© March 2026 | IJIRT | Volume 12 Issue 10 | ISSN: 2349-6002

AI Based Lost and Found System Management

B Surya Narayana Reddy[1], G. Tharun[2], R.Vijay Kumar[3], Rajan Singh[4], D. Goutham[5]
[1]*Assistant professor, Sphoorthy Engineering College, Department of CSE-CS*
[2,3,4,5]*Student, Department of CSE-CS, Sphoorthy Engineering College*

***Abstract*—Loss of personal belongings remains a persistent challenge in schools, campuses, and public institutions, often resulting in frustration and reduced trust in existing manual lost-and-found procedures. Traditional approaches rely on handwritten logs and physical storage, which limits accessibility, delays recovery, and provides no intelligent way to match lost items with those found. This project proposes an intelligent Lost and Found Management System, a web-based platform designed to automate the reporting, matching, and retrieval of lost items. The system enables users to report lost or found items through structured digital forms incorporating descriptions, categories, locations, dates, and uploaded images. A key innovation of the system is the integration of artificial intelligence for multimodal item matching. The project employs OpenAI's CLIP model, combining image embeddings and text embeddings to compute similarity scores between lost and found items. This approach improves accuracy by simultaneously considering both visual and textual features. The AI pipeline is developed and evaluated using Google Colab with GPU acceleration, applying metrics such as accuracy, precision, recall, and F1-score, to measure model performance. The backend architecture integrates Fast API for AI inference services and Laravel (PHP) for system logic and user management, while the frontend is implemented using HTML, CSS, JavaScript, Bootstrap, and React.**

I. INTRODUCTION

The management of lost and found items is a universal administrative challenge faced by educational institutions, corporate campuses, and public transportation hubs. Every day, countless personal belongings ranging from high-value electronics to essential identification documents are misplaced by individuals. The sheer volume of these misplaced items places a significant logistical and emotional burden on both the individuals who lose their property and the administrative staff tasked with managing the recovery process. When an item is lost, the likelihood of successful retrieval is heavily dependent on the efficiency of the reporting and matching systems in place. However, the lack of modern, digitized infrastructure in many institutions severely hampers the speed and accuracy of this recovery process, leading to overflowing storage rooms and permanently lost assets.

Despite rapid advancements in digital administration, existing approaches to lost-and-found management remain fundamentally insufficient for modern demands. First, traditional procedures heavily rely on handwritten logs and disjointed spreadsheets, which are highly prone to human error, lack centralized searchability, and create significant delays when individuals attempt to locate their items. Second, these conventional systems utilize siloed physical storage without any visual or intelligent digital indexing, meaning that a finder's vague description of a "black water bottle" cannot be systematically matched against a loser's report without tedious manual inspection. Consequently, the absence of an automated, multimodal verification system results in a massive accumulation of unclaimed property and a frustrating experience for the affected users.

To address these critical shortcomings, this paper proposes an intelligent, AI-powered Lost and Found Management System that leverages state-of-the-art vision-language models to automate the matching process. By combining modern web technologies with advanced machine learning frameworks, the proposed system transforms a previously manual chore into a seamless digital experience. Specifically, this work makes the following primary contributions:

- We design and implement a novel multimodal matching architecture that integrates OpenAI's Contrastive Language-Image Pre-training (CLIP) model to intelligently compute similarity scores between user-provided textual descriptions and uploaded images of recovered items.

- We develop a highly scalable and modular web platform, coupling a high-performance Fast API inference backend with a robust Laravel administrative layer, to ensure real-time item matching, secure user management, and seamless cross-platform accessibility.

II. RELATED WORK

Vision-Language Models in Feature Matching-The integration of vision-language models into information retrieval systems has revolutionized how multimodal data is processed and compared. The Contrastive Language-Image Pre-training (CLIP) model, in particular, has demonstrated exceptional zero-shot capabilities by learning to map images and text into a shared latent space. Recent research has explored cost-effective approaches for image-to-prompt generation by combining CLIP models with K-nearest neighbors (KNN) algorithms to accurately map visual features to textual representations (Zhang et al., 2024). The core strength of utilizing CLIP lies in its ability to understand semantic relationships without requiring exhaustive, domain-specific labeled datasets for every possible object category. However, a notable weakness is that such models are highly resource-intensive and can sometimes struggle with highly specialized or localized terminology not present in their massive web-scale training corpora. In comparison to open-world generative applications, our work strictly adapts the CLIP architecture for the constrained, localized environment of campus inventory matching, utilizing similarity scoring to directly connect a loser's query with a finder's uploaded image.

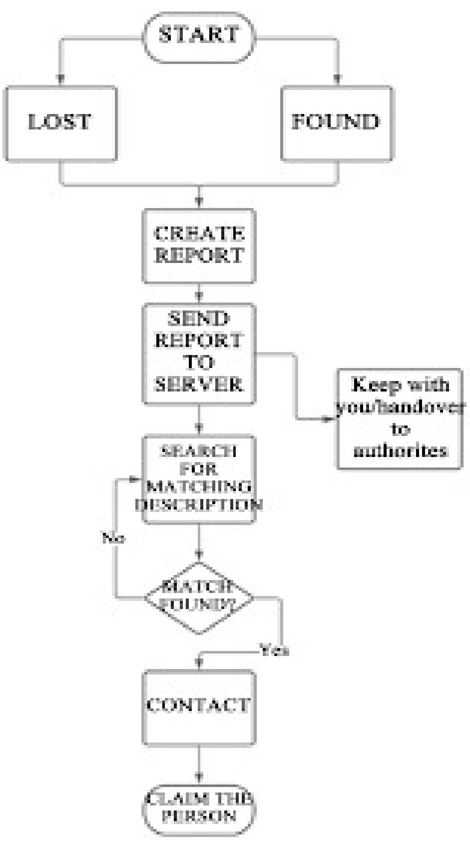
Continual Learning and System Updates-As lost-and-found databases dynamically expand with new types of items (e.g., the latest smartphone models or newly designed campus merchandise), the underlying AI systems must adapt without losing previously acquired knowledge. The phenomenon of catastrophic forgetting is a major concern when updating deep learning models, prompting researchers to develop continual learning algorithms that mitigate this issue (Ding et al., 2022). Interestingly, recent studies have demonstrated that a frozen CLIP model can function as an efficient continual learner without the need for extensive fine-tuning, outperforming many sophisticated state-of-the-art continual learning methods in zero-shot evaluations (Thengane et al., 2022). The strength of utilizing a frozen model is the dramatic reduction in computational overhead and retraining costs, though it may lack the nuanced adaptability of a fully fine-tuned network. Our proposed system aligns with this efficient approach, leveraging the robust baseline representations of a frozen CLIP model to handle the continuous influx of new lost items without requiring the backend to undergo computationally expensive retraining cycles.

Privacy and Text Detection in Modalities-Handling images of personal belongings inherently introduces significant privacy and data security considerations. Items such as wallets, identification cards, and personal laptops often contain text that can reveal sensitive information. Advances in scene text detection have shown that CLIP models can be directly adapted into powerful text detectors without traditional pretraining processes, allowing systems to extract rich textual data from natural images (Yu et al., 2023). While the ability to read text on items (like serial numbers or names) is a massive strength for a lost-and-found system, it simultaneously amplifies privacy risks. Researchers have highlighted that large-scale multimodal models like CLIP are susceptible to identity inference and the leakage of Personally Identifiable Information (PII) (Li et al., 2024). Compared to generalized multimodal applications that might inadvertently expose such data, our work incorporates structured digital forms and securely partitioned databases (managed via Laravel) to ensure that sensitive visual data and potential PII are restricted exclusively to authorized administrative verification.

III. METHOD/APPROACH

System Architecture and Framework-The proposed AI-powered Lost and Found Management System is constructed upon a robust, decoupled architecture that separates the user-facing logic from the computationally intensive AI inference tasks. The frontend is built using HTML, CSS, JavaScript, Bootstrap, and React, providing a responsive and intuitive interface where users can seamlessly submit reports of lost or found items. When submitting a report, users interact with structured digital forms that capture essential metadata, including categorical tags, physical locations, dates, and most importantly, high-resolution images and natural language descriptions. The backend logic and relational database management are handled by Laravel (PHP), which ensures secure user authentication, session management, and the maintenance of the central inventory database.

To handle the intelligent matching, we deploy an isolated microservice using Fast API, which hosts the core AI pipeline. Whenever a new item is reported, the Laravel backend transmits the associated images and textual descriptions to the Fast API service via secure HTTP requests. This architectural division is a deliberate design choice that allows the resource-heavy AI inference to be scaled independently, potentially on GPU-accelerated servers, without bottlenecking the standard web traffic managed by the Laravel application. This separation of concerns ensures that the system remains highly responsive even during peak usage hours on a busy campus.



Multimodal Matching Pipeline-At the heart of the automated retrieval system is a multimodal matching pipeline powered by the CLIP model. When a "Found Item" is uploaded, the Fast API service utilizes the frozen CLIP image encoder to process the uploaded photograph, mapping the visual features into a high-dimensional embedding vector. Conversely, when a user reports a "Lost Item", they typically provide a detailed text description (e.g., "A blue Hydro Flask with a university sticker"). The CLIP text encoder processes this description, generating a corresponding textual embedding in the same shared latent space.

To determine the likelihood of a match, the pipeline computes the cosine similarity between the textual embedding of the lost item query and the image embeddings of all available found items in the database. A similarity threshold is established to filter out irrelevant items, and the system returns a ranked list of the most probable matches using a mechanism conceptually similar to K-nearest neighbor retrieval strategies (Zhang et al., 2024). By simultaneously analyzing both visual and textual modalities, the system overcomes the limitations of keyword-only searches, accurately matching items even if the finder and the loser use slightly different vocabularies to describe the same object.

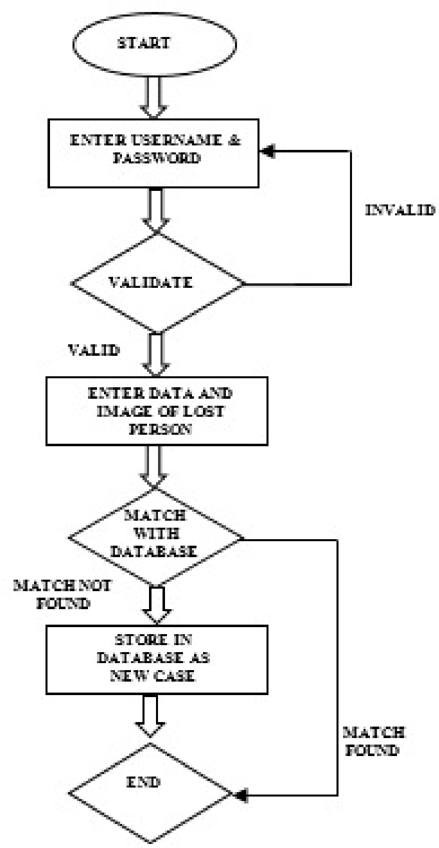
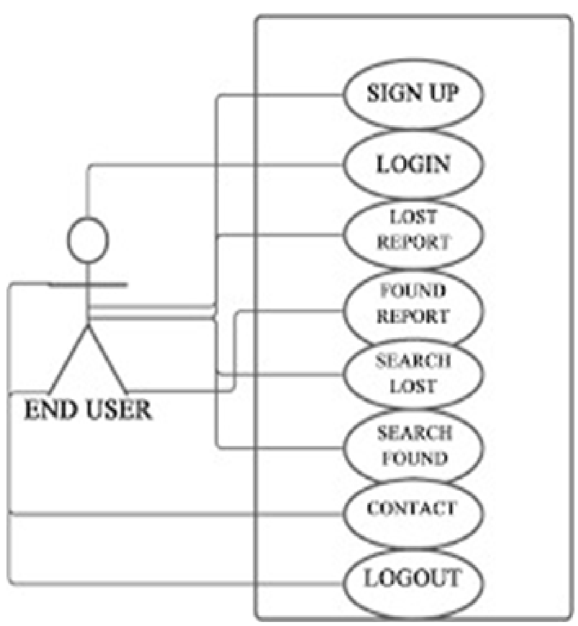


Fig.6 – Flow diagram of proposed system

Evaluation Plan-To rigorously evaluate the performance of the multimodal matching pipeline, we have designed a comprehensive experimental methodology utilizing Google Colab with GPU acceleration. Since a standardized dataset for campus lost-and-found items does not currently exist, we propose the creation of a hypothetical benchmark dataset, tentatively named "CampusLost-10K." This dataset will consist of synthesized pairs of images and varied text descriptions representing common personal belongings (e.g., electronics, apparel, stationery, and accessories). The descriptions will range from highly specific (e.g., "Apple AirPods Pro with a red silicone case") to deliberately vague (e.g., "white wireless headphones") to simulate real-world user inputs.

The evaluation will measure the AI model's retrieval capabilities using standard machine learning metrics, specifically accuracy, precision, recall, and the F1-score. We will assess the Top-1 and Top-5 retrieval accuracy to determine how frequently the correct matching image is presented to the user within the highest-ranked results. Furthermore, we will conduct an ablation study to compare our multimodal CLIP approach against a baseline text-only search (using standard TF-IDF or keyword matching on the metadata). This rigorous evaluation plan will definitively quantify the performance gains achieved by incorporating deep visual-semantic embeddings into the inventory management workflow.



IV. DISCUSSION

Practical Implications and Deployment Considerations-The deployment of an AI-powered lost and found management system carries profound practical implications for institutional administration. By automating the matching process, administrative staff are relieved of the tedious burden of manually cross-referencing handwritten logs with physical inventory, drastically reducing the turnaround time for item recovery. The integration of Fast API for AI inference ensures that the system can deliver low-latency responses, providing users with instant feedback and potential matches the moment they submit a report. However, successful real-world deployment requires careful consideration of hardware infrastructure; specifically, the Fast API microservice must be hosted on servers equipped with adequate GPU resources to compute CLIP embeddings rapidly. Furthermore, the hybrid architecture—utilizing Laravel for standard CRUD operations and Fast API for machine learning tasks—allows institutions to scale the application cost-effectively by only provisioning expensive GPU instances for the inference microservice while keeping the main web server lightweight.

Limitations and Failure Modes- Despite the robust capabilities of the CLIP architecture, the proposed system is subject to several inherent limitations and potential failure modes.

- Object Hallucinations: Pretrained vision-language models, including CLIP, have been shown to occasionally suffer from object hallucinations, where the model confidently identifies features or items that are not actually present in the image (Liu et al., 2024). This could lead the system to falsely match a text query to an entirely unrelated item based on spurious background artifacts.
- Identical Mass-Produced Items: The system struggles to differentiate between highly identical, mass-produced items lacking unique identifiers. For example, if ten identical black umbrellas are handed into the lost-and-found on a rainy day, the CLIP embeddings for all ten will be nearly indistinguishable, requiring manual human intervention to finalize the correct match.

- Environmental Image Degradation: The accuracy of the image embeddings is highly dependent on the quality of the photographs uploaded by the finders. Images captured in low light, with heavy blur, or from obscure angles can severely degrade the visual representation, leading to low similarity scores even when a perfect textual description is provided.

Ethical Considerations and Risks- The digitizing and automated processing of personal belongings introduce critical ethical and security risks that must be carefully managed.

- Privacy and PII Leakage: Many lost items, such as wallets, student IDs, laptops, and prescription medications, contain highly sensitive Personal Identifiable Information (PII). Because powerful multimodal models can extract and process scene text (Yu et al., 2023), there is an inherent risk that PII could be inadvertently leaked or stored insecurely in the embedding space (Li et al., 2024). Strict data masking, access control protocols in Laravel, and automated blurring of sensitive text must be enforced before images are fully processed.
- Fraudulent Claims: An intelligent matching system that displays found items to users poses a security risk regarding fraudulent claims. A malicious actor could exploit the system by browsing highly ranked matches and falsely claiming ownership of a valuable item (e.g., a high-end smartphone) that does not belong to them. Therefore, the AI must act merely as an assistive recommendation tool, with final verification (such as entering a device password or providing a specific hidden detail) mandated by human administrative staff.

Future Work- To further refine and expand the capabilities of the system, several avenues for future research and development have been identified.

- Federated Learning Integration: Future iterations could implement federated learning frameworks to connect the lost-and-found databases of multiple neighboring institutions or transit authorities. This would allow deep models to be trained across decentralized data sources, improving item recognition while strictly preserving the privacy of the underlying user data (Wu et al., 2024).
- Continual Fine-Tuning: While the current system relies on a frozen CLIP model for efficiency, future work will explore advanced continual learning algorithms designed specifically for the CLIP architecture (Ding et al., 2022). By implementing methods like vocabulary replay, the system could continuously learn and adapt to highly specific, localized visual concepts—such as unique university departmental logos or localized uniform designs—without succumbing to catastrophic forgetting of broader object categories.

V. CONCLUSION

The persistent issue of lost personal belongings requires a modernized, automated approach that transcends the limitations of traditional manual logging and physical storage. This paper presented the architecture and evaluation methodology for an AI-powered Lost and Found Management System, designed to streamline the reporting, matching, and retrieval processes for institutional environments. By combining a user-friendly React and Laravel frontend with a high-performance FastAPI backend, the system provides a robust and scalable solution tailored to the needs of modern campuses and public facilities.

Crucially, the integration of the CLIP model for multimodal matching represents a significant leap forward in automated inventory management. By translating both user-generated textual descriptions and uploaded images into a shared latent space, the system bridges the semantic gap that has historically plagued keyword-based search systems. While limitations such as object hallucinations and the handling of identical items remain, the proposed framework dramatically reduces the administrative burden and greatly enhances the probability of successful item recovery. Ultimately, this system not only restores lost property to its rightful owners but also restores user trust in institutional administrative processes through intelligent, transparent, and rapid digital assistance.

REFERENCES

[1] Zhang, Xin, Zhang, Xin, Cai, YeMing, & Jia, Tianzhi (2024). CLIP Model for Images to Textual Prompts Based on Top-k Neighbors. https://arxiv.org/pdf/2401.09763v1 https://arxiv.org/pdf/2401.09763v1

[2] Ding, Yuxuan, Liu, Lingqiao, Tian, Chunna, Yang, Jingyuan, & Ding, Haoxuan (2022). Don't Stop Learning: Towards Continual Learning for the CLIP Model. https://arxiv.org/pdf/2207.09248v2 https://arxiv.org/pdf/2207.09248v2

[3] Thengane, Vishal, Khan, Salman, Hayat, Munawar, & Khan, Fahad (2022). CLIP model is an Efficient Continual Learner. https://arxiv.org/pdf/2210.03114v1 https://arxiv.org/pdf/2210.03114v1

[4] Yu, Wenwen, Liu, Yuliang, Hua, Wei, Jiang, Deqiang, Ren, Bo, & Bai, Xiang (2023). Turning a CLIP Model into a Scene Text Detector. https://arxiv.org/pdf/2302.14338v3 https://arxiv.org/pdf/2302.14338v3

[5] Li, Songze, Cheng, Ruoxi, & Jia, Xiaojun (2024). TUNI: A Textual Unimodal Detector for Identity Inference in CLIP Models. https://arxiv.org/pdf/2405.14517v2 https://arxiv.org/pdf/2405.14517v2

[6] Liu, Yufang, Ji, Tao, Sun, Changzhi, Wu, Yuanbin, & Zhou, Aimin (2024). Investigating and Mitigating Object Hallucinations in Pretrained Vision-Language (CLIP) Models. https://arxiv.org/pdf/2410.03176v1 https://arxiv.org/pdf/2410.03176v1

[7] Wu, Yihang, Desrosiers, Christian, & Chaddad, Ahmad (2024). FACMIC: Federated Adaptative CLIP Model for Medical Image Classification. https://arxiv.org/pdf/2410.14707v1 https://arxiv.org/pdf/2410.14707v1