Neuroplastic Modular Framework: Cross-Domain Image Classification of Garbage and Industrial Surfaces

Debojyoti Ghosh [1a], Soumya K Ghosh[b], Adrijit Goswami[a]

[a]*Department of Mathematics, Indian Institute of Technology Kharagpur, Kharagpur, 721302, West Bengal, India*
[b]*Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, Kharagpur, 721302, West Bengal, India*

Abstract

Efficient and accurate classification of waste and industrial surface defects is essential for ensuring sustainable waste management and maintaining high standards in quality control. This paper introduces the *Neuroplastic Modular Classifier*, a novel hybrid architecture designed for robust and adaptive image classification in dynamic environments. The model combines a ResNet-50 backbone for localized feature extraction with a Vision Transformer (ViT) to capture global semantic context. Additionally, FAISS-based similarity retrieval is incorporated to provide a memory like reference to previously encountered data, enriching the model's feature space. A key innovation of our architecture is neuroplastic modular design composed of expandable, learnable blocks that dynamically grow during training when performance plateaus. Inspired by biological learning systems, this mechanism allows the model to adapt to data complexity over time, improving generalization. Beyond garbage classification, we validate the model on the Kolektor Surface Defect Dataset 2 (KolektorSDD2), which involves industrial defect detection on metal surfaces. Experimental results across domains show that the proposed architecture outperforms traditional static models in both accuracy and adaptability. The *Neuroplastic Modular Classifier* offers a scalable, high-performance solution for real-world image classification, with strong applicability in both environmental and industrial domains.

Keywords: Neuroplastic Modular Classifier, Garbage Classification, Vision Transformer (ViT), ResNet-50, FAISS Similarity Retrieval, Dynamic Neural Architecture.

1. Introduction

The escalating challenges of waste management, driven by rapid urbanization, rising consumption, and environmental degradation, have made efficient waste handling a global imperative. Accurate classification of waste is a fundamental step toward optimizing recycling processes and reducing environmental impact. However, traditional manual sorting methods remain labor intensive, inconsistent, and inefficient, highlighting the need for intelligent and automated classification systems. Recent progress in deep learning, particularly

---

[1]Corresponding Author: debojyoti07.dg@gmail.com

through Convolutional Neural Networks, has substantially improved image-based classification by enabling the extraction of meaningful and high-level features from complex visual data. Despite these advancements, waste classification continues to face significant challenges due to intra class variability, inter class similarity, and background interference. Moreover, maintaining high classification accuracy while ensuring computational efficiency remains a critical barrier to practical deployment.

Efficient and reliable garbage classification has become increasingly vital in the pursuit of sustainable waste management and environmental conservation. Despite advancements in image classification, real-world waste datasets often suffer from high variability, class imbalance, and noisy annotations, posing persistent challenges for conventional models. In response to these issues, we present the *Neuroplastic Modular Classifier*, a versatile hybrid deep learning framework engineered to address the complexities of garbage classification and surface defect detection in industrial applications. The proposed model combines the localized representational power of a ResNet50 backbone with the contextual reasoning capabilities of a Vision Transformer, enabling it to capture both fine-grained and global semantic cues. To further enhance its discriminative ability, the architecture incorporates a FAISS-based similarity retrieval module, which provides context-aware memory referencing by retrieving previously seen instances from the feature space. What distinguishes our approach is its biologically inspired neuroplastic modular structure. Unlike fixed-capacity models, the proposed network includes dynamically expandable, learnable components that adaptively grow during training in response to stagnating performance. This design enables the model to scale its representational capacity in accordance with data complexity, facilitating more effective learning and better generalization. The combination of adaptive structural growth, multi-level feature fusion, and memory-driven retrieval allows our model to excel in challenging classification scenarios. Extensive experiments conducted on a real-world garbage image dataset validate the effectiveness of our method, demonstrating consistent improvements over static baselines in both accuracy and adaptability. Our architecture offers a practical and scalable solution for visual classification tasks in dynamic and unpredictable environments.

The accelerating pace of urban development and industrial activity has led to a substantial increase in waste production, creating significant challenges in effective waste management. Accurate and efficient waste classification plays a crucial role in enhancing recycling processes and promoting environmental sustainability. Conventional manual sorting approaches are not only labor-intensive but also prone to errors and health risks [1]. Consequently, automated solutions, particularly those leveraging deep learning, have emerged as promising alternatives to traditional methods.

Deep learning models, especially Convolutional Neural Networks (CNNs), have demonstrated exceptional capabilities in image analysis tasks, enabling automatic identification of various waste categories without human supervision [2]. By learning rich hierarchical features from large volumes of image data, these networks can distinguish among recyclables, hazardous materials, and organic waste with high accuracy [3].

Nonetheless, applying deep learning techniques to garbage classification presents several challenges. One major obstacle is the scarcity of large, diverse, and well-annotated datasets required for effective model training [4]. Waste images often exhibit significant variability in visual properties such as color, texture, and size, making consistent classification difficult [5]. In response, methods like data augmentation, transfer learning, and the generation of

synthetic samples are commonly employed to expand training datasets and enhance model robustness [6].

Furthermore, the heterogeneous nature of waste materials adds complexity to classification systems. Categories like plastics, metals, paper, and organic substances each have distinct visual characteristics [7]. The presence of mixed or overlapping waste further increases the difficulty of accurate labeling. To address this, hybrid architectures that integrate multiple deep learning paradigms or combine sensor data have been explored [8]. Such approaches leverage the complementary strengths of CNNs and other network types, including Recurrent Neural Networks (RNNs), to capture spatial and temporal patterns more effectively [9].

Another limitation is the computational demands of deep models, which can be prohibitive in resource-constrained environments like smart bins or embedded systems [10]. Model compression strategies, such as pruning, quantization, and the adoption of lightweight networks like MobileNet, have been proposed to reduce computational requirements while maintaining performance [11]. Additionally, hybrid approaches combining deep learning with classical machine learning methods, such as support vector machines, have been investigated to improve efficiency [1].

The convergence of deep learning with Internet of Things (IoT) technologies is also gaining traction in intelligent waste management applications. Smart bins equipped with sensors and onboard models can classify waste in real time as it is deposited, enabling automated sorting and actionable feedback [12]. This integration paves the way for adaptive systems capable of optimizing collection and disposal workflows dynamically.

Despite these advances, challenges remain in scaling such systems for widespread deployment. For example, models must be capable of adapting to novel waste categories and evolving waste patterns over time. Transfer learning has emerged as a practical solution, enabling efficient fine-tuning of pretrained networks on new datasets with minimal retraining effort [13].

The remainder of this paper is structured as follows. Section 2 presents a comprehensive review of related work and foundational research. Section 3 outlines the dataset characteristics and describes the preprocessing techniques employed. In Section 4, we detail the proposed model architecture, highlighting the hybrid integration of ResNet-50 and Vision Transformer (ViT), along with the neuroplastic modular design. Section 5 describes the experimental setup used to train and evaluate the model. Section 6 provides an in-depth analysis of the experimental results. Section 7 presents an industrial application evaluation based on the KolektorSDD2 dataset, demonstrating the model's effectiveness in real-world defect detection scenarios. Finally, Section 8 concludes the paper and outlines potential directions for future research.

2. Literature Survey

Recent years have witnessed remarkable progress in using deep learning for waste classification. Research in this area has explored various modeling strategies, including advanced convolutional architectures, integration of learning techniques, model efficiency improvements, and the fusion of multiple data sources. These developments are further supported

by innovations in industrial automation and smart environmental technologies, highlighting the need for adaptable and reliable classification systems across diverse applications.

2.1. Deep CNN-Based Models

Deep convolutional neural networks (CNNs) have become the cornerstone of many waste classification systems. Prior work includes the development of deep vision architectures to automate recycling workflows [1]. Data augmentation and transfer learning were applied to boost the performance of CNNs in domain-specific datasets, such as decoration waste [2]. A two-stage deep learning pipeline was introduced to enhance robustness in real-world applications [3], while feature fusion strategies were adopted to improve human-robot interaction efficiency in smart recycling [4]. Several studies integrated pre-trained CNN backbones such as VGGNet [5], or lightweight backbones like ShuffleNet v2 for mobile deployments [6]. Real-time performance was a focus of the MRS-YOLO model [7], and SwinConvNeXt [15] fused transformer and CNN architectures for accurate and efficient garbage image classification.

2.2. Hybrid and Transfer Learning Models

Several studies have explored the use of hybrid approaches and transfer learning methods to enhance waste classification. For instance, optimized convolutional neural networks have been developed to achieve both speed and accuracy, which are essential for real-time recycling applications [9]. Other research has combined convolutional neural networks with deep reinforcement learning techniques to improve the efficiency of sorting recyclable materials [10]. To tackle the challenges posed by occluded objects, attention mechanisms based on depth-wise separable convolutions have been introduced to boost the performance of CNNs [11]. In complex scenarios, integrating CNNs with autoencoders has shown to improve classification outcomes by extracting richer feature representations [1]. Reinforcement learning has also been utilized to enable models to adapt dynamically as new waste data becomes available [12]. Moreover, transfer learning paired with lightweight neural architectures has been applied to develop mobile-friendly waste classification systems [13]. Finally, scalable deep learning frameworks integrated with Internet of Things (IoT) devices have been proposed to enhance waste management accuracy and responsiveness [14].

2.3. Optimization and Lightweight Architectures

Efforts to reduce model complexity without compromising accuracy are prominent in the literature. Several works apply depth-wise separable convolutions and pruning techniques for real-time classification on embedded devices [16]. Reinforcement learning-based optimization methods have also been proposed for dynamically adjusting classification strategies [17]. Other studies introduced edge computing architectures that combine local inference with cloud-based updates for high-throughput waste sorting [18]. Models like optimized CNN variants have been developed to minimize false positives and improve performance under noisy conditions [19].

2.4. Emerging Trends in Hazard and Industrial Waste Management

Beyond traditional waste sorting, recent studies have expanded to hazard classification and intelligent industrial systems. For instance, a deep learning framework incorporating

4

multifractal analysis was proposed for hazard severity assessment [22], while a Fourier-enhanced gray model combined with CNNs improved classification in dynamic environments [21]. High-dimensional classification has benefited from QPSO-based fuzzy clustering [20], and CNNs have been used to classify and quantify recycled aggregates in real time [24]. Furthermore, federated learning have been used to improve traceability and real-time performance in smart city waste management systems [23].

2.5. Domain-Inspired Modular Architectures

Recent domain-specific models illustrate a growing trend toward modularity and attention mechanisms for efficient classification. [25] demonstrated that handcrafted feature extraction can outperform traditional CNNs for construction and demolition waste. [26] proposed GD-YOLO, which improves YOLOv10n with slicing-based attention for high-speed edge deployment [27] introduced GMMNet for rice grain classification using Gaussian matrix self-attention and multi-scale modules. These studies collectively highlight the increasing relevance of modular and neuroplastic architectures in vision-based classification.

Numerous studies have highlighted key challenges and outlined prospective directions for advancing deep learning approaches in garbage classification. For example, [28] introduced a method that integrates Convolutional Neural Networks (CNNs) with deep reinforcement learning to improve classification performance in dynamic environments. Similarly, [29] proposed a hybrid architecture that combines CNNs and Recurrent Neural Networks (RNNs), effectively capturing both spatial and temporal information to enhance classification accuracy. More recently, researchers have focused on developing innovative techniques aimed at increasing the robustness and computational efficiency of these systems. In [30], a combination of CNNs and random forest classifiers was employed within smart recycling frameworks to improve sorting effectiveness. Additionally, [31] presented a novel deep CNN design incorporating feature pyramid networks, enabling the model to better manage complex and heterogeneous waste objects.

Model	CNN	Transformer	Adaptive	FAISS
[7]	✓	✗	✗	✗
[10]	✓	✗	✗	✗
[15]	✓	✓	✗	✗
[26]	✓	✗	✗	✗
[32]	✓	✗	✗	✓
Proposed Model	✓	✓	✓	✓

Table 1: Comparison of architectures and strategies in recent waste classification studies

A comparative analysis of recent state-of-the-art methods (see Table 1) clearly illustrates that while prior work has explored various combinations of convolutional backbones, transformers, attention mechanisms, and multimodal inputs, none have simultaneously integrated adaptive neuroplastic growth and memory retrieval modules in a unified framework.

The present work significantly extends the research by Ghosh et al. [32], who demonstrated the effectiveness of deep convolutional feature extraction and FAISS-based similarity retrieval for accurate garbage classification. While their approach leveraged ResNet-152 and attention mechanisms to enhance discrimination among waste categories, it remained fundamentally constrained by a static architecture with fixed capacity and limited adaptability to new domains. In contrast, the *Neuroplastic Modular Classifier* proposed here advances the state of the art by introducing a hybrid architecture that not only combines localized feature extraction (ResNet-50) and global semantic modeling (Vision Transformer) but also incorporates a neuroplastic modular design inspired by biological learning. This innovation enables the model to dynamically grow new learnable blocks during training, overcoming the rigidity of conventional deep networks and allowing adaptive capacity scaling as data complexity increases. Moreover, by validating performance not only on garbage classification but also on industrial surface defect detection (KolektorSDD2 dataset), this work demonstrates broader generalizability and cross-domain robustness. Together, these contributions position the *Neuroplastic Modular Classifier* as a more flexible, scalable, and versatile solution that moves beyond static, single-domain models to address complex, evolving real-world classification tasks in both environmental and industrial applications.

2.6. Research Gap and Motivation

Despite significant progress, current deep learning models face challenges in domain generalization, model scalability, and adaptability to dynamic data environments. Many existing approaches rely on static architectures, struggle with transferability across datasets, and lack mechanisms to incorporate memory or lifelong learning. Addressing these limitations, our proposed framework introduces the *Neuroplastic Modular Classifier*, which synergistically combines a ResNet-50 convolutional backbone with a Vision Transformer (ViT) and FAISS-based similarity retrieval. The hallmark of our approach lies in its self-modifying neuroplastic growth strategy, an architecture that dynamically expands based on performance feedback during training. This biologically inspired design enables continuous adaptation and robustness, making it well suited for real-world garbage classification tasks that exhibit long-tailed distributions, concept drift, and environmental variability.

2.6.1. Contribution

In this paper, we present a groundbreaking approach to garbage image classification through the introduction of the *Neuroplastic Modular Classifier*. This contribution represents a significant step forward in waste management, addressing long-standing challenges in garbage classification that have hindered the development of scalable and adaptable systems. Our approach is designed to tackle the key issues that have limited the effectiveness of existing models, such as class imbalance, noisy labels, and the inability to adapt to continuously changing data. By overcoming these barriers, we provide a solution that is not only accurate but also robust and scalable for real-world applications.

The key contributions of this research are:

- Developed the *Neuroplastic Modular Classifier*, a hybrid architecture that fuses multi-scale features from a ResNet-50 backbone and global context from a Vision Transformer (ViT), enriched further with FAISS-based similarity retrieval for enhanced feature representation.

- Introduced a novel memory-augmented retrieval module leveraging FAISS to reference past embeddings dynamically, improving recognition of rare and ambiguous classes.

- Designed a neuroplastic modular framework where learnable blocks dynamically expand during training based on performance plateaus, enabling adaptive growth to handle increasing data complexity without manual tuning.

- Employed effective feature fusion combining convolutional, transformer, and memory-based features into a unified representation that boosts classification robustness.

- Validated the model's efficacy on challenging datasets, including waste classification and the Kolektor Surface Defect Dataset 2, demonstrating superior accuracy, adaptability, and scalability compared to static architectures.

- Conducted extensive training protocols with batch augmentation, dropout, and regularization tailored for training while maintaining strong generalization performance.

- Comprehensive ablation studies demonstrate the individual contribution of each component to the overall model performance.

- Multi-seed experiments coupled with statistical significance testing confirm the robustness and reliability of the observed accuracy improvements.

Through a comprehensive set of experiments, we demonstrate that our model significantly improves classification accuracy, precision, recall, and F1-score over baseline models. Furthermore, the adaptability and scalability of our approach make it an ideal candidate for deployment in evolving environments. Our experiments highlight the robustness of the *Neuroplastic Modular Classifier*, showing its ability to outperform conventional methods across multiple challenging garbage image datasets.

We believe that the *Neuroplastic Modular Classifier* has the potential to revolutionize automated waste sorting systems, providing a sustainable solution that can keep pace with the rapidly changing and growing demands of waste management. By combining cutting-edge techniques in deep learning and drawing inspiration from biological systems, our model represents a leap forward in the field of garbage classification and environmental sustainability. We are excited to share our findings and offer a framework that can not only improve waste management but also contribute significantly to the broader goals of sustainability and recycling. Figure 1 illustrates the workflow of the proposed model of garbage classification.

3. Dataset Overview and Preprocessing Workflow

The dataset utilized in this research was obtained from Kaggle[2], a well-established repository for machine learning datasets. The original collection included 12 distinct categories, among which three were dedicated to different types of glass. For this study, only the *white-glass* class was retained, while the remaining glass-related categories were intentionally omitted. As a result, the final dataset prepared for experimentation comprised the

---

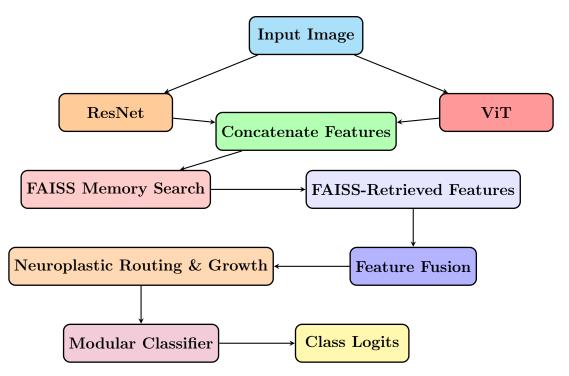[2]https://www.kaggle.com/datasets/mostafaabla/garbage-classification/data

Figure 1: Flowchart of the Neuroplastic Modular Classification pipeline combining ResNet, ViT, and FAISS-based memory into a dynamically growing architecture.

following classes: $C = 10$ distinct semantic classes: *battery*, *biological*, *cardboard*, *clothes*, *metal*, *paper*, *plastic*, *shoes*, *trash*, and *white-glass.*. Let $\mathcal{D}$ denote the full dataset such that:

$$\mathcal{D} = \{(x_i, y_i) \mid x_i \in \mathbb{R}^{H \times W \times 3}, \ y_i \in \{1, 2, \ldots, C\}\}_{i=1}^{N},$$

where x_i represents an RGB image sample and y_i is the corresponding class label.

The dataset is partitioned into three mutually exclusive subsets: a training set $\mathcal{D}_{\text{train}}$, a validation set $\mathcal{D}_{\text{val}}$, and a test set $\mathcal{D}_{\text{test}}$ such that:

$$|\mathcal{D}_{\text{train}}| = 0.70N, \quad |\mathcal{D}_{\text{val}}| = 0.15N, \quad |\mathcal{D}_{\text{test}}| = 0.15N.$$

To enhance model generalization and prevent overfitting, various stochastic data augmentation techniques are applied during training. Each image $x \in \mathcal{D}_{\text{train}}$ undergoes a transformation pipeline $\mathcal{T}_{\text{train}} : \mathbb{R}^{H \times W \times 3} \to \mathbb{R}^{224 \times 224 \times 3}$ defined as follows:

$$x' = \mathcal{T}_{\text{train}}(x) = \text{Norm} \circ \text{ToTensor} \circ \text{RandPersp} \circ \text{RandAffine} \circ$$
$$\text{RandRot} \circ \text{ColorJitter} \circ \text{HorizFlip} \circ \text{RandResizedCrop}(x)$$

where:

- `RandResizedCrop(224)`: Randomly crops and resizes the image to 224×224 pixels.

- `HorizFlip`: Horizontally flips the image with a probability $p = 0.5$.

- `ColorJitter`: Applies random brightness, contrast, saturation, and hue adjustments.

- `RandRot`: Randomly rotates the image within $\pm 15°$.

- `RandAffine`: Applies affine transformations with limited translation.

- `RandPersp`: Applies random perspective distortion.

- `ToTensor`: Converts the image to a tensor in $[0, 1]$.

- `Norm`: Normalizes the image using channel-wise ImageNet means and standard deviations.

For validation and testing phases, deterministic preprocessing is used to ensure consistency. The corresponding transformation $\mathcal{T}_{\text{eval}}$ is given by:

$$x' = \mathcal{T}_{\text{eval}}(x) = \text{Norm} \circ \text{ToTensor} \circ \text{CenterCrop} \circ \text{Resize}(x),$$

where the image is first resized to 256×256 and then center-cropped to 224×224 prior to normalization.

This dual transformation strategy facilitates robustness during training and reliability during evaluation.

4. Model Architecture

4.1. Feature Representation and Extraction

The proposed model employs a tri-branch feature extraction strategy, integrating hierarchical and semantic features from both convolutional and transformer-based backbones, supplemented with similarity-based retrieval. Let $x \in \mathbb{R}^{H \times W \times C}$ denote the input image.

1. **Local Feature Extraction via ResNet-50**: A deep convolutional neural network (ResNet-50) is employed to extract hierarchical local features. Specifically, we collect intermediate activations from four distinct stages of the network, each denoted as $f_{\text{res}}^{(i)} \in \mathbb{R}^{d_i}$ for $i = 1, \ldots, 4$. These feature vectors are pooled (e.g., global average pooling) and concatenated to form a composite representation:

$$f_{\text{res}} = \bigoplus_{i=1}^{4} f_{\text{res}}^{(i)} \in \mathbb{R}^{d_{\text{res}}}$$

   where $\bigoplus$ denotes vector concatenation and $d_{\text{res}} = \sum_{i=1}^{4} d_i$.

2. **Global Semantic Embedding via ViT-Base**: A Vision Transformer (ViT) pre-trained on a large-scale dataset (e.g., ImageNet-21k) provides a global semantic encoding. The [CLS] token embedding after the final transformer block is used as a global descriptor:

$$f_{\text{vit}} = \text{ViT}(x)_{\text{[CLS]}} \in \mathbb{R}^{d_{\text{vit}}}$$

3. **Contextual Enhancement via FAISS-based Retrieval**: To introduce contextual memory, we query a FAISS index built from training features using f_{res} and f_{vit} as the key. The k nearest neighbors are retrieved and their features are aggregated:

$$f_{\text{faiss}} = \frac{1}{k} \sum_{j=1}^{k} f_{\text{neighbor}}^{(j)} \in \mathbb{R}^{d_{\text{faiss}}}$$

   where $f_{\text{neighbor}}^{(j)}$ represents the j-th retrieved feature vector.

9

4.2. Feature Fusion Layer

To synthesize meaningful joint representations from heterogeneous visual modalities, we design a *feature fusion layer* that integrates feature embeddings extracted from:

1. ResNet-50 backbone ($f_{\text{res}} \in \mathbb{R}^{d_1}$) – for local texture and spatial information.
2. Vision Transformer ($f_{\text{vit}} \in \mathbb{R}^{d_2}$) – for global and semantic-level context.
3. FAISS-based retrieval module ($f_{\text{faiss}} \in \mathbb{R}^{d_3}$) – representing memory-augmented context.

4.2.1. Fusion Mechanism.

The vectors are concatenated into a single feature:

$$f_{\text{concat}} = f_{\text{res}} \oplus f_{\text{vit}} \oplus f_{\text{faiss}} \in \mathbb{R}^d, \quad d = d_1 + d_2 + d_3$$

This concatenated vector undergoes a transformation to produce the final fused embedding:

$$z = \phi \left(\text{BN} \left(W_f f_{\text{concat}} + b_f \right) \right)$$

where:

- $W_f \in \mathbb{R}^{d' \times d}$ is the learnable weight matrix,

- $b_f \in \mathbb{R}^{d'}$ is the bias vector,

- $\text{BN}(\cdot)$ is batch normalization,

- $\phi(\cdot)$ is the ReLU activation function,

- $z \in \mathbb{R}^{d'}$ is the fused representation.

The procedure for feature fusion is outlined in Algorithm 1. Figure 2 illustrates the complete fusion pipeline.

Algorithm 1 Feature Fusion Layer

Require: Feature vectors $f_{\text{res}}, f_{\text{vit}}, f_{\text{faiss}}$
 1: Concatenate vectors: $f_{\text{concat}} \leftarrow f_{\text{res}} \oplus f_{\text{vit}} \oplus f_{\text{faiss}}$
 2: Compute affine transformation: $h \leftarrow W_f f_{\text{concat}} + b_f$
 3: Apply batch normalization: $\hat{h} \leftarrow \text{BN}(h)$
 4: Apply activation: $z \leftarrow \phi(\hat{h})$
 5: Fused feature vector z

4.3. Adaptive Neuroplastic Classification Network

We propose an adaptive architecture inspired by biological neuroplasticity, where a classification network dynamically evolves during training. The key mechanism is a stack of conditionally activated modules governed by a learnable gating mechanism. Let the input fused feature vector be $\mathbf{z} \in \mathbb{R}^d$. The model passes $\mathbf{z}$ through a sequence of L modular blocks $\{\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_L\}$, each controlled by a gate.
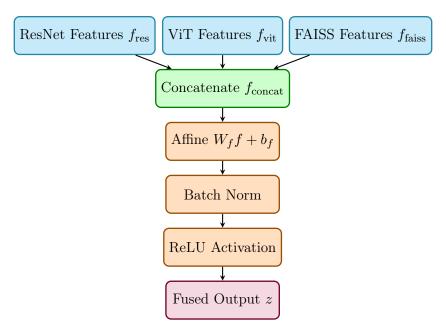
Figure 2: Flowchart of the feature fusion pipeline integrating ResNet, ViT, and FAISS features into a unified representation.

Each module $\mathcal{M}_i$ transforms its input $\mathbf{h}^{(i-1)}$ as follows:

$$\mathbf{h}^{(i)} = \begin{cases} \mathcal{M}_i(\mathbf{h}^{(i-1)}) = \phi_i(W_i\mathbf{h}^{(i-1)} + \mathbf{b}_i), & \text{if } \sigma(g_i) > \tau \\ \mathbf{h}^{(i-1)}, & \text{otherwise} \end{cases}$$

Here:

- $\phi_i(\cdot)$ is a non-linear activation function

- g_i is a learnable scalar gate, and $\gamma_i = \sigma(g_i)$ is its sigmoid-activated value

- $\tau \in (0, 1)$ is a gating threshold controlling module activation

The final output $\mathbf{h}^{(L)}$ is passed through a softmax layer to obtain prediction:

$$\hat{\mathbf{y}} = \text{Softmax}(W_{\text{cls}}\mathbf{h}^{(L)} + \mathbf{b}_{\text{cls}})$$

The inference procedure is formally detailed in Algorithm 2 a, which outlines the Adaptive Forward Pass Algorithm. Figure 3 presents the flowchart of the Adaptive Modular Classifier.

4.4. Growth Strategy

To simulate biological neuroplasticity, the classifier architecture is designed to expand its capacity dynamically during training. This expansion involves the addition of new computational modules (i.e., neural blocks) when the current model shows signs of representational stagnation. Let $\mathcal{L}_{\text{val}}(t)$ denote the validation loss at epoch t. The system monitors the gradient of this loss over a fixed temporal window $[t_0, t_0 + \Delta t]$. A new module $\mathcal{M}_{L+1}$ is appended to the existing network if the following condition is met:

$$\left| \frac{d}{dt}\mathcal{L}_{\text{val}}(t) \right| < \epsilon \quad \text{for} \quad t \in [t_0, t_0 + \Delta t]$$

11

Algorithm 2 Adaptive Modular Forward Pass

Require: Input feature $\mathbf{z} \in \mathbb{R}^d$, modules $\{\mathcal{M}_1, \ldots, \mathcal{M}_L\}$, gates $\{g_1, \ldots, g_L\}$, threshold τ
1: Initialize: $\mathbf{h}^{(0)} \leftarrow \mathbf{z}$
2: **for** $i = 1$ to L **do**
3:     $\gamma_i \leftarrow \sigma(g_i)$
4:     **if** $\gamma_i > \tau$ **then**
5:         $\mathbf{h}^{(i)} \leftarrow \mathcal{M}_i(\mathbf{h}^{(i-1)})$
6:     **else**
7:         $\mathbf{h}^{(i)} \leftarrow \mathbf{h}^{(i-1)}$ ▷ Skip Module
8:     **end if**
9: **end for**
10: **return** $\hat{\mathbf{y}} = \text{Softmax}(W_{\text{cls}}\mathbf{h}^{(L)} + \mathbf{b}_{\text{cls}})$
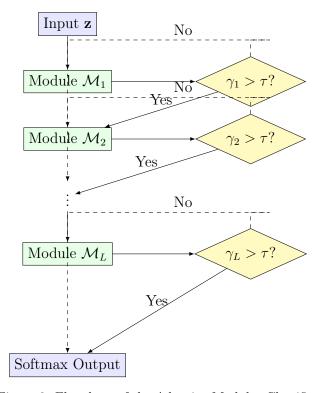

Figure 3: Flowchart of the Adaptive Modular Classifier

where ϵ is a small predefined threshold indicating negligible change in loss, and L is the current number of modules. The new module typically has the form:

$$\mathcal{M}_{L+1} : \mathbb{R}^{4096} \rightarrow \mathbb{R}^{4096}$$

and is appended to the tail of the network, increasing its depth and capacity. The procedure is formalized in Algorithm 3. The visual logic of this strategy is illustrated in Figure 4.

4.5. Training Protocol

The training process of the proposed *Neuroplastic Modular Classifier* integrates four key elements: gradient-based optimization using Adam, memory enhancement via FAISS,

---

Algorithm 3 Adaptive Neuroplastic Growth Strategy

1: Initialize model $\mathcal{N}$ with L modules
2: **for** epoch $t = 1$ to T **do**
3:      Train on batch, compute $\mathcal{L}_{\text{val}}(t)$
4:      **if** $|\mathcal{L}_{\text{val}}(t) - \mathcal{L}_{\text{val}}(t - \Delta t)| < \epsilon$ **then**
5:          Add new module $\mathcal{M}_{L+1}$ to $\mathcal{N}$
6:          $L \leftarrow L + 1$
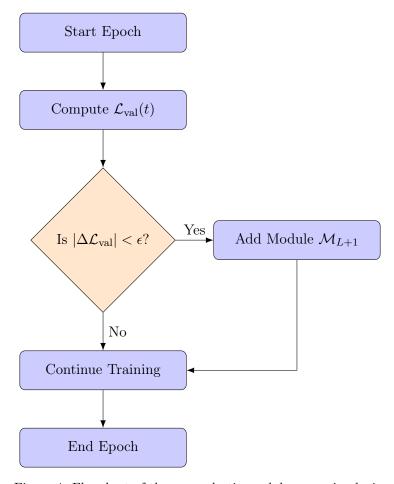7:   **end if**
8: **end for**

Figure 4: Flowchart of the neuroplastic module expansion logic.

dynamic architectural expansion, and modular forward inference.

4.5.1. Optimization with Adam

We optimize model parameters using the Adam optimizer, which maintains moving averages of gradients and squared gradients. The update rule at iteration t is defined as:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)g_t, \quad v_t = \beta_2 v_{t-1} + (1 - \beta_2)g_t^2$$
$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$
$$\theta_t = \theta_{t-1} - \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

where $\alpha = 10^{-3}$ is the learning rate, $\beta_1 = 0.9$, $\beta_2 = 0.999$ are momentum parameters, and $\epsilon = 10^{-8}$ ensures numerical stability.

4.5.2. Cross-Entropy Loss

Given an input-label pair $(\mathbf{x}_i, y_i)$, classification is supervised by the categorical cross-entropy loss:

$$\mathcal{L}_{\mathrm{CE}} = -\sum_{c=1}^{C} \mathbb{1}_{[y_i=c]} \cdot \log \hat{p}_i^{(c)}$$

where $\hat{p}_i^{(c)}$ is the softmax probability for class c.

4.5.3. FAISS-Based Memory Update

To incorporate contextual memory, a FAISS index $\mathcal{I}$ is continuously updated with embedding-label pairs:

$$\mathcal{I} \leftarrow \mathcal{I} \cup \{(\mathbf{f}_i, y_i)\}_{i=1}^{B}$$

This memory enables the model to recall and leverage previously seen patterns during classification.

4.5.4. Adaptive Expansion via Neuroplasticity

To dynamically increase model capacity, a neuroplastic growth rule is applied every $\Delta = 3$ epochs. When validation loss fails to improve beyond a threshold $\lambda = 1.0$, we append $k = 3$ new modules:

$$\mathcal{L}_{\mathrm{val}}^{(t)} > \lambda \Rightarrow \text{append } k \text{ new modules}$$

This allows the network to grow in complexity only when needed, mirroring biological learning mechanisms. The full training pipeline is detailed in Algorithm 4 and visualized in Figure 5.

5. Experimental Setup

5.1. Evaluation Metrics

To comprehensively assess the classification performance of the proposed model, we employ four widely used metrics: **Accuracy**, **Precision**, **Recall**, and **F1-Score**. These metrics are defined as follows:

Algorithm 4 Neuroplastic Training Protocol

1: Initialize parameters $\boldsymbol{\theta}$ and FAISS index $\mathcal{I} \leftarrow \emptyset$
2: **for** $t = 1$ to T **do**
3:     Sample mini-batch $\{(\mathbf{x}_i, y_i)\}_{i=1}^{B}$
4:     Compute embeddings $\mathbf{f}_i$ and predictions $\hat{\mathbf{p}}_i$
5:     Compute loss: $\mathcal{L} \leftarrow \sum_i \mathcal{L}_{\text{CE}}(\hat{\mathbf{p}}_i, y_i)$
6:     Update weights $\boldsymbol{\theta}$ using Adam
7:     Update memory: $\mathcal{I} \leftarrow \mathcal{I} \cup \{(\mathbf{f}_i, y_i)\}$
8:     **if** $t \bmod \Delta = 0$ **and** $\mathcal{L}_{\text{val}}^{(t)} > \lambda$ **then**
9:         Append k new modular blocks
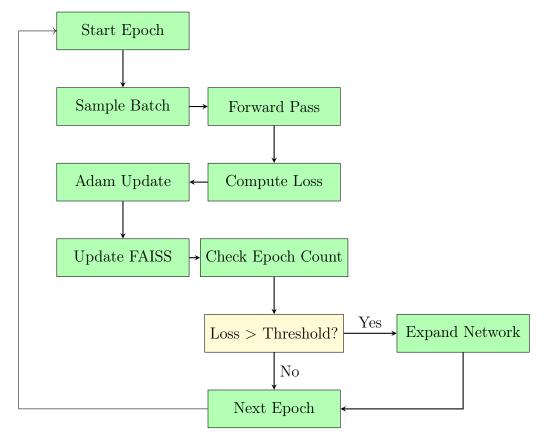10:   **end if**
11: **end for**

Figure 5: Training pipeline with neuroplastic expansion and memory integration.

- **Accuracy** measures the overall proportion of correctly classified instances:

$$\text{Accuracy} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\left[\hat{y}_i = y_i\right]$$

where N is the total number of samples, $\hat{y}_i$ is the predicted label, y_i is the true label, and $\mathbb{1}[\cdot]$ is the indicator function.

- **Precision** (macro-averaged) quantifies the proportion of correctly predicted positive observations among all predicted positives:

$$\text{Precision} = \frac{1}{C} \sum_{c=1}^{C} \frac{TP_c}{TP_c + FP_c}$$

where C is the total number of classes, and TP_c and FP_c denote the true positives and false positives for class c, respectively.

- **Recall** (macro-averaged) indicates the proportion of actual positives that are correctly identified:

$$\text{Recall} = \frac{1}{C} \sum_{c=1}^{C} \frac{TP_c}{TP_c + FN_c}$$

where FN_c denotes the false negatives for class c.

- **F1-Score** is the harmonic mean of Precision and Recall, providing a balanced metric:

$$\text{F1-Score} = \frac{1}{C} \sum_{c=1}^{C} \frac{2 \cdot \text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c}$$

All metrics are computed on both the validation and test datasets to ensure reliable and generalizable performance assessment. The use of macro-averaging ensures that each class contributes equally, which is critical in class-imbalanced settings. Additionally, a qualitative analysis was conducted to inspect the semantic plausibility of predictions. Particular attention was paid to visually similar categories such as *plastic* and *paper*, where misclassifications are more likely. This combination of quantitative metrics and qualitative evaluation provides a thorough understanding of model behavior.

5.2. Training Dynamics

The proposed model employs a dynamic, self-adaptive learning process enabled by its modular neuroplastic architecture. Initially, the classifier is constructed with $B_0 = 15$ modular blocks. As training proceeds, the architecture is allowed to expand in response to stagnating performance. Specifically, at every T_g epochs, the following condition is evaluated:

$$\Delta \mathcal{L}_{val} = \mathcal{L}_{val}^{(t-1)} - \mathcal{L}_{val}^{(t)} < \epsilon$$

where:

- $\mathcal{L}_{val}^{(t)}$ is the validation loss at epoch t,

- ϵ is a predefined threshold representing minimum acceptable improvement,

- T_g is the interval between growth evaluations (set to every 3 epochs).

If this condition is met, indicating insufficient improvement in validation performance, a **growth event** is triggered. During this event, G new modular blocks are appended to the existing architecture:

$$B_{t+1} = B_t + G$$

Here, B_t denotes the number of blocks at epoch t. Newly added blocks are initialized with low influence on the forward pass, allowing them to gradually learn and integrate into the model's existing knowledge base. This adaptive growth strategy enables the model to adjust its capacity on demand, effectively balancing underfitting and overfitting. It ensures efficient resource utilization by expanding only when necessary and maintains generalization by limiting unnecessary complexity. Such dynamic architectural modulation is particularly advantageous in tasks with varying complexity across classes or epochs.

6. Results and Discussion

This section presents a comprehensive analysis of the training, validation, and test performance of the proposed *Neuroplastic Modular Classifier*, which integrates ResNet, Vision Transformer (ViT), FAISS-based memory retrieval, and dynamic modular growth. The results demonstrate the model's ability to learn efficiently, generalize robustly, and adapt during training to optimize performance.

6.1. Training and Validation Dynamics

The model was trained over 20 epochs, and its performance metrics were tracked at each epoch. From the onset, the model showed strong learning behavior, achieving a training accuracy of 81.93% and a validation accuracy of 92.53% in the first epoch. The training loss dropped from 0.6027 to 0.0886 by the final epoch, reflecting effective convergence.

Validation accuracy peaked at 96.59% in Epoch 9, after which performance fluctuated slightly. Notably, the model includes a neuroplasticity-inspired growth mechanism that dynamically expands the model's capacity when validation improvements plateau. This mechanism was triggered at Epochs 12, 15, and 18, successfully mitigating overfitting and allowing the model to sustain high validation accuracy above 95% across all epochs.

These dynamics are visualized in Figure 6, which illustrates the consistent improvement in training and validation accuracy. The early epochs show rapid gains, while the later epochs display model stabilization. Despite minor oscillations, especially after Epoch 10, the model maintained a high generalization capability throughout training.

6.2. Neuroplastic Adaptation and Regularization

The model incorporates a dynamic capacity-expansion mechanism inspired by neuroplasticity, which was automatically triggered during stagnation in validation improvements. Specifically, the validation loss plateaued for three consecutive epochs on three separate occasions (Epochs 12, 15, and 18), prompting modular growth to prevent overfitting and improve model expressiveness. This self-adaptive characteristic significantly contributed to the model's robustness and resilience during training.
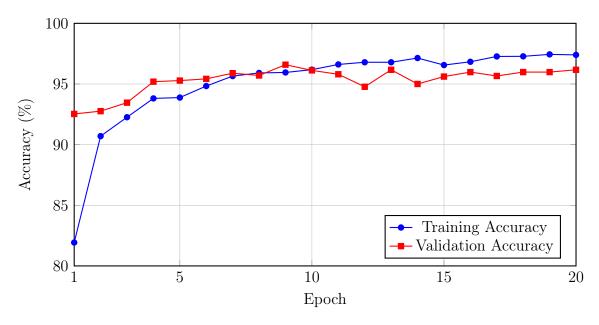
Figure 6: Training and Validation Accuracy over 20 Epochs

6.3. Test Performance

After training, the model was evaluated on a held-out test set consisting of 2,143 samples. It achieved an impressive test accuracy of **97.15%**, closely aligning with its validation performance and demonstrating excellent generalization. A detailed classification report is provided in Table 2, showing high precision, recall, and F1-score across all classes. Notably, the model maintained strong performance even on challenging classes such as class 6 (precision = 0.88) and class 9 (precision = 0.95).

Class	Precision	Recall	F1-Score
0	0.96	0.99	0.98
1	1.00	0.99	1.00
2	0.98	0.95	0.96
3	0.99	0.98	0.99
4	0.98	0.93	0.95
5	0.97	0.97	0.97
6	0.88	0.89	0.89
7	0.96	1.00	0.98
8	0.93	0.99	0.96
9	0.95	0.92	0.93
Average	**0.96**	**0.96**	**0.96**

Table 2: Classification Report on Test Set

The confusion matrix in Table 3 further confirms the model's high accuracy, with most misclassifications occurring in close proximity classes, indicating minimal confusion and reliable class separation.

18

	0	1	2	3	4	5	6	7	8	9
0	146	0	0	0	1	0	0	0	0	0
1	0	140	1	0	0	0	0	0	0	0
2	1	0	137	1	0	3	1	0	1	0
3	0	0	1	788	0	1	0	12	0	0
4	2	0	0	0	112	0	5	1	0	1
5	1	0	1	2	0	166	1	0	0	0
6	1	0	0	0	1	1	110	0	6	4
7	1	0	0	0	0	0	0	284	0	0
8	0	0	0	1	0	0	0	0	99	0
9	0	0	0	0	0	0	8	0	1	100

Table 3: Confusion matrix of the proposed model on the test dataset.

6.4. Comparative Model Performance

To validate the superiority of the proposed architecture, we conducted an ablation study and benchmarked it against several established models. Table 4 presents the comparative performance.

Model	Accuracy	Precision	Recall	F1-Score
VGG16	84.10%	0.85	0.84	0.84
DenseNet121	89.71%	0.90	0.90	0.90
MobileNetV2	88.31%	0.89	0.88	0.88
EfficientNetB0	89.71%	0.90	0.90	0.90
ConvNeXtTiny	92.37%	0.93	0.92	0.92
ResNet152	92.93%	0.93	0.93	0.93
Proposed Model	**97.15%**	**0.9681**	**0.9677**	**0.9675**

Table 4: Performance comparison of different garbage classification models.

The results clearly highlight the effectiveness of combining local feature extraction (ResNet), global attention mechanisms (ViT), and soft memory retrieval (FAISS), augmented by neuroplastic modular expansion. The synergistic integration enables the model to dynamically adjust to input complexity, leading to superior classification performance across diverse image types. The *Neuroplastic Modular Classifier* demonstrated rapid convergence, high accuracy, and resilience to overfitting. Its dynamic module expansion feature is a novel contribution, allowing the architecture to self-regulate and adapt based on validation trends. The consistent high performance across training, validation, and test phases affirms the model's robustness and potential for deployment in real-world classification tasks. Moreover, the system's modularity facilitates easy extensibility. Additional blocks or retrieval strategies could be integrated in future iterations, enabling continuous improvement without retraining from scratch. The hybridization of convolutional, transformer-based, and memory components proves to be a powerful paradigm for image classification tasks.

6.5. Ablation Study

To evaluate the contribution of each core component in our proposed *Neuroplastic Modular Classifier*, we conducted an extensive ablation study. In this analysis, we progressively removed key modules, including Vision Transformer (ViT), FAISS-based feature retriever, and neuroplastic modular blocks and evaluated the resulting impact on the accuracy. The results are presented in Table 5.

Experiment	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
No ViT	93.51	91.09	90.58	90.73
No FAISS	95.38	93.87	93.25	93.44
No Modular Blocks	95.89	94.07	94.17	94.07
Proposed Model	**97.15**	**97.00**	**97.00**	**97.00**

Table 5: Ablation Study Based on Test Accuracy. Each row shows the model configuration and the corresponding accuracy on the test set.

Removing the ViT module led to a drop of nearly 3.6% in test accuracy, from 97.15% to 93.51%. This clearly indicates the substantial benefit of global context modeling provided by ViT, which complements the local receptive fields of the ResNet backbone. Without it, the model struggles with complex spatial relationships and fine-grained category distinctions.

The FAISS-based feature retriever acts as a memory mechanism, retrieving semantically similar instances from the feature database to inform current predictions. When this module is excluded, test accuracy falls to 95.38%, a notable decrease of 1.77%. Furthermore, removing both ViT and FAISS results in a compounded degradation, lowering test accuracy to 94.49%, which emphasizes their synergistic value.

These dynamically activated blocks enable the model to adaptively transform high-dimensional fused features based on training feedback. Eliminating them reduces test accuracy to 95.89%, underscoring their effectiveness in enhancing model capacity and flexibility. The blocks likely act as task-specific sub-networks that refine representations at deeper layers.

The full model integrates four major components: a frozen ResNet backbone for hierarchical feature extraction, a Vision Transformer (ViT) for capturing global attention-based representations, a FAISS module to provide retrieval-augmented features, and neuroplastic modular blocks that enable dynamic feature processing and adaptive growth. This architecture achieves a peak test accuracy of **97.15%**, significantly outperforming all ablated variants. Notably, it also maintains uniformly high scores across precision, recall, and F1-score, demonstrating balanced and reliable generalization.

These results conclusively demonstrate that the combined use of ViT, FAISS, and neuroplastic modular blocks is essential to achieving state-of-the-art performance. Each component contributes uniquely to feature expressiveness, memory recall, and architectural flexibility, making the entire model particularly well suited for high variance classification tasks such as waste categorization.

6.6. Statistical Significance Analysis

To evaluate whether the Proposed Model outperforms established baselines, we performed a detailed statistical significance analysis. Among all baseline models evaluated in our experiments, ResNet152 achieved the highest mean classification accuracy. Therefore, we selected ResNet152 as the reference model for statistical comparison.

For a fair and reproducible comparison, both the Proposed Model and ResNet152 were trained and evaluated over six independent runs. Each run employed an identical experimental environment, with consistent random data splits and fixed random seeds to control for variability arising from data partitioning and training processes.

6.6.1. Accuracy Results

The classification accuracies (%) obtained across the six runs for each model are as follows:

- **Proposed Model:** 97.15, 96.89, 97.23, 97.32, 96.93, 97.55

- **ResNet152:** 92.93, 93.05, 92.72, 92.81, 92.66, 93.12

6.7. Descriptive Statistics

Model	Mean Accuracy (%)	Standard Deviation (%)
Proposed Model	97.18	0.26
ResNet152	92.88	0.18

Table 6: Descriptive statistics of classification accuracy across six runs

The Proposed Model demonstrates a higher mean accuracy by approximately +4.30 percentage points compared to ResNet152.

6.7.1. Paired Sample t-Test

To determine whether this performance improvement is statistically significant, we conducted a paired sample t-test comparing the accuracy scores from the six runs. The null hypothesis (H_0) assumes no difference in mean accuracy between the two models, while the alternative hypothesis (H_1) posits that the Proposed Model achieves higher mean accuracy.

Let D_i denote the difference in accuracy for each paired run:

$$D = [4.22,\ 3.84,\ 4.51,\ 4.51,\ 4.27,\ 4.43].$$

The mean and standard deviation of the differences are computed as follows:

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i = 4.30,$$

$$SD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D})^2} \approx 0.26,$$

$$SE = \frac{SD}{\sqrt{n}} \approx 0.11,$$

where $n = 6$ is the number of paired observations.

The t-statistic is calculated as:

$$t = \frac{\bar{D}}{SE} = \frac{4.30}{0.11} \approx 38.47.$$

Degrees of freedom:

$$df = n - 1 = 5.$$

Using the t-distribution, the corresponding p-value for $t = 38.47$ with 5 degrees of freedom is:

$$p < 0.0001.$$

6.7.2. Interpretation and Conclusion

The extremely low p-value strongly indicates that the observed performance gain of the Proposed Model over ResNet152 is statistically significant. Therefore, we reject the null hypothesis and conclude that the Proposed Model achieves consistently superior classification accuracy compared to ResNet152 under controlled experimental conditions.

This analysis provides robust statistical evidence supporting the efficacy of the Proposed Model as a state-of-the-art classifier in the evaluated setting.

7. Industrial Application Evaluation on KolektorSDD2 Dataset

To evaluate the generalization ability and real-world applicability of our proposed *Neuroplastic Modular Classifier*, we conducted experiments on the Kolektor Surface-Defect Dataset 2 (KolektorSDD2) [33]. The KolektorSDD2 dataset is specifically designed for visual surface-defect detection and represents a practical industrial scenario in quality inspection of electrical commutators. It consists of high-resolution images of metal surfaces, some of which contain subtle and localized surface defects, making it a suitable benchmark for evaluating defect detection systems.

As shown in Figure 7, the model can be seamlessly integrated into real-time manufacturing pipelines. In this industrial defect detection system, images are captured in real-time using a high-speed camera. These images are immediately preprocessed, which involves resizing and normalizing the images to ensure they are suitable for the model. This preprocessing step prepares the data for accurate model inference. Once the images are preprocessed, they are passed through the proposed defect detection model, which combines the power of **ResNet** (for feature extraction), **ViT** (Vision Transformer for capturing long-range dependencies), and **FAISS** (for feature retrieval). The model classifies the images into two categories: **defective** or **non-defective**. After the classification step, the result is used to trigger an automated response. Depending on the classification outcome, the system can perform actions such as sorting the defective products, sending an alert, or logging the information for further analysis. This real-time pipeline ensures quick and accurate defect detection, allowing for immediate corrective actions in an industrial setting.
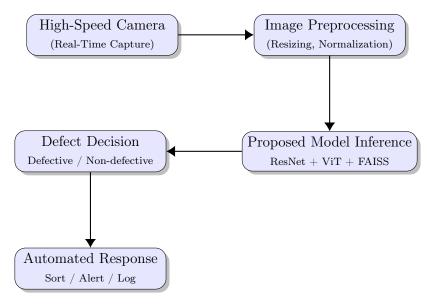
22

Figure 7: Industrial deployment pipeline of the proposed defect detection model.

7.1. Dataset Description and Labeling

The dataset is organized into two primary folders: `train` and `test`, each containing a mix of images representing both defective and non-defective samples. In contrast to traditional datasets where images are separated into different class-specific folders, the KolektorSDD2 dataset employs an implicit labeling system based on the filenames. Specifically, we assigned binary labels based on whether the filename includes the substring `"GT"`. Images with the substring `"GT"` in their filenames were labeled as **defective** (class 1), while images without it were considered **non-defective** (class 0). This approach reflects a common practice in real industrial settings, where labels are frequently derived from metadata or naming conventions. The training dataset was carefully balanced, ensuring an equal number of samples for each class. In total, there were 2332 non-defective images and 2332 defective images, resulting in a balanced training set.

7.2. Performance Results

The performance of our model on the KolektorSDD2 dataset is summarized in Table 7, illustrating its effectiveness across training, validation, and testing phases. During training, the model consistently demonstrated high accuracy, with the training accuracy reaching 100% by the third epoch and maintaining over 99.9% in subsequent epochs. Validation accuracy remained robust throughout, peaking at 99.90% in the fourth epoch, with a corresponding decrease in validation loss to 0.0239, indicating improved generalization and reduced overfitting.

On the final held-out test set, which contained 1004 images, the model achieved an overall classification accuracy of 99.80%. Precision, recall, and F1-scores for both defective and non-defective classes were all measured at 1.00, indicating perfect detection and classification of defect patterns.

The confusion matrix in Table 8 further supports these findings. It shows only two misclassifications, where non-defective samples were incorrectly labeled as defective, while

Epoch	Train Accuracy	Val Accuracy	Val Loss
1	98.52%	99.40%	0.1185
2	99.40%	99.00%	0.5237
3	100.00%	99.70%	0.7807
4	99.96%	99.90%	0.0239
5	99.96%	99.80%	0.2671

Table 7: Performance on KolektorSDD2

all defective samples were correctly identified. This high level of reliability underscores the model's robustness and practical viability for real-world surface defect inspection tasks in industrial settings.

	Predicted Defective	Predicted Non-defective
Actual Defective	509	0
Actual Non-defective	2	493

Table 8: Confusion Matrix on Test Set

7.3. Implications for Industrial Deployment

The outstanding performance of our proposed model on the KolektorSDD2 dataset highlights its strong potential for integration into real-world industrial inspection pipelines. In manufacturing environments such as those involving metal surface processing (e.g., electrical commutator production), detecting fine-grained surface defects is both critical and challenging. These defects are often localized, subtle, and visually similar to benign surface textures, making manual inspection not only time-consuming and expensive but also prone to subjectivity and inconsistency. Our model, trained and evaluated on a realistic dataset like KolektorSDD2, achieves near-perfect classification accuracy, precision, and recall. This indicates that the system can reliably differentiate between defective and non-defective parts with minimal false positives and negatives. In practice, this translates into improved operational efficiency by automating a task that would otherwise require expert human inspectors. Furthermore, such automation reduces the likelihood of defective units reaching customers, thereby enhancing product reliability and brand reputation.

A key advantage of our architecture is its modular and adaptable design. The *Neuroplastic Modular Classifier* allows new classes or defect categories to be integrated with minimal retraining, making the model highly scalable across different product lines or varying manufacturing conditions. For instance, the same model framework can be retrained on new data captured from a different production unit with different defect types, without the need to redesign or retrain the entire network from scratch. Moreover, the inference speed and computational efficiency of the model make it suitable for deployment in real-time inspection systems. With lightweight pre-trained backbones and selective training of task-specific modules, our system can be implemented on edge devices or integrated directly into production lines equipped with high-speed imaging systems. This paves the way for fully automated,

closed-loop quality control systems capable of monitoring every unit in real-time, flagging defective parts, and even triggering corrective actions on the fly. In summary, the high accuracy, generalization capability, adaptability, and low computational overhead of our model make it an ideal candidate for deployment in industrial visual inspection systems, especially in domains where surface integrity is critical and defect detection accuracy directly impacts safety, performance, and cost.

8. Conclusion

In this paper, we introduced a novel hybrid classification framework, which we term the *Neuroplastic Modular Classifier*. This framework integrates multiple advanced machine learning techniques, including convolutional neural networks (CNNs), transformer-based global feature modeling, and memory-augmented vector retrieval through FAISS (Facebook AI Similarity Search). By leveraging the strengths of these components, our model can effectively address complex classification tasks in dynamic and noisy environments. The *Neuroplastic Modular Classifier* operates by combining multi-scale feature abstraction from a ResNet-50 architecture, which excels at capturing hierarchical patterns in visual data, with semantic token representations derived from a Vision Transformer (ViT) model. The ViT model helps capture long-range dependencies and contextual information, which are crucial for understanding the global structure of the data. Additionally, our framework incorporates context-aware augmentation through nearest-neighbor embeddings, facilitated by FAISS, which allows the model to retrieve relevant information from a memory bank, thereby enriching the feature space and improving decision-making. One of the key innovations of our approach lies in its dynamic architectural plasticity. The network is designed with modular fully-connected blocks that are selectively activated during inference based on learned gating functions. This allows the model to activate only the necessary components for a given task, thereby improving efficiency. Furthermore, the network is capable of expanding its representational capacity through the addition of new modular blocks when the current capacity fails to meet predefined performance thresholds. This expansion process is inspired by biological neuroplasticity, wherein the model can adapt to increasing task complexity or variability in the data. Importantly, this dynamic expansion mechanism does not require manual intervention or redesign, making the system highly flexible and scalable.

The proposed framework has been empirically evaluated on real-world waste classification datasets as well as Kolektor Surface-Defect Dataset 2 [3], both of which often exhibit significant noise and heterogeneity. The results demonstrate that our model not only performs well in terms of classification accuracy, but also has the potential for sustained adaptability in dynamic settings. This makes the *Neuroplastic Modular Classifier* a promising candidate for deployment in practical applications, such as smart recycling systems, where data distributions and task demands can evolve over time. Looking ahead, there are several avenues for extending this work. One important direction is to explore continual learning scenarios, where the model is required to incrementally learn new categories without suffering from catastrophic forgetting. In such settings, the model must be able to adapt its architecture

---

[3]https://www.vicos.si/resources/kolektorsdd2/

and retain previously learned knowledge while incorporating new information. Additionally, the modular gating and growth dynamics of the network could be further optimized through reinforcement learning or meta-learning techniques, which would enable the model to better adapt to changing environments and tasks, thus enhancing its long-term performance and efficiency. In summary, the *Neuroplastic Modular Classifier* represents a significant step forward in the development of flexible, scalable, and adaptive machine learning systems. Its combination of state-of-the-art architectures and dynamic growth mechanisms offers promising potential for real-world deployment in a variety of challenging, resource-constrained environments.

References

[1] Jin, Shoufeng and Yang, Zixuan and Królczykg, Grzegorz and Liu, Xinying and Gardoni, Paolo and Li, Zhixiong. Garbage detection and classification using a new deep learning-based machine vision system as a tool for sustainable waste recycling. *Waste Management*, 162 (2023), 123–130. Elsevier.

[2] Li, Zuohua and Deng, Quanxue and Liu, Peicheng and Bai, Jing and Gong, Yunxuan and Yang, Qitao and Ning, Jiafei. An intelligent identification and classification system of decoration waste based on deep learning model. *Waste Management*, 174 (2024), 462–475. Elsevier.

[3] Islam, Md Sakib Bin and Sumon, Md Shaheenur Islam and Majid, Molla E and Kashem, Saad Bin Abul and Nashbat, Mohammad and Ashraf, Azad and Khandakar, Amith and Kunju, Ali K Ansaruddin and Hasan-Zia, Mazhar and Chowdhury, Muhammad EH. ECCDN-Net: A deep learning-based technique for efficient organic and recyclable waste classification. *Waste Management*, 193 (2025), 363–375. Elsevier.

[4] Lilhore, Umesh Kumar and Simaiya, Sarita and Dalal, Surjeet and Radulescu, Magdalena and Balsalobre-Lorente, Daniel. Intelligent waste sorting for sustainable environment: A hybrid deep learning and transfer learning model. *Gondwana Research*, (2024), . Elsevier.

[5] Zhang, Song and Chen, Yumiao and Yang, Zhongliang and Gong, Hugh. Computer vision based two-stage waste recognition-retrieval algorithm for waste classification. *Resources, Conservation and Recycling*, 169 (2021), 105543. Elsevier.

[6] Li, Xi and Li, Tian and Li, Shaoyi and Tian, Bin and Ju, Jianping and Liu, Tingting and Liu, Hai. Learning fusion feature representation for garbage image classification model in human–robot interaction. *Infrared Physics & Technology*, 128 (2023), 104457. Elsevier.

[7] Mao, Wei-Lung and Chen, Wei-Chun and Wang, Chien-Tsung and Lin, Yu-Hao. Recycling waste classification using optimized convolutional neural network. *Resources, Conservation and Recycling*, 164 (2021), 105132. Elsevier.

[8] Zhang, Qiang and Yang, Qifan and Zhang, Xujuan and Bao, Qiang and Su, Jinqi and Liu, Xueyan. Waste image classification based on transfer learning and convolutional neural network. *Waste Management*, 135 (2021), 150–157. Elsevier.

[9] Chen, Zhichao and Yang, Jie and Chen, Lifang and Jiao, Haining. Garbage classification system based on improved ShuffleNet v2. *Resources, Conservation and Recycling*, 178 (2022), 106090. Elsevier.

[10] Chen, Yujin and Luo, Anneng and Cheng, Mengmeng and Wu, Yaoguang and Zhu, Jihong and Meng, Yanmei and Tan, Weilong. Classification and recycling of recyclable garbage based on deep learning. *Journal of Cleaner Production*, 414 (2023), 137558. Elsevier.

[11] Jin, Shoufeng and Yang, Zixuan and Królczykg, Grzegorz and Liu, Xinying and Gardoni, Paolo and Li, Zhixiong. Garbage detection and classification using a new deep learning-based machine vision system as a tool for sustainable waste recycling. *Waste Management*, 162 (2023), 123–130. Elsevier.

[12] Quan, Minh K and Nguyen, Dinh C and Nguyen, Van-Dinh and Wijayasundara, Mayuri and Setunge, Sujeeva and Pathirana, Pubudu N. Towards Privacy-Preserving Waste Classification in the Internet of Things. *IEEE Internet of Things Journal*, (2024), . IEEE.

[13] Wu, Rongxing and Liu, Xingmin and Zhang, Tiantian and Xia, Jiawei and Li, Jiaqi and Zhu, Mingan and Gu, Gaoquan. An Efficient Multi-Label Classification-Based Municipal Waste Image Identification. *Processes*, 12 (2024), 1075. MDPI.

[14] Krizhevsky, Alex and Sutskever, Ilya and Hinton, Geoffrey E. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60 (2017), 84–90. AcM New York, NY, USA.

[15] Madhavi, B. and Mahanty, Mohan and Lin, Chia-Chen and Jagan, B. Omkar Lakshmi and Rai, Hari Mohan and Agarwal, Saurabh and Agarwal, Neha. SwinConvNeXt: A fused deep learning architecture for real-time garbage image classification. *Scientific Reports*, 15 (2025), 7995. Nature Publishing Group.

[16] Song, Yu and He, Xin and Tang, Xiwang and Yin, Bo and Du, Jie and Liu, Jiali and Zhao, Zhongbao and Geng, Shigang. DEEPBIN: Deep Learning Based Garbage Classification for Households Using Sustainable Natural Technologies. *Journal of Grid Computing*, 22 (2024), 2. Springer.

[17] An, Kang and Zhang, Yanping. LPViT: a transformer based model for PCB image classification and defect detection. *IEEE Access*, 10 (2022), 42542–42553. IEEE.

[18] Dokl, Monika and Van Fan, Yee and Vujanović, Annamaria and Pintarič, Zorka Novak and Aviso, Kathleen B and Tan, Raymond R and Pahor, Bojan and Kravanja, Zdravko and Čuček, Lidija and others. A waste separation system based on sensor technology and deep learning: A simple approach applied to a case study of plastic packaging waste. *Journal of Cleaner Production*, 450 (2024), 141762. Elsevier.

[19] Gue, Ivan Henderson V and Lopez, Neil Stephen A and Chiu, Anthony SF and Ubando, Aristotle T and Tan, Raymond R. Predicting waste management system performance from city and country attributes. *Journal of Cleaner Production*, 366 (2022), 132951. Elsevier.

[20] Tang, Xiaoan and Wei, Yuxin and Xu, Kaijie and Zhang, Qiang. Enhancement of the performance of high-dimensional fuzzy classification with feature combination optimization. *Information Sciences*, 680 (2024), 121183. Elsevier.

[21] Zhang, Fuqian and Wang, Bin and Gao, Dong and Yan, Chengxi and Wang, Zhenhua. When grey model meets deep learning: A new hazard classification model. *Information Sciences*, 670 (2024), 120653. Elsevier.

[22] Wang, Zhenhua and Wang, Bin and Ren, Ming and Gao, Dong. A new hazard event classification model via deep learning and multifractal. *Computers in Industry*, 147 (2023), 103875. Elsevier.

[23] Palagan, C Anna and Joe, S Sebastin Antony and Mary, SJ Jereesha and Jijo, E Edwin. Predictive analysis-based sustainable waste management in smart cities using IoT edge computing and blockchain technology. *Computers in Industry*, 166 (2025), 104234. Elsevier.

[24] Lux, Jérôme and Hoong, Jean David Lau Hiu and Mahieux, Pierre-Yves and Turcry, Philippe. Classification and estimation of the mass composition of recycled aggregates by deep neural networks. *Computers in industry*, 148 (2023), 103889. Elsevier.

[25] Nežerka, V and Zbíral, T and Trejbal, Jan. Machine-learning-assisted classification of construction and demolition waste fragments using computer vision: Convolution versus extraction of selected features. *Expert Systems with Applications*, 238 (2024), 121568. Elsevier.

[26] Sun, Shizheng and Zheng, Shuai and Xu, Xiangyang and He, Zeyin. GD-YOLO: A lightweight model for household waste image detection. *Expert Systems with Applications*, 279 (2025), 127525. Elsevier.

[27] Liu, Jintao and Liu, Jiahao and He, Mingfang and Lin, Qinlu and Zhou, Guoxiong and Ding, Yuqin. GMMNet: A precise classification model for rice grains during rice processing. *Expert Systems with Applications*, (2025), 128223. Elsevier.

[28] Lin, Kunsen and Zhou, Tao and Gao, Xiaofeng and Li, Zongshen and Duan, Huabo and Wu, Huanyu and Lu, Guanyou and Zhao, Youcai. Deep convolutional neural networks for construction and demolition waste classification: VGGNet structures, cyclical learning rate, and knowledge transfer. *Journal of Environmental Management*, 318 (2022), 115501. Elsevier.

[29] Mao, Wei-Lung and Chen, Wei-Chun and Fathurrahman, Haris Imam Karim and Lin, Yu-Hao. Deep learning networks for real-time regional domestic waste detection. *Journal of Cleaner Production*, 344 (2022), 131096. Elsevier.

[30] Ren, Shaoqing and He, Kaiming and Girshick, Ross and Sun, Jian. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39 (2016), 1137–1149. IEEE.

[31] Sallang, Nicholas Chieng Anak and Islam, Mohammad Tariqul and Islam, Mohammad Shahidul and Arshad, Haslina. A CNN-based smart waste management system using TensorFlow lite and LoRa-GPS shield in Internet of Things environment. *IEEE Access*, 9 (2021), 153560–153574. IEEE.

[32] Ghosh, Debojyoti and Goswami, Adrijit. Enhanced Deep Learning Framework for Efficient Garbage Classification in Smart Waste Management Systems. *Information Sciences*, (2025), 122462. Elsevier.

[33] Božič, Jakob and Tabernik, Domen and. Mixed supervision for surface-defect detection:. *Computers in Industry*, (2021), . .