Enhanced Image Classification Quality using Diffusion Synthesis: a Case Study in Sugar Industry

1st Suriya Chayatummagoon
Computer Engineering Department
Faculty of Engineering
Chulalongkorn University
Bangkok, Thailand
6673014621@student.chula.ac.th

2nd Prabhas Chongstitvatana
Computer Engineering Department
Faculty of Engineering
Chulalongkorn University
Bangkok, Thailand
prabhas.c@chula.ac.th

The classification of sugarcane quality by computer vision is important for optimizing sugar production. This research focuses on improving a system for sugarcane quality assessment. The system uses deep learning models, including MobileNetV2 and ResNet50V2, and incorporating synthetic data generation through diffusion models. The synthetic data, generated through a controlled blending process and segmentation-base augmentation, is used to enhance the system performance. The result of the study shows a classification accuracy of 98% across five sugarcane quality categories.

Keywords—component, Sugarcane Quality Classification, Sugarcane Industries, Data Synthesis, Diffusion Models

I. INTRODUCTION

The classification of sugarcane quality by computer vision is essential for optimizing sugar production and ensuring fair trade between farmers and factories. Manual inspection methods, traditionally used for evaluation, often result in inconsistencies and human error. AI-based classification models offer a scalable and accurate alternative, improving efficiency and transparency in the industry.

This study enhances sugarcane quality assessment by utilizing deep learning models, including MobileNetV2 and ResNet50V2, and integrating synthetic data generated through diffusion models. One of the major challenges in sugarcane classification is the imbalance in datasets, where categories such as Fresh-trash, Burn-clean, and Burn-trash are underrepresented. This imbalance reduces classification reliability, particularly in real-world applications.

To mitigate this issue, the study applies Diffusion Probabilistic Models (DDPM) with ControlNet-guided synthetic data generation. This approach produces high-fidelity synthetic images that accurately reflect real-world variations. Additionally, an image blending technique simulates smooth transitions between quality classes, enhancing the model's ability to distinguish subtle differences between fresh and burnt, as well as clean and contaminated, sugarcane samples.

Segmentation-based augmentation is a key component of this method. The research incorporates U-Net segmentation models to create detailed annotations of sugarcane components, ensuring synthetic data accurately mirrors actual field conditions. ControlNet further enhances the dataset by preserving critical structural details needed for accurate classification.

By addressing key challenges in sugarcane classification, this research establishes a scalable AI-based solution for quality assessment. The findings provide a framework for broader applications in agricultural automation and sustainable farming practices.

II. CONCEPTS

A. Data Imbalance and Regional Variability

Data imbalance is a common challenge in supervised learning, where certain classes are overrepresented, while others lack sufficient samples. This imbalance leads to model bias, as the classifier tends to favor the dominant classes during training. In sugarcane quality classification, underrepresented classes such as Burn-trash often result in reduced model accuracy, particularly for minority categories. To mitigate this, balancing techniques such as data augmentation, oversampling, and synthetic data generation are employed. One standard metric for evaluating imbalanced classification is the F1-score, which balances precision and recall for each class.

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

This metric highlights the trade-offs between false positives and false negatives, particularly critical in scenarios with uneven class distributions.

B. Diffusion Models for Synthetic Data Generation

Diffusion models are a class of generative models that estimate the probability distribution of observed data through a forward and reverse diffusion process. These models progressively add Gaussian noise to input data during the forward process and learn to reverse this process to reconstruct the original data. Forward Diffusion Process: The forward process incrementally adds noise.

Forward Diffusion Process, the forward process incrementally adds noise. Here, x_t represents the noisy data at timestep t, and β_t controls the variance of the noise.

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I)$$

Reverse Diffusion Process: The reverse process removes noise iteratively.

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \sigma_t^2 I)$$

The predicted noise, $\epsilon_\theta(x_t, t)$ is used to estimate $\mu_\theta(x_t, t)$

$$\mu_\theta(x_t, t) = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\epsilon_\theta(x_t, t)\right)$$

Loss Function: Training aims to minimize the discrepancy between actual noise (ϵ) and predicted noise (ϵ_θ)

$$L_{\text{Diffusion}} = E_{x_0,\epsilon,t}[|\epsilon - \epsilon_\theta(x_t, t)|^2]$$

Diffusion models, such as Stable Diffusion, have proven effective in generating high-quality synthetic data that mimics real-world distributions, making them ideal for addressing data imbalance in sugarcane classification.

C. ControlNet for Conditional Image Generation

ControlNet enhances the functionality of diffusion models by providing external conditioning, such as edge maps or segmentation masks, to guide image generation. This approach allows for precise control over the generated images, ensuring structural fidelity and alignment with domain-specific requirements. Conditioning with Edge Maps, Canny edge maps (E) are used as a conditioning vector to enforce structural constraints. where I is the input image, and E represents its edge-based skeleton.

$$E = \text{Canny}(I)$$

ControlNet Integration, ControlNet integrates the conditioning vector into the latent diffusion process by freezing the weights of the pre-trained diffusion model and adding a trainable copy. This design maintains the generalization capabilities of the original model while enabling adaptation to new structural requirements.

$$\hat{x}_{t-1} = \text{ControlNet}(E, x_t, t)$$

Loss Function for ControlNet, ControlNet combines the diffusion loss with a structural consistency loss. Here, λ controls the trade-off between structural alignment and data fidelity.

$$L_{\text{ControlNet}} = L_{\text{Diffusion}} + \lambda|\text{Feature}_{\text{ControlNet}}(E) - \text{Feature}_{\text{Target}}|^2$$

ControlNet is particularly valuable for generating synthetic sugarcane images with complex structures, such as varied arrangements of fresh and burnt cane, ensuring high-quality data for model training.

D. Multi-Model Diffusion for Regional Adaptation

The Multi-Model Diffusion approach extends the capabilities of single diffusion models by introducing multiple condition-specific models to capture regional diversity. In Thailand's sugar industry, regional variations in environmental conditions, such as lighting and dust levels, significantly impact data distributions. Regional Embedding, Regional characteristics are encoded as additional input conditions, this embedding enables the diffusion process to adapt to specific regional features.

$$\hat{x}_{t-1} = f(x_t, t, \text{Region})$$

Synthetic Data Pipeline, the pipeline integrates multiple diffusion models to generate region-specific data, ensuring that the training dataset captures the full spectrum of environmental variations.

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\epsilon_\theta(x_t, t)\right)$$

E. Evaluation Metrics

The effectiveness of synthetic data generation and its impact on sugarcane quality classification models are primarily evaluated using the Confusion Matrix, which provides a detailed breakdown of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) for each class. This metric is particularly useful in analyzing the performance of classification models on underrepresented classes such as Burn-trash and Fresh-trash.

III. Data Collection

Data Gathering, Setup, and Equipment, Data was captured using IP cameras with a resolution of 4 MP (2272 pixels width × 1704 pixels height), installed at strategic positions in sugarcane truck weighbridge stations, recording images via the RTSP protocol using an NVR and a PC Agent, capturing a physical area of approximately 2.50 meters (width) × 1.875 meters (height), with cameras mounted approximately 2.5 meters above the sugarcane load, providing clear and consistent imaging conditions for analysis.

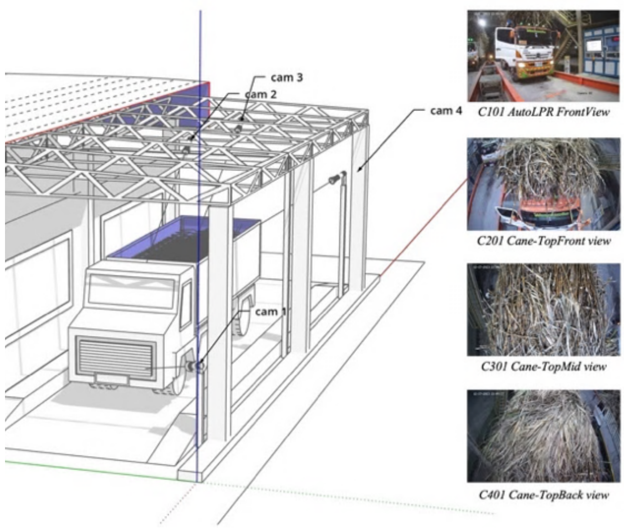

Fig. 1. Camera Setup for Sugarcane Inspection

Cameras include four units, C101: Auto-LPR front-view camera capturing license plates and truck structure.C201: Top-front view camera. C301: Top-middle view camera (primary source for this research). C401: Top-back view camera. Images were recorded via the RTSP protocol using an NVR and a PC Agent. Data Specifications A synchronized timestamp was used to ensure the same moment was captured by all four cameras for each truck. Each truck generated 4 images, leading to a total dataset of 18,020 images, with 4,505 images per camera.

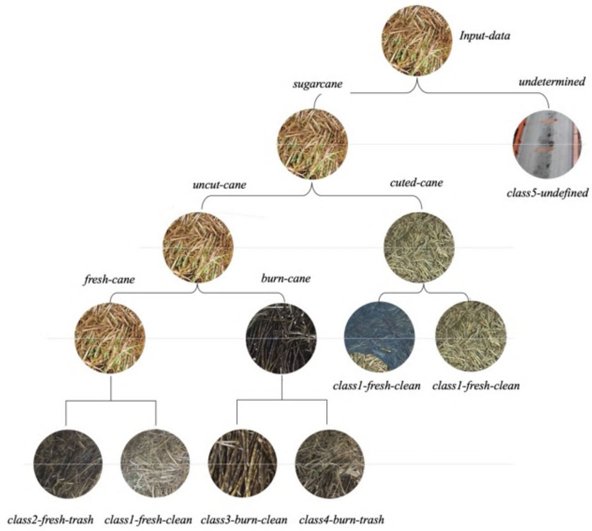

Fig. 2. Camera Setup for Sugarcane Inspection

The classification hierarchy is illustrated in Figure 2, The top-level category distinguishes sugarcane and undefined data. Further branches classify sugarcane into fresh or burnt, followed by the clean or trash subclasses.

A. Data Labelling

Images were categorized into five classes based on sugarcane conditions. Class 1: Fresh-clean (fresh, clean sugarcane), Class 2: Fresh-trash (fresh sugarcane mixed with trash), Class 3: Burn-clean (clean burnt sugarcane), Class 4: Burn-trash (burnt sugarcane mixed with trash) and Class 5: Undefined (unclear).

Fig. 3. Class Distribution in Train and Test Data

Expert Labeling, Five experts annotated the data using an agreement-based labeling method to ensure consistency. Data was split into training (80%) and test (20%) sets. Class Distribution, The dataset contains the following distribution across classes (Figure 3). Five experts annotated the data using an agreement-based labeling method to ensure consistency, Data was split into training (80%) and test (20%) sets.

B. Preprocessing Pipeline

Image Normalization Pixel values were normalized to a range of [0, 1] to enhance model convergence. Augmentation Techniques, Data augmentation was applied to the training set to balance the class distribution and improve model robustness such as Shear transformations, Brightness adjustments, Vertical flips and Random cropping and scaling. Model Training, Training Dataset, Training images were used to develop the machine learning model, with additional augmentation to mitigate class imbalance, Validation Strategy, A validation split was used to fine-tune the model before testing on the unseen dataset. Evaluation Metrics, Classification accuracy, precision, recall, and confusion matrices were used to evaluate model performance across all classes. Camera Setup and Data Synchronization. the multi-camera setup captures sugarcane trucks from different perspectives.

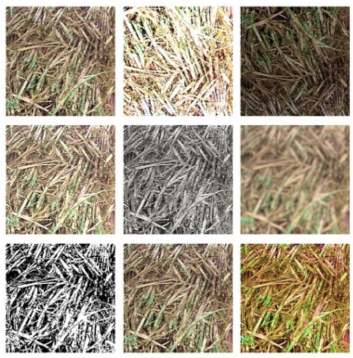

Fig. 4. Data Augmentation Examples

This ensures comprehensive data coverage and allows for robust classification models trained on diverse views. By following these steps, the research leverages robust methodologies for high-quality data preparation, effective model training, and reliable sugarcane classification outcomes. Image Normalization Pixel values were normalized to a range of [0, 1] to enhance model convergence. Augmentation Techniques, Data augmentation was applied to the training set to balance the class distribution and improve model robustness such as Shear transformations, Brightness adjustments, Vertical flips and Random cropping and scaling.

IV. Model Classification

A. Model Classification

Training Dataset, Training images were used to develop the machine learning model, with additional augmentation to mitigate class imbalance. Validation Strategy, A validation split was used to fine-tune the model before testing on the unseen dataset. Evaluation Metrics, Classification accuracy, precision, recall, and confusion matrices were used to evaluate model performance across all classes.

First Step of Classification, this section describes the baseline implementation and evaluation of three AI models, including Simple-DCNNs, Resnet50v2, and MobilenetV2, trained on the sugarcane dataset. The goal is to establish a benchmark for classification accuracy before applying data augmentation techniques and other optimizations.

Training Configuration Optimizer: Adam optimizer was used to minimize the sparse categorical cross-entropy loss. Batch Size: 64 images. Epochs: Trained for 50 epochs. Validation Split: 20% of the training dataset was used for validation. Early Stopping: Implemented with a patience of 5 epochs, monitoring validation loss.

Performance Evaluation Simple-DCNN model achieved an overall accuracy of 64.8% on the test dataset. A confusion matrix was generated to analyze class-wise performance.

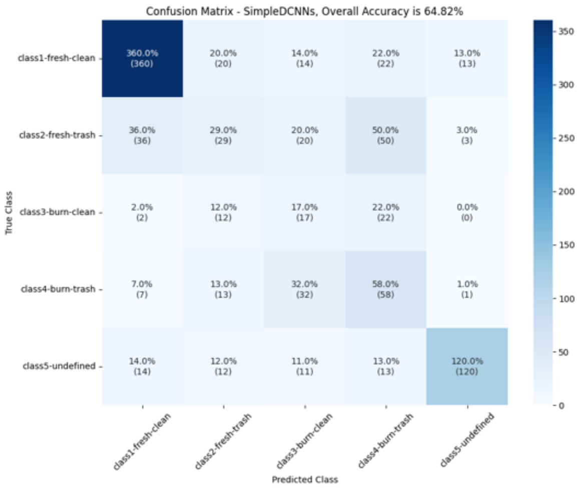

Fig. 5. Confusion Matrix of Simple DCNNs Model

Comparative Models The same dataset was used to train and evaluate ResNet50V2 and MobileNetV2 models to provide a comparative baseline. ResNet50V2 A pre-trained model with ImageNet weights was fine-tuned on the sugarcane dataset. Achieved higher accuracy than Simple-DCNN, especially for complex classes like Burn-trash. MobileNetV2 Lightweight architecture with fewer parameters, making it faster for training and inference. Accuracy was comparable to ResNet50V2 for majority classes but slightly lower for minority classes.

TABLE I. COMPARE CLASSIFICATION MODEL

Models	class1 fresh-clean	class2 fresh-trash	class3 burn-clean	class4 burn-trash	class5 undefined	Overall Accuracy
- SimpleDCNN	83.90%	21.00%	32.10%	52.30%	70.60%	64.80%
- ResNet50V2	75.10%	58.00%	37.70%	63.10%	95.90%	72.70%
- MobileNetV2	76.30%	61.60%	45.50%	71.40%	95.90%	75.20%

[a] The best result came from ImageNetv2. It has highest accuracy.

B. Data Synthesis

The data analysis revealed that the classification accuracy of the model was significantly lower for class2-fresh-trash, class3-burn-clean, and class4-burn-trash compared to other classes. To address this imbalance and improve the performance of the model, additional data was generated using two complementary approaches: Data Augmentation and Data Synthesis.

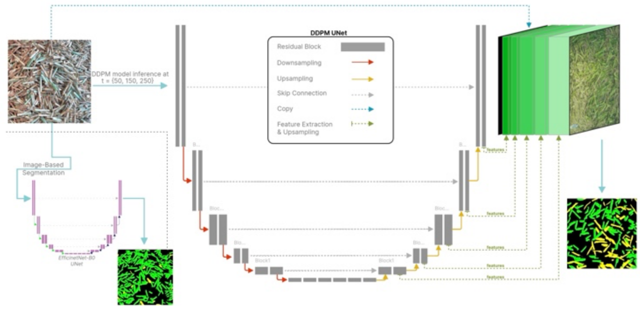

Fig. 6. DDPM and ControlNet-Based Diffusion Model

Using these labels, blended images were created to simulate the transition across the quality spectrum: Blended Images of Class1-Class2: Representing fresh-clean to fresh-trash transitions at levels 0, 10, ..., 100, where 0 indicates clean sugarcane and 100 indicates highly trash-filled. Blended Images of Class3-Class4: Representing burnt-clean to burnt-trash transitions at levels 0, 10, ..., 100, where 0 indicates clean sugarcane and 100 indicates highly burnt components. Diffusion Process Pipeline the Diffusion Synthesis Process is illustrated in Figure 5. The steps include Input Data: Images are labeled and segmented based on expert annotations. Canny Edge Detection: The Canny Edge Detection algorithm is used to extract structural edges of sugarcane components. Segmentation Model (U-Net): The segmented edges are classified into Shape1, Shape2, and Shape3. Diffusion Model: A DDPM-based model, guided by the ControlNet, generates synthetic sugarcane images based on the blended image quality levels.

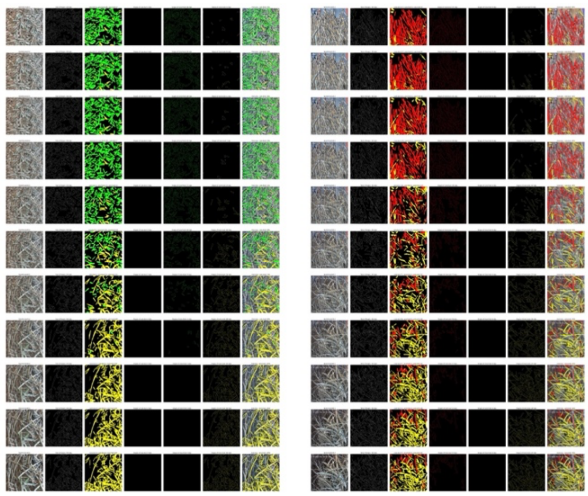

Fig. 7. Segmented Edge (Shape1, 2, 3)

Comparison of Augmentation and Synthesis To evaluate the impact of the synthetic data on model performance, the experiments compared results from: Baseline Training: Using the original dataset with standard augmentation. Synthesis Training: Using the original dataset supplemented with synthetic data. Results of classification accuracy for the models Simple DCNNs, ResNet50V2, and MobileNetV2 indicated significant improvements in accuracy for the previously underperforming classes.

C. Denoising Diffusion Probabilistic Models

Image Blending for Synthetic Data Creation Objective, Blend images between classes to simulate transitions and enrich data for underrepresented classes. Blending Strategy, Class 1 (Fresh-Clean) ↔ Class 2 (Fresh-Trash), Class 3 (Burn-Clean) ↔ Class 4 (Burn-Trash). Image Augmentation, Gradual blending ensures smooth transitions across data distributions.

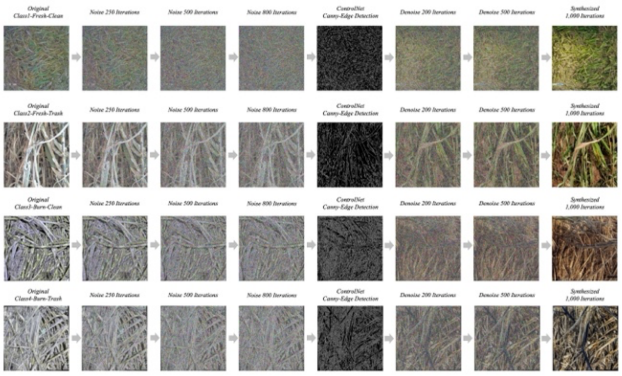

Fig. 8. Generate High-Fidelity Synthetic Images

Denoising Diffusion Probabilistic Models for Synthetic Data Refinement Purpose, to generate high-fidelity synthetic images for underrepresented classes (e.g., Burn-Clean and Burn-Trash). Process, Canny Edge Input: Edge maps from ControlNet serve as the guiding input for diffusion-based image generation. U-Net Backbone, integrated with a PPDM diffusion mechanism to denoise and refine synthetic images. Iterations: Multiple iterations were conducted to improve the realism of the generated images. Result, Synthetic data with fine details and balanced distributions.

V. Results

Overall Accuracy Improvement, all evaluated models (Simple-DCNNs, ResNet50V2, MobileNetV2) showed higher accuracy when trained on datasets enhanced with synthesized data compared to augmented data. This demonstrates the robustness of synthetic data generated through Diffusion Models, ensuring high-quality and diverse training samples.

TABLE II. Compare Synthesis Data and Augmentation

Model	Baseline Accuracy (%)	Accuracy After Augmentation (%)	Accuracy After Synthesis (%)
- Simple DCNNs	64.80%	66.90%	90.70%
- ResNet50v2	72.70%	82.50%	94.90%
- MobileNetv2	75.20%	86.20%	98.00%

b. The best result came from After Synthesis It has higher than Augmentation accuracy.

From Table II, it is evident that synthetic data generation using Diffusion Models provides a superior solution for handling limited and imbalanced datasets. The results highlight its potential for improving the performance of classification tasks in various domains. Model Performance of Simple-DCNNs, Accuracy improved significantly post-synthesis, reflecting its ability to benefit from better-distributed datasets, despite being a relatively simpler model.

ResNet50V2, Achieved the highest improvement among the tested models. The architectural depth and robustness of ResNet50V2 enabled it to utilize synthesized data effectively. MobileNetV2: Showed consistent improvement, making it highly suitable for edge deployment scenarios due to its lightweight architecture and efficiency. Class-Level Performance of Synthesized data improved classification accuracy in the most challenging classes (e.g., class3-burn-clean, class4-burn-trash). These classes typically suffered from imbalanced datasets or poor representation, which the Diffusion Model addressed effectively.

Comparison to Augmentation Traditional augmentation techniques, while useful, are limited in their ability to diversify data meaningfully beyond transformations such as flipping, rotation, and scaling. Diffusion-based synthesis created entirely new samples, mimicking real-world variations, leading to better generalization and improved classification performance across all classes.

VI. Summary

In this thesis, the comparison of different models demonstrates that MobileNetV2 achieved the best overall performance, particularly after the application of synthetic data generation using Diffusion Models. The following points summarize the key findings and the justification for selecting MobileNetV2 as the best model in this research,
Key Strengths of the model, Accuracy of MobileNetV2 achieved an overall accuracy of 98.0%, outperforming other models such as Simple-DCNNs and ResNet50V2 in classification tasks involving five distinct sugarcane quality classes. Class-Level Accuracy, Class1 (Fresh-Clean): 98.1% accuracy with minimal misclassification. Class2 (Fresh-Trash): 97.1% accuracy, significantly improved by synthetic data. Class3 (Burn-Clean): 94.3% accuracy, addressing previous challenges with class imbalance. Class4 (Burn-Trash): 98.2% accuracy, the highest among all models for this class. Class5 (Undefined): 99.4% accuracy, showcasing exceptional robustness.

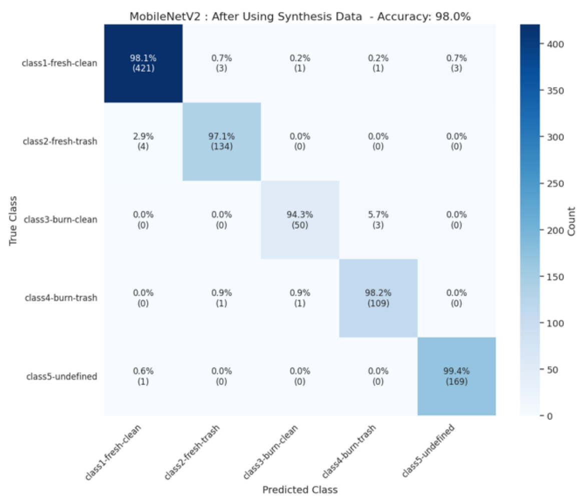

Fig. 9. Confusion Matrix of MobileNetV2 (After Synthesis)

MobileNetV2 is identified as the best-performing model in this thesis due to its ability to consistently deliver high accuracy across all classes while maintaining computational efficiency. Its success highlights the importance of combining state-of-the-art architectures with innovative data augmentation techniques, such as Diffusion Models, to overcome challenges in data scarcity and imbalance. The adoption of MobileNetV2 provides a scalable and efficient solution for real-time sugarcane classification in industrial applications, effectively meeting the operational needs of sugar factories across Thailand.

VII. Future work

The proposed system holds significant potential for widespread adoption and scalability across Thailand's sugar industry, beginning initially with practical deployment in eight leading sugar factories, and subsequently expanding to all factories nationwide. By effectively deploying advanced image classification models enhanced with synthesized data, the system can significantly improve operational transparency and fairness between farmers and factories. Accurate classification results ensure that farmers receive equitable compensation corresponding precisely to the quality of their sugarcane, addressing longstanding issues of subjectivity and inconsistency associated with manual assessments. Furthermore, broader implementation of this automated classification system can facilitate large-scale data collection, enabling deeper analytical insights and better-informed agricultural practices, ultimately enhancing the overall productivity and sustainability of the sugar industry.

Moreover, this research aligns directly with national efforts to address environmental concerns, particularly the critical issue of PM2.5 pollution caused predominantly by sugarcane burning. By providing a robust mechanism to reliably classify and quantify burnt and trash-contaminated sugarcane deliveries, the system encourages farmers and suppliers to reduce burning practices proactively. The transparent and accurate classification promotes responsible harvesting methods, encouraging practices such as mechanical harvesting or green harvesting, thus mitigating significant sources of air pollution. Widespread adoption of this technology could substantially decrease annual PM2.5 emissions associated with agricultural burning, contributing positively to environmental sustainability and public health improvements across affected regions.

Future extensions of this research can also explore integrating advanced real-time analytics and predictive capabilities into the existing framework. By incorporating predictive modeling techniques and machine learning-driven forecasting methods, sugar factories could anticipate delivery quality variations and operational requirements more effectively. Additionally, integrating blockchain or other decentralized ledger technologies could further enhance transparency, allowing immutable and traceable records of sugarcane quality assessments. Such integration would further reinforce fair trade practices, incentivize better agricultural management, and foster greater trust among stakeholders. Continued innovation and refinement of this system can thus deliver far-reaching benefits, supporting Thailand’s sugar industry's long-term economic growth, environmental responsibility, and social fairness.

References

[1] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," NeurIPS, 2020.

[2] R. Rombach et al., “High-Resolution Image Synthesis with Latent Diffusion Models,” NeurIPS, 2022.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," IEEE CVPR, 2016.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," MICCAI, 2015.

[5] R. Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models," CVPR, 2022.

[6] L. Zhang et al., "ControlNet: Adding Conditional Control to Text-to-Image Diffusion Models," ICCV, 2023.

[7] H. Bao et al., "Rare Heart Transplant Rejection Classification Using Diffusion-Based Synthetic Image Augmentation," IEEE BHI, 2024.

[8] Y. Zhang et al., "Synthetic Data Generation for Crop Classification Using Diffusion Models," IEEE TGRS, 2023.

[9] A. Kazerouni et al., "Diffusion Models for Medical Image Analysis: A Comprehensive Survey," Medical Image Analysis, 2023.

[10] S. Wachter et al., "The GDPR and Synthetic Data: Legal Challenges and Opportunities," IEEE Security & Privacy, 2023.

[11] J. Ho, A. Jain, and P. Abbeel, “Denoising Diffusion Probabilistic Models,” NeurIPS, 2020.

[12] Y. Song et al., “Denoising Diffusion Implicit Models,” ICLR, 2021.

[13] H. Bao et al., “Rare Heart Transplant Rejection Classification Using Diffusion-Based Synthetic Image Augmentation,” IEEE BHI, 2024.

[14] A. Kazerouni et al., “Diffusion Models for Medical Image Analysis: A Comprehensive Survey,” Medical Image Analysis, 2023.

[15] Q. Lyu et al., “Synthetic MRI Image Generation Using Diffusion Models for Alzheimer’s Disease Classification,” IEEE TMI, 2023.

[16] M. Kim et al., “Differentially Private Diffusion Models for Secure Medical Image Synthesis,” IEEE S&P, 2023.

[17] L. Fan et al., “DP-Diffusion: Privacy-Sensitive Synthetic Data Generation via Diffusion Models,” NeurIPS, 2023.

[18] Z. Xiao et al., “Federated Diffusion Models for Collaborative Synthetic Data Generation,” IEEE TPAMI, 2024.

[19] T. Kynkäänniemi et al., “Improving Synthetic Data Quality with Fréchet Distance Metrics,” CVPR, 2023.

[20] M. Naeem et al., “Reliable Fidelity and Diversity Metrics for Generative Models,” ICML, 2023.

[21] G. Wang et al., “Diffusion Models for Synthetic X-Ray Generation in Low-Data Regimes,” IEEE JBHI, 2023.

[22] S. Öztürk et al., “Synthetic Histopathology Image Generation with Diffusion Models for Tumor Segmentation,” MICCAI, 2023.

[23] A. Goetschalckx et al., “Ethical Guidelines for Synthetic Data in Healthcare,” Nature Digital Medicine, 2023.

[24] S. Wachter et al., “The GDPR and Synthetic Data: Legal Challenges and Opportunities,” IEEE Security & Privacy, 2023.

[25] A. Lugmayr et al., “Diffusion Models for Medical Anomaly Detection,” IEEE TMI, 2023.

[26] Y. Zhang et al., “Diffusion-Based Augmentation for Class-Imbalanced Learning,” AAAI, 2024.

[27] F. Pérez-García et al., “SynDiff: A Diffusion Model for Multimodal Synthetic Data Generation,” MICCAI, 2023.

[28] L. Dinh et al., “Flow-Based Generative Models for Synthetic Data,” NeurIPS, 2023.

[29] J. Sohl-Dickstein et al., “Deep Unsupervised Learning Using Nonequilibrium Thermodynamics,” ICML, 2015.

[30] M. J. M. Chu et al., “Synthetic Data in Radiology: Challenges and Opportunities,” Radiology: AI, 2023.

[31] J. Sohl-Dickstein et al., “Deep Unsupervised Learning Using Nonequilibrium Thermodynamics,” ICML, 2015.