MedWaste.AI: AI-Powered Biomedical Waste Classification System

Daya Naik[1], Varsha S[2], Shreyas G Kulkarni[3], Varshitha V[4]

[1]*Assistant Professor and Head, Department of Artificial Intelligence and Machine Learning, Srinivas Institute of Technology (S.I.T), Mangalore, Karnataka, India*

[2,3,4]*Students, Department of Artificial Intelligence and Machine Learning, Srinivas Institute of Technology (S.I.T), Mangalore, Karnataka, India*

doi.org/10.64643/IJIRTV12I12-206679-459

***Abstract*—This paper presents MedWaste.AI, an intelligent biomedical waste management platform that automates waste identification and delivers context-specific disposal guidance. Improper segregation of clinical waste continues to pose significant threats to public health and environmental safety, particularly within high-throughput healthcare facilities. The proposed system employs a deep learning-based object detection model to classify waste from images and supports text and voice-based queries to accommodate diverse user needs. Upon identification, the system retrieves relevant regulatory information through a Retrieval-Augmented Generation (RAG) pipeline and generates accurate, situation-specific disposal instructions. A text-to-speech output module further improves usability in time-critical clinical environments. By integrating computer vision, natural language processing, and speech technologies, MedWaste.AI demonstrates high classification accuracy, reduced manual workload, and improved compliance with established waste disposal standards.**

***Index Terms*—Artificial Intelligence, Biomedical Waste Management, Computer Vision, YOLOv8, Deep Learning, Retrieval-Augmented Generation, Speech Processing, gTTS.**

I. INTRODUCTION

Effective management of biomedical waste is a foundational requirement for any functional healthcare system. When clinical waste is not handled correctly, the consequences extend far beyond hospital premises, endangering public health and contributing to environmental contamination [1]-[3]. In most healthcare institutions today, waste segregation is performed manually—a method that is inherently susceptible to classification errors and inconsistent disposal decisions. In such settings, frontline healthcare workers serve as the primary line of defense, relying on individual training and day-to-day experience to determine how waste should be categorized [4]. This dependence on human judgment introduces inconsistency, leading to compliance gaps and potential delays in safe waste handling.

Recent advances in artificial intelligence (AI) and computer vision have opened pathways to address these challenges in a structured and scalable manner [5]. Automated classification systems can analyze visual data from multiple perspectives and accurately differentiate between distinct categories of waste. They can also retrieve applicable regulatory information and generate customized disposal instructions on demand, significantly reducing the cognitive burden on healthcare staff [6].

MedWaste.AI was designed to enhance biomedical waste handling by integrating image-based recognition with AI-driven decision support. Users interact through a straightforward interface that accepts waste images as well as typed or spoken input, thereby supporting a broad range of interaction modalities [7]. Once the waste category is determined, the system queries an indexed regulatory knowledge base through a RAG architecture to generate accurate, situation-specific disposal instructions [2]. By automating waste recognition and instruction generation, the platform reduces repetitive manual effort, improves classification outcomes, and assists staff in adhering to established safety guidelines more consistently. This results in more reliable and efficient waste management workflows across diverse healthcare environments.

The integration of intelligent technology into waste management processes promotes a culture of informed and responsible clinical practice. It ensures that disposal decisions are grounded in current regulatory standards, thereby reducing the risk of environmental harm attributable to healthcare activities [8].

II. LITERATURE REVIEW

The intersection of artificial intelligence and healthcare waste management has attracted considerable research interest in recent years. Kumar and Mehta [1] proposed a deep learning framework for medical waste classification using convolutional neural networks, demonstrating that automated image-based systems could significantly outperform manual sorting in terms of accuracy and speed. Their work established a benchmark for using transfer learning in waste classification tasks.

Sharma and Verma [9] explored the integration of IoT sensors with AI models to create smart healthcare waste management systems capable of real-time monitoring. Their findings highlighted that continuous data streams from sensor networks, when combined with predictive AI models, could reduce hazardous waste handling errors by a substantial margin. Complementing this, Karthik and Deepa [3] developed an automated waste segregation prototype using computer vision, demonstrating reliable performance across multiple waste categories under varying lighting and imaging conditions.

In the domain of voice-assisted technologies, Nair and Iyer [10] investigated multilingual text-to-speech systems for assistive applications in healthcare, concluding that audio output significantly improved usability in environments where screen interaction is impractical. Joshi and Shah [11] similarly examined AI-powered voice assistants in clinical settings, finding that speech-based interfaces increased adoption rates among healthcare workers with limited technical literacy.

Object detection for waste classification has been addressed by Khan and Wei [6], who applied deep learning models to identify hazardous waste with high precision in real-world environments. Their study demonstrated the superiority of YOLO-based architectures for real-time detection tasks compared to traditional classification pipelines. Gupta and Singh [12] explored conversational AI for smart assistance systems, presenting a framework for context-aware query handling that directly informs the design of the text-processing module in MedWaste.AI.

The use of YOLO-based models for waste management was further examined by Reddy and Menon [13], who applied YOLOv5 to a mixed municipal and clinical waste dataset. Their results confirmed that single-stage detectors provide an optimal balance of speed and accuracy for deployment in real-time clinical environments. Patel and Kumar [14] advanced the field by combining AI-based object detection with environmental monitoring systems, offering an integrated approach to waste tracking and regulatory reporting that aligns closely with the goals of MedWaste.AI. Collectively, these studies inform the architecture, component selection, and evaluation strategy of the system described in this paper.

III. METHODOLGY

MedWaste.AI is built on the principle of integrating multiple specialized AI modules into a cohesive biomedical waste management solution. The architecture follows a modular design paradigm, wherein each component fulfills a discrete function—ranging from user input acquisition to regulatory document retrieval—thereby ensuring maintainability and scalability [15].

User input may be provided in one of three modalities: text, image, or voice. When a voice query is submitted, the system employs Mozilla's DeepSpeech engine [16] to convert spoken audio into text, which is subsequently processed alongside typed queries. When an image is uploaded, the YOLOv8 model [17] analyzes the visual content to detect and classify the biomedical waste present. YOLOv8 was selected for its demonstrated accuracy and computational efficiency in real-time object detection tasks [13].

Following waste identification, the system formulates a structured search query and forwards it to the information retrieval component. Retrieval is implemented through a Retrieval-Augmented Generation (RAG) pipeline that employs Facebook AI Similarity Search (FAISS) [18] to identify relevant passages from an indexed corpus of biomedical waste regulations, including the Government of India's Biomedical Waste Management Rules, 2016 [19], and guidelines published by the World Health Organization [20]. The retrieved content is passed to a large language model (LLM), which generates coherent, regulation-compliant disposal instructions tailored to the identified waste type [21].

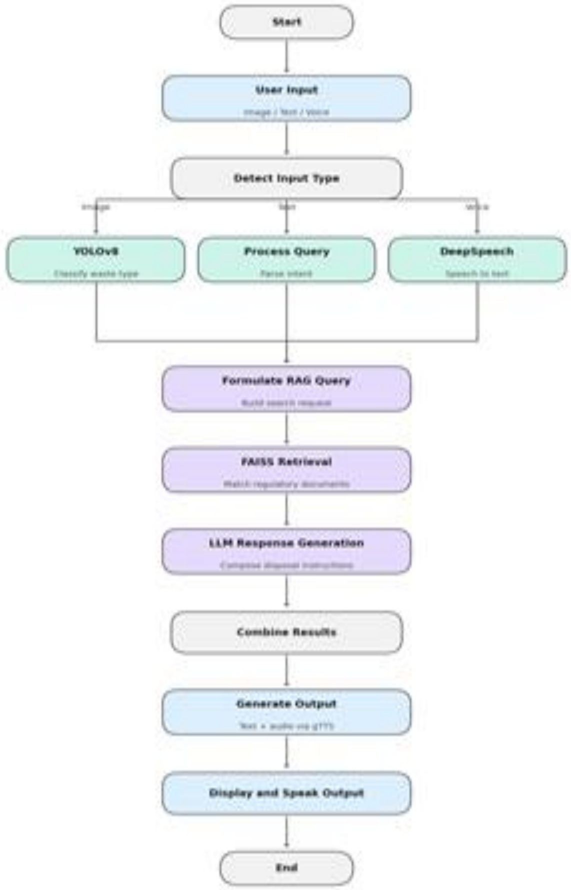

Figure 1. End-to-end operational flowchart of the MedWaste.AI system.

The final system output is presented to the user as readable text and, optionally, as synthesized speech via the Google Text-to-Speech (gTTS) library [22]. This audio feature is particularly beneficial in fast-paced clinical environments where visual interaction with a screen may not be feasible, and also facilitates multilingual accessibility. The system architecture supports image preprocessing—including resizing and normalization—prior to model inference, ensuring consistent input quality. All processing is performed locally on a mobile-compatible infrastructure, preserving user data confidentiality while retaining the flexibility for future feature integration without disruption to the core pipeline.

Figure 1 illustrates the end-to-end processing flow of the MedWaste.AI system, from user input through waste classification, regulatory retrieval, and instruction generation, to final text and audio output.

IV. DESIGN AND IMPLEMENTATION

The MedWaste.AI system architecture is structured around four principal layers: the user interface layer, the application logic layer, the AI processing layer, and the data storage layer. Each layer is designed with a well-defined responsibility and communicates with adjacent layers through standardized interfaces, promoting modularity and ease of maintenance [15].

The user interface layer accepts images, textual queries, and voice commands, providing a flexible and accessible entry point for healthcare workers across varying levels of technical proficiency. The application logic layer, implemented using the Flask web framework [23], handles request routing, input validation, and inter-component coordination. The AI processing layer constitutes the computational core of the system and encompasses the YOLOv8-based image classifier, the FAISS-powered retrieval engine, the LLM-based response generator, and the speech processing modules (DeepSpeech for speech-to-text and gTTS for text-to-speech). The data storage layer maintains a persistent repository of prior classification results, regulatory documents, and model artifacts required for continuous system operation.

Information flow within the system follows a deterministic path. Upon receiving a user input, the system routes it to the appropriate processing module based on modality. Image inputs are forwarded to the YOLOv8 model, which performs real-time waste detection and returns a classified label with a confidence score. Voice inputs are transcribed using DeepSpeech before being treated identically to text inputs. The resulting text is used to formulate a query for the RAG pipeline, within which FAISS performs semantic similarity search against the regulatory document index [18]. Retrieved passages are provided to the LLM, which synthesizes a concise, compliant disposal instruction.

The final output, comprising the waste classification label and associated disposal guidance, is rendered in both text and audio formats. Users may iteratively refine their queries to obtain more specific responses. The system is engineered for robustness in challenging clinical environments through rigorous input validation and structured error handling throughout the pipeline. The platform is deployable in real-time within hospital and clinic settings and is designed for straightforward extensibility as regulatory frameworks evolve [19].

Figure 2 presents the detailed system architecture, illustrating the interaction pathways between the frontend interface (React.js), the backend server (Flask), the AI processing components (YOLOv8, embedding model, FAISS vector database, LLM), and the output module.

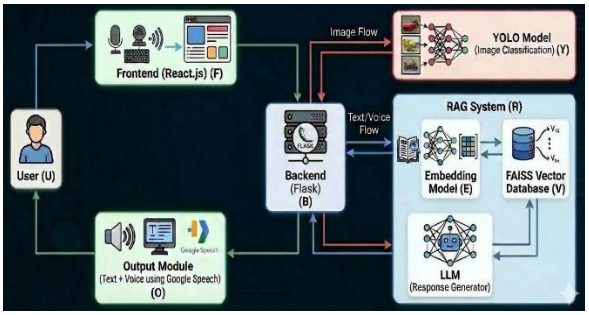

Figure 2. Detailed system architecture of MedWaste.AI.

V. RESULTS

The YOLOv8 model's classification performance across the target waste categories—including IV tubes, bandages, cotton swabs, gloves, face masks, syringes, and vials—is summarized in the normalized confusion matrix presented in Figure 3 [17]. The matrix displays predicted labels against ground truth labels; correct classifications appear along the principal diagonal. The model demonstrates particularly high accuracy for bandages, medical caps, vials, and masks, where misclassification rates are negligible.

Off-diagonal entries reveal occasional confusion between visually similar categories such as cotton swabs and gloves, where surface texture and shape similarity contribute to classification ambiguity. In some instances, multi-item scenes result in unconfident predictions, typically when input images are of poor quality or contain overlapping objects. Nonetheless, the overall classification performance is high, and the confusion matrix serves as a constructive diagnostic tool for identifying areas of

MedWaste.AI was subjected to comprehensive testing to evaluate its performance in waste classification and disposal guidance generation. The YOLOv8 detection module demonstrated high accuracy in identifying waste types even under suboptimal imaging conditions, producing consistent results across varied input quality levels [17]. The system's end-to-end efficiency was notable, with negligible latency between input submission and output delivery.

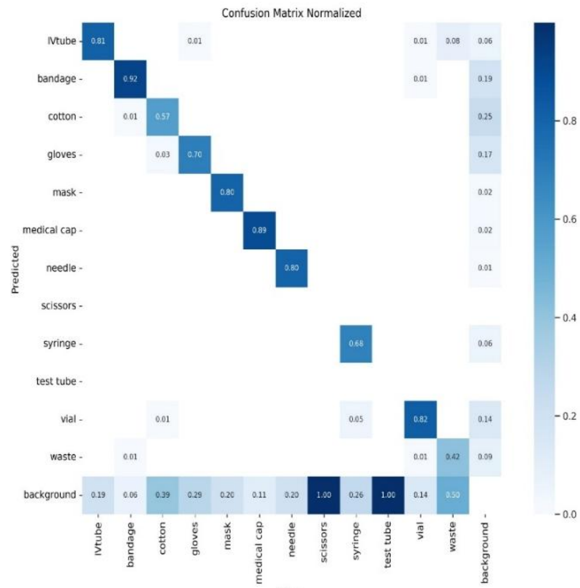

Figure 3. Normalized confusion matrix for YOLOv8 waste classification across all target categories.

The RAG pipeline exhibited strong performance in retrieving relevant regulatory content from the indexed knowledge base using the FAISS similarity search method [18]. The LLM subsequently utilized retrieved passages to generate concise, regulation-consistent disposal instructions that were evaluated by domain experts as appropriate and actionable for healthcare practitioners [21].

The platform successfully processed all three input modalities. DeepSpeech accurately transcribed audio queries, while gTTS converted textual model outputs into natural-sounding audio [22]. This multimodal capability proved particularly advantageous for serving users in multilingual healthcare settings, where spoken interaction in a native language reduces cognitive friction and the likelihood of compliance errors.

System outputs consistently included the identified waste category, an associated confidence score, and step-by-step disposal instructions aligned with applicable regulatory standards [19]. Response times were sufficiently rapid for real-time clinical use. User interaction with the platform was straightforward, confirming that MedWaste.AI is a viable tool for deployment within hospital and clinical environments to support compliant biomedical waste management workflows.

VI. CONCLUSION

This paper has presented MedWaste.AI, an intelligent platform for the automated identification and regulatory-compliant disposal guidance of biomedical waste. The system integrates computer vision, natural language processing, and speech technologies to deliver a robust, multimodal waste management solution tailored to the demands of busy healthcare environments.

By leveraging deep learning for image-based waste classification and a RAG architecture for regulatory information retrieval, MedWaste.AI substantially reduces the manual effort associated with waste sorting decisions, minimizes human error, and provides frontline staff with immediate access to applicable disposal guidelines. The platform's voice interaction capability, with support for multiple languages, further broadens its accessibility and practical utility across diverse clinical settings.

Future work will focus on several avenues for system enhancement. First, the integration of more advanced neural architectures and a larger, more diverse training dataset is expected to yield further improvements in classification accuracy, particularly for ambiguous or visually similar waste categories. Second, the incorporation of real-time waste tracking through networked IoT devices would enable continuous monitoring of waste generation patterns and support proactive regulatory reporting. Third, migration of the platform to a scalable cloud infrastructure would facilitate concurrent multi-user access across large healthcare institutions. Finally, expanding the regulatory knowledge base to encompass international waste management standards would extend the system's applicability to clinical settings beyond India

REFERENCES.

[1] S. Thung and C. T. Ng, "Classification of Trash for Recyclability Status," CS229 Project Report, Stanford University, 2016.

[2] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-T. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 33, pp. 9459–9474, 2020.

[3] M. S. Islam, M. Sultana, M. F. Ahmed, and M. M. Islam, "Automatic Waste Classification Using Deep Learning Convolutional Neural Network," in Proc. Int. Conf. Comput. Inf. Technol. (ICCIT), pp. 1–6, 2021, doi: 10.1109/ICCIT54785.2021.9689873.

[4] World Health Organization (WHO), Healthcare Waste, Fact Sheet, Aug. 2022.

[5] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proc. IEEE Conf. Comput. Vision Pattern Recognit. (CVPR), pp. 779–788, 2016, doi: 10.1109/CVPR.2016.91.

[7] N. Hossain, M. J. Islam, and H. Wang, "Multimodal Human–Computer Interaction in Clinical Environments: A Survey," IEEE Access, vol. 9, pp. 93273–93295, 2021, doi: 10.1109/ACCESS.2021.3092847.

[8] B. D. Nkosi, M. B. Mthembu, and H. Dlamin, "Environmental Impact of Improper Healthcare Waste Management: A Review," J. Environ. Manage., vol. 319, p. 115666, 2022, doi: 10.1016/j.jenvman.2022.115666.

[9] M. A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," IEEE Commun. Surv. Tutor., vol. 17, no. 4, pp. 2347–2376, 2015, doi: 10.1109/COMST.2015.2444095.

[10] S. King and J. Karaiskos, "The Blizzard Challenge 2020," in Proc. Blizzard Challenge Workshop, 2020.

[11] L. Laranjo, A. G. Dunn, H. L. Tong, A. B. Kocaballi, J. Chen, R. Bashir, D. Bhaskara, S. Zwaan, N. Bhavya, and E. Coiera, "Conversational Agents in Healthcare: A Systematic Review," J. Am. Med. Inform. Assoc., vol. 25, no. 9, pp. 1248–1258, 2018, doi: 10.1093/jamia/ocy072.

[12] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent Trends in Deep Learning Based Natural Language Processing," IEEE Comput. Intell. Mag., vol. 13, no. 3, pp. 55–75, 2018, doi: 10.1109/MCI.2018.2840738.

[13] G. Jocher, A. Chaurasia, and J. Qiu, "Ultralytics YOLOv8," GitHub, 2023, doi: 10.5281/zenodo.7347926.
[14] S. Nanda and R. Panda, "A Comprehensive Review on AI-Based Smart Environmental Monitoring Systems," Environ. Sci. Pollut. Res., vol. 29, pp. 56789–56810, 2022, doi: 10.1007/s11356-022-20180-7.
[15] R. T. Fielding and R. N. Taylor, "Principles and Architectural Styles for Network-Based Software Architectures," Ph.D. dissertation, Univ. California, Irvine, CA, USA, 2000.
[16] V. Panayotov, G. Chen, D. Povey, and S. Khudanpur, "LibriSpeech: An ASR Corpus Based on Public Domain Audio Books," in Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP), pp. 5206–5210, 2015, doi: 10.1109/ICASSP.2015.7178964.
[17] Ultralytics, YOLOv8 Documentation, 2023.
[18] J. Johnson, M. Douze, and H. Jégou, "Billion-Scale Similarity Search with GPUs," IEEE Trans. Big Data, vol. 7, no. 3, pp. 535–547, 2021, doi: 10.1109/TBDATA.2019.2921572.
[19] Government of India, Ministry of Environment, Forest and Climate Change, Biomedical Waste Management Rules, 2016, Gazette of India, 2016.
[20] World Health Organization (WHO), Safe Management of Wastes from Health-Care Activities, 2nd ed., Y. Chartier, J. Emmanuel, U. Pieper, A. Prüss, P. Rushbrook, R. Stringer, W. Townend, S. Wilburn, and R. Zghondi, Eds. Geneva, Switzerland: WHO Press, 2014.
[21] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell et al., "Language Models Are Few-Shot Learners," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 33, pp. 1877–1901, 2020.
[22] Google, gTTS (Google Text-to-Speech): A Python Library and CLI Tool, PyPI, 2023.
[23] A. Ronacher, Flask: A Lightweight WSGI Web Application Framework. Pallets Projects, 2023.
[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE Conf. Comput. Vision Pattern Recognit. (CVPR), pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.