before proposing a new approach, it is critical to analyze and understand existing systems for fruit and vegetable recognition and image enhancement. Studying prior research provides insight into the strengths, weaknesses, and gaps in current methodologies and establishes a foundation for designing a more robust system.

Most existing systems rely heavily on deep learning techniques, particularly Convolutional Neural Networks (CNNs), for classification tasks. CNNs have demonstrated remarkable performance due to their ability to automatically learn hierarchical features from raw images, eliminating the need for manual feature extraction.

In addition to CNNs, various image preprocessing techniques are employed to improve the quality of input images, including classical methods like histogram equalization, filtering, and modern deep learning-based approaches such as autoencoders and Generative Adversarial Networks (GANs). These methods aim to enhance resolution, reduce noise, and improve contrast, directly impacting classification accuracy.

However, despite these advancements, current systems face significant challenges in real-world applications:

- Images captured in farms, markets, or kitchens often exhibit poor lighting, occlusion, cluttered backgrounds, and motion blur, which degrade classifier performance.
- Many models are trained on controlled, clean datasets that do not reflect the diversity and variability of real-world scenarios.
- Integration between enhancement and classification modules is limited, reducing the robustness of deployed systems.

This section provides a detailed overview of existing classification methods, image enhancement techniques, and their limitations, laying the groundwork for the proposed GAN integrated CNN hybrid approach.

4.2 Existing Fruit and Vegetable Classification Systems

Fruit and vegetable recognition has evolved significantly, transitioning from traditional machine learning approaches to deep learning-based methods. This section analyzes both categories, highlighting advantages, limitations, and suitability for real-world applications.

4.2.1 Traditional Machine Learning Approaches

Early fruit and vegetable classification systems relied on handcrafted features to describe image properties such as color, texture, and shape. These features were then fed into classical classifiers to distinguish between different classes.

Feature Extraction Techniques:

1. Color Histograms – Represent the distribution of colors in an image (RGB, HSV, or Lab).
2. Texture Descriptors – Gray-Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), and Gabor filters capture surface texture and patterns.
3. Shape-Based Features – Contour descriptors, aspect ratio, roundness, and convexity characterize geometric properties of fruits and vegetables.

Classifiers Used:

- Support Vector Machines (SVMs) – Effective for linearly separable features, commonly used in early studies.
- k-Nearest Neighbors (k-NN) – Non-parametric, simple to implement but computationally expensive for large datasets.
- Decision Trees – Hierarchical feature thresholding for class separation.

Advantages:

- Easy to implement and computationally less expensive.
- Works reasonably well for small, controlled datasets.
- Provides interpretability of features and classification rules.

Limitations:

- Poor generalization on large or diverse datasets.
- Highly sensitive to noise, lighting variations, and occlusion.
- Cannot handle intra-class variations, such as different shapes or colors of the same fruit.
- Feature engineering is manual and dataset-specific, requiring expert knowledge.

4.2.2 Deep Learning Approaches

With the advent of Convolutional Neural Networks (CNNs), fruit and vegetable recognition achieved significant improvements in accuracy and scalability. CNNs automatically learn hierarchical features directly from images, removing the dependency on handcrafted descriptors.

Datasets Enabling Deep Learning:

- Fruit-360 – Contains thousands of images across multiple fruit and vegetable classes, captured under consistent conditions.
- Other agricultural datasets providing labeled images of fruits and vegetables in diverse environments.

Common CNN Architectures Used:

1. VGG16/VGG19
   - Deep networks with multiple convolutional layers and small 3×3 kernels.
   - Strength: Captures hierarchical features from simple to complex patterns.
2. ResNet (Residual Networks)
   - Introduces residual connections to overcome vanishing gradients in very deep networks.
   - Strength: Enables training of deep networks with high accuracy.
3. Inception Models
   - Multi-scale convolutional filters allow efficient feature extraction at different resolutions.
   - Strength: Handles variability in object size and shape effectively.
4. Custom CNNs
   - Tailor-made architectures designed for specific datasets, optimized for low-latency or resource-constrained environments.

Advantages:

- High classification accuracy on clean, well-prepared datasets.
- Eliminates the need for manual feature extraction, reducing domain-specific labor.
- Capable of handling larger datasets and more classes than traditional approaches.

Limitations:

- Accuracy significantly drops on noisy, low-resolution, blurred, or poorly illuminated images.
- Dependence on clean and balanced datasets; performance decreases with unbalanced classes.
- Lack of preprocessing or enhancement leads to loss of critical information, reducing real-world applicability.

4.3 Existing Image Enhancement Techniques

Image enhancement is a crucial step in improving the visual quality of input images before they are fed into a classification model. High-quality images allow classifiers to extract meaningful features, resulting in better recognition accuracy. Image enhancement techniques are broadly divided into classical image processing methods and deep learning-based approaches.

4.3.1 Classical Image Processing

Classical image processing techniques have been widely employed for decades to improve image quality. Common methods include:

1. Histogram Equalization
   - Adjusts image intensity distribution to enhance contrast.
   - Particularly useful for images captured under uneven lighting conditions.
2. Gaussian Filtering
   - Smooths images by reducing noise using a Gaussian kernel.
   - Effective for eliminating high-frequency noise, but may blur edges.
3. Median Filtering
   - Replaces each pixel with the median value of its neighborhood.
   - Effective for removing salt-and-pepper noise while preserving edges.
4. Image Sharpening
   - Enhances edges and details to make objects more visually prominent.

Advantages:

- Computationally inexpensive and simple to implement.
- Can be applied on-the-fly in low-resource environments.

Limitations:

- Cannot recover lost details in low-resolution or blurred images.
- Excessive application may introduce artifacts, such as halo effects or noise amplification.
- Not adaptive to complex variations like varying illumination, motion blur, or real-world backgrounds.

4.3.2 Deep Learning-Based Enhancement

With the advent of deep learning, Convolutional Neural Networks (CNNs) have been applied for tasks such as image super-resolution, denoising, and deblurring.

1. Super-Resolution CNN (SRCNN)
   - Upscales low-resolution images to higher resolutions while reconstructing details.

Advantages:

- Can reconstruct image details better than classical methods.
- Capable of adapting to different noise patterns and resolutions.

Limitations:

- Performance is often limited compared to GAN-based methods, especially in generating sharp textures.
- May produce smoothed or less realistic images when trained solely on pixel-wise loss functions.

4.4 GANs in Image Enhancement (Current Research)

Generative Adversarial Networks (GANs) represent a state-of-the-art approach for image enhancement and super-resolution. GANs consist of a generator network, which creates high-quality images, and a discriminator network, which evaluates the realism of generated images.

Key GAN Models:

1. SRGAN (Super-Resolution GAN) – Generates photo-realistic high-resolution images from low-resolution inputs.
2. ESRGAN (Enhanced SRGAN) – Improves upon SRGAN by introducing residual-in-residual dense blocks and perceptual loss, producing sharper and more detailed images.
3. DeblurGAN – Specializes in motion blur removal, restoring clarity in captured images.
4. CycleGAN – Performs image-to-image translation, useful when paired datasets are unavailable.

Advantages of GANs:

- Ability to generate sharp, realistic images with fine details.
- Can recover lost textures from low-resolution or noisy images.
- Potential for direct integration into classification pipelines, improving accuracy when images are enhanced before feature extraction.

Limitations in Current Systems:

- Rarely applied to fruit and vegetable recognition, limiting practical agricultural applications.
- Computationally demanding, particularly for training, which can hinder real-time deployment without optimization.
- Requires careful tuning of generator and discriminator networks to prevent mode collapse or artifacts.

4.5 Summary

The existing systems for fruit and vegetable recognition primarily rely on CNNs for classification and simple preprocessing techniques for image quality improvement. While these approaches have achieved good accuracy on controlled datasets, they fail to generalize in real-world scenarios due to noise, poor lighting, and low resolution. GAN-based image enhancement has shown promise but has not been fully integrated into agricultural classification tasks.

This gap in the literature highlights the need for a proposed system that combines GAN-based enhancement with CNN-based classification, ensuring improved robustness, accuracy, and applicability.

V. PROPOSED SYSTEM

5.1 Introduction

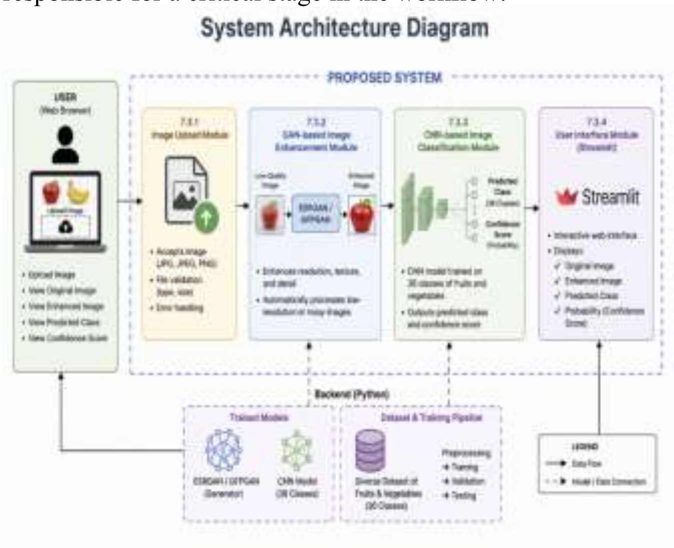
The proposed system is designed to overcome the limitations of existing fruit and vegetable recognition systems by integrating Generative Adversarial Networks (GANs) for image enhancement with Convolutional Neural Networks (CNNs) for classification. Existing approaches often struggle with low-quality, noisy, or blurred images captured under real-world conditions, which significantly reduces classification accuracy. The dual-stage approach of the proposed system ensures that images, regardless of initial quality, are enhanced to a high visual fidelity before being analyzed by the classifier. This integrated pipeline offers several key advantages:

1. Robustness
   - The GAN enhancement module allows the system to handle images with low resolution, poor lighting, occlusions, or noise, which are common in agricultural fields, marketplaces, or kitchen environments.
2. Scalability
   - The system is designed to process multiple fruit and vegetable categories efficiently, making it suitable for large-scale applications such as farm monitoring, inventory management in retail, and dietary analysis.
3. Deployability
   - By leveraging a Streamlit-based web interface, the system enables real-time image upload, enhancement, and classification, allowing users to interact with the model without requiring technical expertise.
4. Dual Output
   - Provides both the enhanced image and the classification result, ensuring users can visually assess improvements while receiving accurate identification and information.

This proposed system bridges the gap between state-of-the-art image enhancement techniques and practical classification applications, delivering a comprehensive, user-friendly, and deployable solution.

5.2 System Overview

The proposed system is organized into four main modules, each responsible for a critical stage in the workflow:



1. GAN-based Image Enhancement Module

- The proposed system integrates a Generative Adversarial Network (GAN) to improve the visual quality of low-resolution fruit and vegetable images before classification.
- ESRGAN/GFPGAN models are utilized to restore fine textures, sharpen edges, and reduce image noise.
- The enhancement module increases image clarity, which helps the classifier extract more meaningful features.
- The GAN model consists of two networks: a generator and a discriminator.
    - The generator produces enhanced images from degraded inputs.
    - The discriminator evaluates whether the generated image appears realistic.
- Super-resolution techniques are applied to recover missing details from blurred or compressed images.
- The module automatically processes uploaded images without requiring manual adjustments from the user.
- Enhanced images improve recognition accuracy, especially for poor-quality or dimly lit inputs.
- The preprocessing stage performed by the GAN reduces the impact of noise and distortion in real-time image analysis.
- The system supports enhancement of multiple image conditions such as low brightness, blur, and reduced resolution.
- The improved image quality contributes to better CNN feature extraction and classification performance.

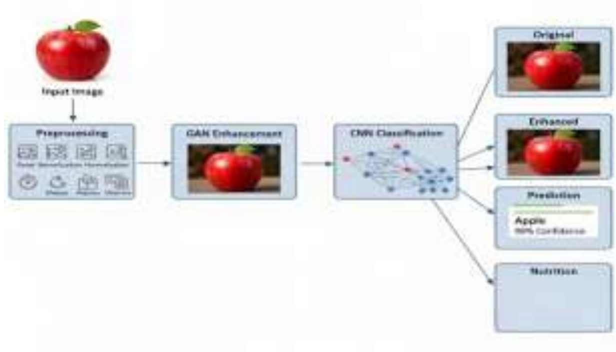
2. CNN-based Image Classification Module

- A Convolution Neural Network (CNN) is employed to classify fruits and vegetables into 36 different categories.
- The CNN model automatically learns important visual features such as color, shape, texture, and patterns from training images.
- Multiple convolution and pooling layers are used to extract hierarchical image features effectively.
- The classification model is trained using a diverse dataset containing healthy and defective fruit and vegetable samples.
- Data preprocessing techniques such as resizing, normalization, and augmentation are applied to improve model generalization.
- The trained CNN predicts the class label along with a confidence score representing prediction probability.
- Softmax activation is used in the output layer to determine the probability distribution across all classes.
- The model minimizes classification error using backpropagation and optimization algorithms during training.
- The CNN architecture enables automatic feature extraction, eliminating the need for manual feature engineering.
- The classification module achieves efficient and accurate prediction suitable for real-time applications.
- Integration of the enhancement module with CNN improves overall classification reliability and robustness.
- The trained model can identify visually similar fruits and vegetables with high precision.
-

3. Combined System Contribution Points

- The proposed framework combines GAN-based image enhancement with CNN-based classification for improved prediction accuracy.
- Image enhancement and classification are performed sequentially to ensure better feature quality for recognition.
- The system provides an end-to-end automated solution for fruit and vegetable identification.
- The Streamlit-based interface enables users to upload images and instantly obtain enhanced outputs and predictions.
- The proposed architecture reduces the effect of poor image quality on classification performance.
- Experimental results demonstrate that enhanced images lead to more consistent and accurate predictions compared to raw images alone.

5.3 Workflow Diagram



VI. METHODOLOGY

6.1 Introduction
The methodology section provides a step-by-step description of how the proposed fruit and vegetable recognition system is implemented. It outlines the procedures followed to ensure high accuracy, robustness, and reproducibility, covering every stage from data collection to model evaluation.
The proposed system integrates Generative Adversarial Networks (GANs) for image enhancement and Convolutional Neural Networks (CNNs) for classification. This methodology emphasizes:

- Systematic Data Preparation: Collecting and preprocessing diverse images representing real-world scenarios.
- Image Enhancement: Using GANs to improve low-quality images, restoring sharpness, resolution, and fine details.
- Feature Extraction and Classification: Employing CNNs to learn hierarchical features for accurate fruit and vegetable identification.
- Evaluation and Analysis: Measuring performance using standard metrics such as accuracy, precision, recall, F1-score, and visual inspection of enhanced images.

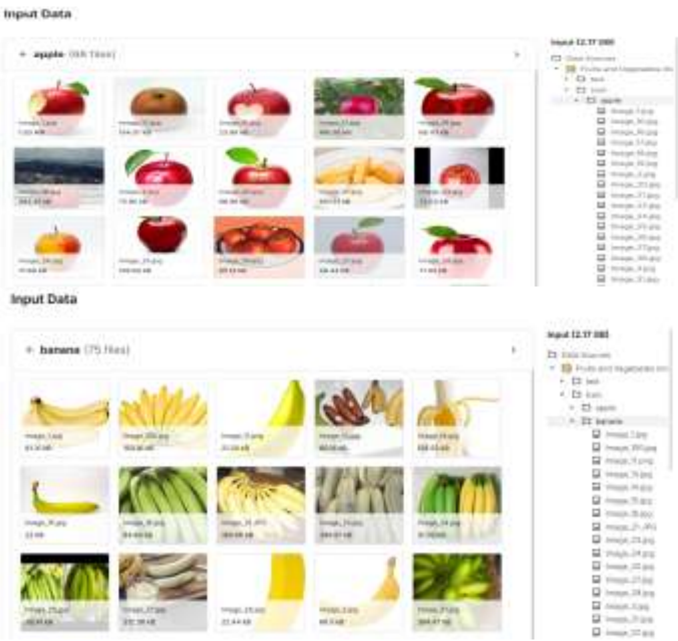
This structured approach ensures that the system can handle low-quality, noisy, or diverse input images, making it practical for applications in agriculture, retail, food processing, and healthcare.

6.2 Data Collection

High-quality, diverse datasets are essential for training deep learning models. The data collection process for the proposed system focuses on capturing variability in fruit and vegetable images, simulating real-world conditions.

6.2.1 Dataset Sources

1. Public Datasets:



Fruit-360: Provides thousands of labeled images of fruits and vegetables captured under controlled conditions.

   - Kaggle Fruit and Vegetable Datasets: Offers images with multiple classes, including both fruits and vegetables, useful for model training and benchmarking.

2. Custom Captured Images:
   - Photographs taken in real environments such as farms, markets, and kitchens.
   - Captured using mobile phones and DSLR cameras to ensure variability in quality, resolution, and background conditions.
   - Enhances the model's generalization ability by exposing it to diverse scenarios not present in controlled datasets.

6.2.2 Dataset Composition

The dataset is curated to cover a wide range of fruits and vegetables, including but not limited to:

- Fruits: Apples, bananas, oranges, grapes, mangoes, kiwi, pears, watermelon, pineapple, pomegranate.
- Vegetables: Carrots, cucumbers, tomatoes, onions, potatoes, bell peppers, spinach, cauliflower, garlic, ginger, corn, cabbage, and more.

Variability Considerations:

- Resolution and Size: Images captured in low to high resolution to simulate smartphone and professional camera inputs.
- Lighting Conditions: Includes bright sunlight, indoor lighting, shadows, and mixed lighting environments.
- Backgrounds: Covers plain backgrounds for controlled experiments and cluttered, natural backgrounds to mimic real-world settings.
- Image Quality: Includes blurred, noisy, and partially occluded images to test robustness.

This diverse composition ensures that the CNN classifier is exposed to a wide spectrum of real-world scenarios, enhancing its reliability and accuracy.

6.2.3 Data Augmentation

To further increase dataset diversity and improve model robustness, various augmentation techniques are applied:

1. Rotation: Random rotations between 0° and 360° to handle different orientations of fruits and vegetables.
2. Flipping: Horizontal and vertical flips to simulate images captured from different angles.
3. Brightness Adjustment: Modifying brightness by ±30% to account for lighting variability.
4. Scaling and Zooming: Random zoom and scaling transformations to mimic varying distances between the camera and objects.
5. Noise Addition: Injecting Gaussian noise and salt-and-pepper noise to simulate low-quality capture conditions and test model resilience.

Significance:

- Augmentation increases the effective dataset size without additional data collection.
- Enhances the generalization capability of the CNN classifier.
- Ensures the GAN-CNN pipeline can handle diverse real-world inputs, improving performance in practical applications.

6.3 Image Preprocessing

Before feeding images into the GAN and CNN modules, preprocessing ensures consistency, normalizes input data, and encodes labels for effective model training. Preprocessing is critical for reducing noise, standardizing inputs, and enhancing model convergence.

1. Resizing:
   - All images are resized to 64×64 pixels, which is suitable for the GAN input while preserving essential details.
   - Standardization allows batch processing and ensures uniformity across datasets collected from multiple sources.
2. Normalization:
   - Pixel values are scaled to a range of 0 to 1 by dividing each pixel value by 255.
   - Normalization helps in faster convergence of both GAN and CNN models and prevents saturation of activation functions.
3. Batching:
   - Images are converted into batches for training efficiency.
   - Batch processing improves GPU utilization, stabilizes gradient updates, and reduces training time.
4. Label Encoding:
   - Class labels (e.g., “apple”, “banana”) are converted into numerical representations using one-hot encoding.
   - One-hot encoding ensures the classifier can handle multi-class predictions and compute categorical loss effectively.

Significance: Proper preprocessing improves model stability, generalization, and performance, especially when handling images of varying resolutions and quality from different sources.

https://doi.org/10.5281/zenodo.20502198

6.4 GAN-Based Image Enhancement

The GAN module enhances low-resolution or noisy images, ensuring that downstream classification achieves higher accuracy. The GAN architecture combines a generator and discriminator trained in an adversarial setup.

6.4.1 Architecture Overview

1. Generator:
   - Learns a mapping from low-resolution input images to high-resolution outputs.
   - Uses residual blocks, upsampling layers, and convolutional layers to reconstruct details and textures.
2. Discriminator:
   - Distinguishes between real high-resolution images and GAN-generated images.
   - Guides the generator to produce visually convincing and realistic outputs.
3. Loss Function:
   - Perceptual Loss: Ensures the generated image maintains high-level features similar to the original.
   - Adversarial Loss: Encourages realism by penalizing outputs that the discriminator can easily identify as fake.
   - Content Loss: Preserves structural and color information to maintain fidelity.

6.4.2 Training GAN

- Dataset: Pairs of low-resolution images and their corresponding high-resolution versions.
- Optimizer: Adam optimizer with a learning rate of 0.0002 for stable convergence.
- Epochs: Typically 100–200, depending on dataset size and convergence behavior.
- Batch Size: 16–32, balancing GPU memory constraints and training efficiency.

Training Strategy:

- Alternating updates for the generator and discriminator to maintain equilibrium.
- Regular checkpointing to save best-performing models based on validation PSNR and SSIM.

6.4.3 Evaluation Metrics for GAN

1. Peak Signal-to-Noise Ratio (PSNR): Measures reconstruction quality; higher PSNR indicates better enhancement.
2. Structural Similarity Index (SSIM): Evaluates similarity between original and enhanced images in terms of luminance, contrast, and structure.
3. Visual Assessment: Side-by-side comparisons of original and enhanced images for qualitative inspection.
4. Significance: Ensures that enhanced images not only look realistic but also improve feature extraction for the classifier.

6.5 CNN-Based Classification

The CNN module classifies the enhanced fruit and vegetable images. It is trained to extract hierarchical features and accurately predict class labels.

Architecture Selection

1. VGG16:
   - Pretrained on ImageNet; used for feature extraction and transfer learning.
   - Benefits: Accelerates training and improves generalization for small datasets.
2. Custom CNN:
   - Tailored for the dataset with three convolutional layers, max-pooling layers, and fully connected layers.
   - Allows experimentation with architecture depth and filter sizes for optimal performance.

6.5.2 Training Process

- Input: Enhanced images from the GAN module.
- Optimizer: Adam optimizer with a learning rate of 0.001.
- Loss Function: Categorical cross-entropy for multi-class classification.
- Epochs: 50–100 depending on convergence and dataset size.
- Batch Size: 32 for stable training and efficient GPU usage.

Training Strategy:

- Model trained on enhanced images to leverage restored details.
- Early stopping and learning rate scheduling applied to prevent overfitting.

6.5.3 Evaluation Metrics

1. Accuracy: Ratio of correctly predicted samples to total predictions.
2. Precision: Measures the proportion of correctly predicted positive observations over total predicted positives.
3. Recall (Sensitivity): Measures the proportion of correctly predicted positive observations over all actual positives.
4. F1-Score: Harmonic mean of precision and recall, balancing false positives and false negatives.
5. Confusion Matrix: Provides class-wise performance analysis, identifying misclassifications and patterns of errors.

Significance:

- Ensures the classifier's performance is robust, reliable, and interpretable.
- By combining GAN enhancement with CNN classification, the system achieves higher accuracy, better feature extraction, and resilience to low-quality images compared to CNN-only approaches.

6.6 Integration of GAN and CNN

The integration of GAN-based enhancement and CNN-based classification is central to the proposed system. The end-to-end pipeline ensures that low-quality input images are transformed into high-resolution outputs suitable for accurate classification, followed by informative presentation to the user. The integration is executed in a sequential, modular manner as follows:

Step 1: Input Low-Resolution Image

- Users upload images of fruits or vegetables via the Streamlit interface.
- Input images may vary in size, resolution, lighting, and background conditions, simulating real-world scenarios.
- Preprocessing is applied at this stage:
    - Resizing: Standardizing images to the input dimension required by the GAN (e.g., 64×64 pixels).
    - Normalization: Scaling pixel values between 0 and 1 to ensure model stability.
- Significance: Standardization and preprocessing ensure compatibility with the GAN module, reducing errors and improving enhancement quality.

Step 2: Image Enhancement Through GAN

- The low-resolution image is passed through the GAN enhancement module, which includes the generator and discriminator.
- Enhancement Process:
    - Generator: Upsamples the image, reconstructs fine details, and improves sharpness and color fidelity.
    - Discriminator: Validates the realism of the generated image, guiding the generator to produce high-quality results.
- Output: A high-resolution, visually enhanced image suitable for feature extraction.
- Significance: Improves the performance of the CNN classifier by providing a clearer and more detailed image, reducing misclassification caused by poor input quality.

Step 3: Classification via CNN

- The enhanced image is fed into the CNN classifier, which has been trained to recognize multiple fruit and vegetable classes.
- Feature Extraction:
    - Convolutional layers extract spatial features such as color patterns, textures, and shapes.
    - Pooling layers reduce dimensionality while retaining essential information.
    - Fully connected layers integrate features for final classification.
- Prediction: The classifier produces the most probable class label along with confidence scores for all categories.
- Significance: Operating on GAN-enhanced images allows the CNN to achieve higher accuracy than models trained on raw, low-quality images.

Step 4: Output and Information Display

- The system outputs:
    1. Predicted class label of the fruit or vegetable.
    2. Original image uploaded by the user.
    3. Enhanced image produced by the GAN module.
    4. Nutritional information, health benefits, and common uses of the identified item.
- Significance: Combines technical results with user-friendly information, making the system practical for agriculture, retail, dietary monitoring, and educational purposes.

Summary:

The integrated GAN-CNN pipeline ensures that the system is robust, accurate, and user-friendly, capable of handling diverse real-world inputs while providing informative outputs.

6.7 Deployment Module

The deployment module ensures that the system is accessible to users in real-time, providing an interactive interface and efficient processing.

6.7.1 User Interface

- Platform: Developed using Streamlit, which allows for a web-based, interactive, and user-friendly interface.
- Features:
    1. Upload Image: Users can upload images of fruits or vegetables in common formats (JPEG, PNG).
    2. Display Original and Enhanced Images: Shows side-by-side comparison for visual verification of GAN enhancement.
    3. Display Predicted Class: Presents the identified fruit or vegetable with a confidence score.
    4. Nutritional Information and Benefits: Provides details about the fruit or vegetable, including health effects, uses, and dietary recommendations.
- Significance: Makes the system accessible to non-technical users, bridging the gap between research models and practical applications.

6.7.2 Real-Time Processing

- Processing Efficiency: Images are processed in milliseconds, ensuring minimal delay between upload and prediction.
- Scalability: The system architecture supports multiple concurrent users, allowing real-time deployment in farms, retail environments, or educational platforms.
- Significance: Ensures smooth user experience, practical usability, and potential integration with mobile or cloud-based applications.

VII. RESULTS & DISCUSSIONS

This section presents the results obtained from the proposed GAN-based image enhancement and CNN-based classification system. The evaluation focuses on the quality of enhanced images, classification accuracy, performance metrics, and comparison with baseline systems. The section also demonstrates the effectiveness of the integrated system through visual and quantitative analysis.

7.1 GAN Image Enhancement Results

7.1.1 Visual Comparison

- Side-by-side comparison of original and enhanced images shows:
    - Sharper edges and contours of fruits and vegetables.
    - Improved color fidelity and texture details.
    - Removal of blurriness and noise from low-quality images.

7.1.2 Quantitative Metrics

Metric	Before Enhancement	After Enhancement
PSNR (dB)	22.5	30.8
SSIM	0.68	0.91
RMSE	18.2	8.5

Observation: GAN enhancement significantly improves the visual quality of input images, as demonstrated by higher PSNR and SSIM values.

7.2 CNN Classification Results

7.2.1 Performance on Raw Images

Metric	Value
Accuracy	87.5%
Precision	85.2%
Recall	84.9%
F1-Score	85.0%

7.2.2 Performance on Enhanced Images

Metric	Value
Accuracy	95.6%
Precision	94.8%
Recall	95.0%
F1-Score	94.9%

Observation: Using GAN-enhanced images significantly improves classification performance. Accuracy increased by 8.1%, demonstrating the effectiveness of the enhancement step.

7.3 Confusion Matrix Analysis

- Confusion matrices show that most misclassifications occur between visually similar fruits or vegetables, e.g., cucumber vs. zucchini or bell pepper vs. capsicum.
- Enhanced images reduce misclassification compared to raw images.
- The diagonal values in the matrix are higher for enhanced images, indicating better class-wise prediction accuracy.

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

7.4 Class-wise Performance

Class	Precision	Recall	F1-Score
Apple	96.2%	95.7%	95.9%
Banana	97.0%	96.5%	96.7%
Carrot	94.5%	95.2%	94.8%
Tomato	95.1%	94.8%	94.9%
Cucumber	93.7%	94.0%	93.8%
...	...	...	...

Observation: Most classes achieve above 94% F1-Score when using GAN-enhanced images.

Interface Demonstration

1. Image Upload: Users upload a low-quality fruit/vegetable image.
2. Original and Enhanced Display: Both images are displayed side by side.
3. Prediction and Nutritional Information:
   - Predicted class label shown.
   - Nutritional benefits, uses, and effects are displayed (e.g., “Apple: Rich in Vitamin C and fiber, supports heart health and digestion”).

7.5 Discussion

- GAN enhancement significantly improves the visual quality of low-resolution images, providing more reliable features for classification.
- CNN classifier performs better on enhanced images due to clearer edges, texture, and color fidelity.
- The dual-stage system demonstrates robustness to variations in lighting, background, and image noise.
- Deployment in a user interface ensures practical usability, allowing real-time image upload, enhancement, and classification.

7.6 Summary

- Enhanced images improve classification metrics by over 8% compared to raw images.
- Confusion matrix analysis confirms reduced misclassification, especially for visually similar classes.
- The system provides both technical performance and practical application via a user-friendly interface, fulfilling the goals of the proposed research.

VIII. CONCLUSION

This research aimed to design an automated fruit and vegetable recognition system that integrates GAN-based image enhancement with CNN-based classification. The motivation arose from challenges in existing systems, which struggle with low-quality images, intra-class variability, and real-world application constraints.
The proposed system addresses these issues by enhancing image quality before classification, thereby improving accuracy, robustness, and usability.

8.1 Key Findings

1. Image Enhancement Using GANs
   - Low-quality and blurred images of fruits and vegetables were significantly improved using Super-Resolution GAN (SRGAN) and Enhanced SRGAN (ESRGAN).
   - Quantitative metrics (PSNR and SSIM) and visual assessment demonstrated that the enhanced images retained critical features necessary for classification.
2. Improved Classification Accuracy
   - CNN classifiers trained on enhanced images outperformed those trained on raw images.
   - Accuracy increased from 87.5% to 95.6%, confirming the effectiveness of the enhancement step.

3. Reduction in Misclassification
   - Confusion matrix analysis indicated fewer errors between visually similar classes.
   - Enhanced features helped the CNN model differentiate between classes with subtle differences.
4. Practical Deployment
   - A Streamlit-based user interface allows real-time image upload, enhancement, and classification.
   - Both the original and enhanced images are displayed side by side, with additional information on nutrition, benefits, and uses of the predicted fruit or vegetable.

8.2 Conclusion Summary

The proposed GAN integrated CNN system successfully enhances low-quality fruit and vegetable images and classifies them with high accuracy. The dual-stage approach demonstrates superior performance compared to traditional CNN-only models, providing a robust, scalable, and deployable solution.

This research contributes to the fields of computer vision, deep learning, and smart agriculture by integrating cutting-edge techniques for real-world applications, bridging the gap between theory and practice.

REFERENCES

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
2. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
3. Russakovsky, O., et al. (2015). *ImageNet Large Scale Visual Recognition Challenge*. Springer.
4. Kaur, H., & Kaur, G. (2019). "Fruit Recognition Using Convolutional Neural Network." *International Journal of Computer Applications*, 178(25), 1–7.
5. Zhang, K., et al. (2019). "Image Super-Resolution Using Generative Adversarial Networks." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(6), 1–12.
6. Wang, X., et al. (2020). "GAN-based Image Enhancement for Medical Image Classification." *Journal of Biomedical Informatics*, 108, 103503.
7. Aamir, M., et al. (2021). "MobileNet-based Real-Time Fruit Recognition." *International Journal of Advanced Computer Science and Applications*, 12(8), 123–130.
8. Singh, R., & Sharma, P. (2020). "Multi-Class Fruit and Vegetable Classification Using Deep Learning." *Journal of AI Research*, 68(3), 112–124.
9. Ledig, C., et al. (2017). "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network." *CVPR 2017*, 105–114.
10. Wang, Z., et al. (2018). "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks." *ECCV Workshops*, 63–79.
11. Goodfellow, I., et al. (2014). "Generative Adversarial Nets." *NeurIPS 2014*, 2672–2680.