Optimizing Waste Classification Through Large-Scale Language Models in Deep Learning

Jianchun Qi
Auckland University of Technology, New Zealand

Minh Nguyen
Auckland University of Technology, New Zealand

Wei Qi Yan
http://orcid.org/0000-0002-3946-4617
Auckland University of Technology, New Zealand

ABSTRACT

In this article, the authors introduce a novel multimodal approach that harnesses the capabilities of large language models (LLMs) to enrich waste classification with semantic insights. The method generates descriptive prompts specifically tailored to the waste imagery, which are then used to infer semantic attributes relevant to classification tasks. These descriptions facilitate a transformative converter architecture that bridges textual and visual domains, enabling the model to interpret waste images with enhanced precision. They present the first multimodal waste classification model that leverages the LLM-generated textual descriptors alongside visual features. Extensive testing shows that the approach outperforms existing models, achieving a top accuracy improvement of 62.20%. A comprehensive suite of ablation studies further underscores the method's efficacy and resilience, confirming its potential to advance waste management by integrating the complementary strengths of both image and textual data.

INTRODUCTION

Waste management is an increasingly important topic. Inadequate waste management mechanisms pose a great challenge to the protection of the ecological environment, the improvement of public health, and the safeguarding of human health. For example, the current open waste dumps, this waste management method is prone to produce hazardous chemicals (Mohanraj, Senthilkumar, Chandrasekar & Arulmozhi, 2023), pollute the soil and water, and damage the ecosystem, in addition to becoming a breeding ground for pathogens, which can easily lead to the spread of infectious diseases (Amasuomo & Baird, 2016)

DOI: 10.4018/407411

(Ferronato & Torretta, 2019). Proper waste disposal practices hold significant implications for ecological sustainability, resource efficiency, and public health enhancement. At present, waste classification is a crucial step in waste management, aiding in resource recycling and minimizing resource depletion. It contributes to reducing the reliance on waste incineration and landfills, thereby lessening pollution and safeguarding ecosystems (Gundupalli, Hait & Thakur, 2017). However, traditional methods of waste classification, which are typically semi-manual or semi-automatic, struggle to keep pace with the increasing volumes of waste, often resulting in inefficient sorting and adverse health effects on workers. Consequently, there is a pressing need to incorporate more sophisticated technologies, such as artificial intelligence, into waste management. The integration of advanced AI-driven classification techniques can lead to more effective, efficient, and health-conscious waste management practices. This, in turn, supports economic growth and environmental protection, steering us toward the sustainable coexistence of humanity and nature (Qiu et al., 2022) (Shi, Tan, Wang & Wang, 2021).

Although deep learning models for waste classification are constantly being improved and have obtained significant classification and detection results, there is still room for improvement. A slew of waste classification models, such as the optimized DenseNet121 and ResNet-10 using fusion schemes, have waste classification accuracies as high as over 85% (Ahmad, Khan & Al-Fuqaha, 2020) (Mao, Chen, Wang & Lin, 2021). However, the datasets they use are only simple recyclable waste categories, such as glass, cardboard, plastic, paper, and metal, which cannot measure the real waste classification application scenarios. According to the waste classification standard, waste should be classified into four categories, namely, recyclable waste, wet waste, dry waste, and hazardous waste. While the ETHSeg model groups the four categories of waste based on X-rays, the classification accuracy of small objects in waste remains low (Qiu et al., 2022). Thus, the lack of a waste dataset, the low accuracy of small object waste classification, and the intensive manual annotation effort due to the wide variety of waste categories are the important challenges faced by artificial intelligence in waste classification tasks.

Artificial Intelligence (AI), particularly large-scale language models, has shown remarkable promise in various applications, including image classification tasks. The cutting-edge GPT-4, for instance, can perform image classification, text-to-image conversion, and image-to-text translation. Its superior generalization and zero-shot learning abilities enable the processing of complex datasets with high efficiency. Building on this potential, our book chapter introduces a pioneering approach to waste classification by harnessing the semantic capabilities of large language models. We utilize MiniGPT-4 to generate textual descriptions of waste images, then input into the pre-trained language model RoBERTa (Liu et al., 2019) (Zhu, Chen, Shen, Li & Elhoseiny, 2023). Concurrently, we process the waste images directly through the Swin Transformer model (Liu et al., 2021). The culmination of our efforts is the novel Image-Text Aware Adaptive Attention mechanism, which integrates outputs from these dual pathways.

Despite these advancements, large language models are not without their challenges in practical scenarios. They are composed of multi-layer neural networks with hundreds of millions of parameters, necessitating substantial computational resources and extended training periods. This complexity results in considerable training expenses. Furthermore, an increase in model parameters can complicate the model's interpretability and elevate its complexity (Singla, 2023).

Therefore, to solve this problem and avoid manually collecting image description information from a large language model, we introduced MiniGPT-4 into our model through an API interface, aiming to leverage the rich semantics of the large language model in a simplified way to create an efficient and highly accurate waste classification model. Our description-driven approach to image classification shows promise, particularly when image data is