Abstract

Automated waste classification has emerged as a critical requirement for modern waste management systems, enabling efficient recycling, reducing landfill dependency, and minimizing environmental pollution. However, real-world waste images exhibit substantial intra-class variability, inter-class similarity, contamination, deformation, and environmental noise, making accurate classification an inherently challenging task. Additionally, conventional deep learning models often rely solely on visual cues, lack mechanisms for uncertainty handling, operate with fixed architectural capacity, and inadequately integrate semantic or contextual knowledge. These limitations restrict their deployment in robust, real-world recycling and smart city infrastructures.

This thesis addresses these challenges by developing a suite of advanced hybrid deep learning architectures that combine memory-augmented retrieval, neuroplastic modularity, cross-modal semantic reasoning, and gated

Convolutional Neural Network(CNN) – Transformer fusion. First, a Facebook AI Similarity Search(FAISS)-augmented classification framework incorporates nearest-neighbor embedding retrieval to refine decisions for ambiguous samples. Second, a neuroplastic modular classifier enables dynamic architectural expansion, inspired by biological learning, to adapt model capacity to evolving data complexity. Third, a semantic-augmented Vision Transformer introduces environmental and material attribute embeddings through cross-modal attention, improving interpretability and classification performance. Finally, a gated CNN–Transformer fusion model leverages Squeeze-and-Excitation(SE)-enhanced CNN features, global Transformer representations, and multi-level auxiliary supervision to achieve efficient and robust classification under noisy real-world conditions.

Comprehensive experiments demonstrate that the proposed models consistently outperform conventional baselines across multiple metrics, achieving higher accuracy, improved generalization, and stronger resilience to challenging visual conditions. Collectively, the contributions of this thesis advance the state of automated waste classification by introducing flexible, semantically informed, and context-aware deep learning architectures suitable for deployment in intelligent recycling systems, industrial automation pipelines, and smart city environments.

In addition to environmental applications, this thesis extends the proposed context-aware deep learning paradigm to urban demand forecasting. A knowledge graph-augmented stacking framework is developed for accurate bike-sharing demand prediction, capturing spatiotemporal patterns, urban behavior, and infrastructure dynamics. This demonstrates the adaptability of the proposed architectures across sustainability-driven domains beyond computer vision.