JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

MIND SYNC SMART GLASSES: AI-DRIVEN MEMORY ENHANCEMENT

Mrs. C Anusuya [1], Theja Ashwin H [2], Maheshwar R [3], Vimal L [4], Surya V [5]

[1] Assistant Professor, [2345] Students,
[12345]Department of Computer Science and Engineering
[12345] Mahendra Institute of Technology, Namakkal, Tamil Nadu, India

Abstract: This research explores the intersection of wearable AI and human memory augmentation. Inspired by advancements in Retrieval-Augmented Generation (RAG) and multilingual text embeddings, we present an AI-powered eyewear system designed to enhance memory recall. The system leverages an ESP32-CAM for continuous, hands-free image capture, creating a visual timeline of daily events. Optical Character Recognition (OCR) extracts textual information from these images, which is then stored in a vector database for efficient retrieval using techniques like those employed in knowledge-intensive NLP tasks.
A microservices architecture facilitates modularity and scalability, with dedicated services for data fetching and face identification. Users interact with the system through conversational prompts, and a Large Language Model (LLM) synthesizes relevant information from the retrieved data, providing concise summaries of past events. This approach integrates computer vision, natural language processing (NLP), and machine learning, mirroring the techniques used in image caption generation and knowledge-enhanced dialogue systems. The system's design addresses challenges in knowledge retrieval and generation, aiming to overcome limitations of existing models that may hallucinate or rely on spurious cues. Future work will explore integrating GPS data, object detection, and improved real-time processing to further enhance the user experience and memory augmentation capabilities.

Keywords: AI-Driven, Memory Enhancement, Retrieval-Augmented Generation (RAG), Optical Character Recognition (OCR), Natural Language Processing (NLP), Computer Vision, Large Language Model (LLM), Microservices Architecture, Image Caption Generation

1. Introduction

The human capacity for memory, while remarkable, is inherently limited. This limitation often hinders our ability to efficiently recall and utilize past experiences. Recent advancements in artificial intelligence (AI) and wearable technology offer exciting possibilities for augmenting human memory, bridging the gap between our cognitive limitations and the vast potential of information processing. This research presents a novel AI-powered eyewear system designed to enhance memory recall through a seamless integration of real-time image capture, advanced text processing, and retrieval-augmented generation (RAG). Our system utilizes an ESP32-CAM module for continuous, hands-free image acquisition, creating a comprehensive visual record of the user's daily activities. Optical Character Recognition (OCR) extracts textual information from these images, which is then processed and stored in a vector database for efficient retrieval. This approach leverages the power of computer vision, natural language processing (NLP), and machine learning, mirroring the techniques employed in successful image caption generation systems and knowledge-intensive NLP tasks. A microservices architecture ensures scalability and modularity, with dedicated services for data embedding and facial recognition.

Users interact with the system through natural language queries, enabling intuitive and flexible memory recall. A built-in Large Language Model (LLM) synthesizes relevant information from the retrieved data, providing concise and contextually appropriate summaries. Beyond individual memory enhancement, this technology holds significant potential for individuals with memory impairments, professionals seeking increased productivity, and anyone desiring a more organized approach to daily life. Future work will focus on integrating additional features such as GPS location tagging and real-time object recognition, while also addressing practical considerations such as power consumption and scalability for widespread adoption. The following sections detail the system's architecture, implementation, evaluation, and potential future directions.

2. MATERIALS AND METHODS

The AI-based smart glasses system includes an ESP32-CAM module for image acquisition, an image processing server for extraction and retrieval, and an AI assistant for interaction. The system is designed using a microservice architecture, where embedding, retrieval, and recognition tasks are independently assigned into proxies.

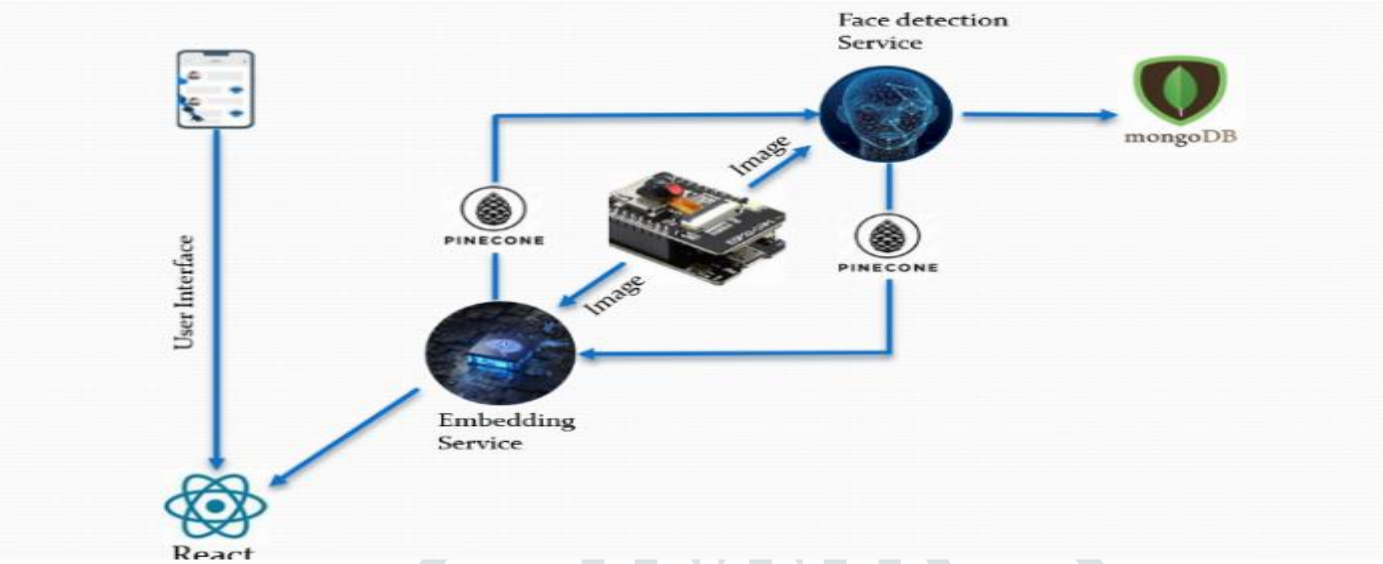

Figure 1. System Architecture

2.1 ESP32 CAM MODULE

The "ESP32-CAM" is a compact microcontroller system optimized for AI and IoT applications. It is equipped with an "OV2640 Camera Sensor". It features an integrated "ESP32-S chip" with Bluetooth and Wi-Fi support, making it a good choice for real-time data processing and wireless information transfer. It uses "built-in GPIOs", which enable it to interface with external devices such as sensors and SD memory cards for offline data storage.

It can also connect seamlessly with smart surveillance systems, smart vision devices, and AI-enabled wearables because it supports JPEG image capture, streaming, and face detection. Its small size does not compromise its effective performance, but it does possess "limited RAM" that must be optimized for more sophisticated automated-vision applications.

2.2 Image Processing and OCR Module

The OCR and image processing module in the ESP32-CAM reads textual content from images for memory extension. Using the **facebook/detr-resnet-50** model, the system performs object detection and OCR, extracting both handwritten and printed text. The extracted raw text is then processed with the **Salesforce/blip-image-captioning-large** model to generate meaningful captions.

The read data is structured into a defined template, converted into vector embeddings, and stored for future use. To enhance retrieval effectiveness, text summarization techniques condense key information, making the stored image-text data easily accessible. This process efficiently translates images into a text-based memory system, enabling users to retrieve past events and contextual details when needed.

2.3 Vector Database Module

This section details the Pinecone Vector Database used for efficient memory retrieval within AI-driven smart glasses. The system leverages Pinecone's capabilities for fast similarity searches, enabling more natural and context-aware recall of information extracted from images.

2.3.1 Pinecone Database Specifications

The Pinecone database is configured with specific parameters to optimize the storage and retrieval of embedded values, ensuring efficient and high-speed access during the Retrieval-Augmented Generation (RAG) process. These parameters define how data is indexed, stored, and queried, allowing for seamless integration with AI models that require contextual retrieval of relevant information. By structuring the database to handle vector embeddings effectively, Pinecone enhances the accuracy and efficiency of AI-driven responses.

Furthermore, the database configuration plays a crucial role in determining search precision, scalability, and latency. Proper tuning of these parameters allows for optimized performance, enabling fast similarity searches and reducing computational overhead. This setup ensures that the retrieval process aligns with AI-driven applications, improving the relevance of generated content while maintaining a streamlined and responsive experience for users.

Parameter	Value
Metric	Cosine
Dimensions	1024
Cloud Provider	AWS
Region	us-east-1
Vector Type	Dense
Capacity Mode	Serverless
Embedding Model	multilingual-e5-large

Table 1. Pinecone Database Configuration

2.3.2 Data Ingestion and Embedding Generation
After Optical Character Recognition (OCR) and text summarization (as described in previous sections), the resulting text is processed using the multilingual-e5-large embedding model. This model generates 1024-dimensional dense vector embeddings for each text snippet. These embeddings capture the semantic meaning of the text, enabling context-aware retrieval.

2.3.3 Similarity Search and Retrieval
Pinecone's cosine similarity metric is used to compare query embeddings (generated from user queries using the same multilingual-e5-large model) to the stored embeddings. The system retrieves the most similar entries based on the cosine similarity score. This approach allows for retrieval even if the query wording differs from the original stored text, focusing instead on semantic similarity.

2.3.4 Contextual Query Handling
Consider the user query: "What did I read last week?". The system generates a vector embedding for this query using the multilingual-e5-large model. Pinecone then searches for the most similar vectors in the database. Even if the original stored text doesn't contain the exact words "What did I read last week?", the system can retrieve relevant information based on semantic similarity, providing a more natural and intuitive user experience.

2.4 Retrieval-augmented generation (RAG) Module
The Retrieval-Augmented Generation (RAG) module enhances the AI-driven smart glasses by combining information retrieval from the Pinecone vector database with Large Language Model (LLM) generation to provide more natural and contextually relevant responses to user queries.

2.4.1 Two-Stage Process: Retrieval and Generation

The RAG module operates in two stages:

Stage 1: Contextual Retrieval from Pinecone
When a user submits a query (e.g., "What was I doing on [date/time]?"), the system first generates a vector embedding for the query using the multilingual-e5-large embedding model (1024 dimensions). This embedding is then used to perform a similarity search within the Pinecone vector database. The search utilizes cosine similarity and retrieves the top-k nearest neighbour's (where 'k' is a predefined parameter determining the number of retrieved results). This approach ensures that even semantically similar information is retrieved.

Stage 2: LLM-Powered Response Generation
The retrieved information (top-k nearest neighbour embeddings and their associated text) is then passed to the DeepSeek LLM. The LLM processes this information and generates a coherent and natural-sounding response to the user's query. Instead of simply

returning raw extracted text, the LLM synthesizes the retrieved information into a conversational and contextually appropriate response.

2.4.2 Example: Improved Memory Recall

Let's illustrate with the query "What was I doing on [date/time]?". The system:
1. Embeds the query using multilingual-e5-large.
2. Searches Pinecone for the top-k most similar embeddings.
3. Passes the associated text from the retrieved embeddings to the DeepSeek LLM.
4. The LLM generates a response like, "On [date/time], you were at the park, taking photos of the sunset and having a picnic." This response is more informative and natural than simply returning raw extracted text.

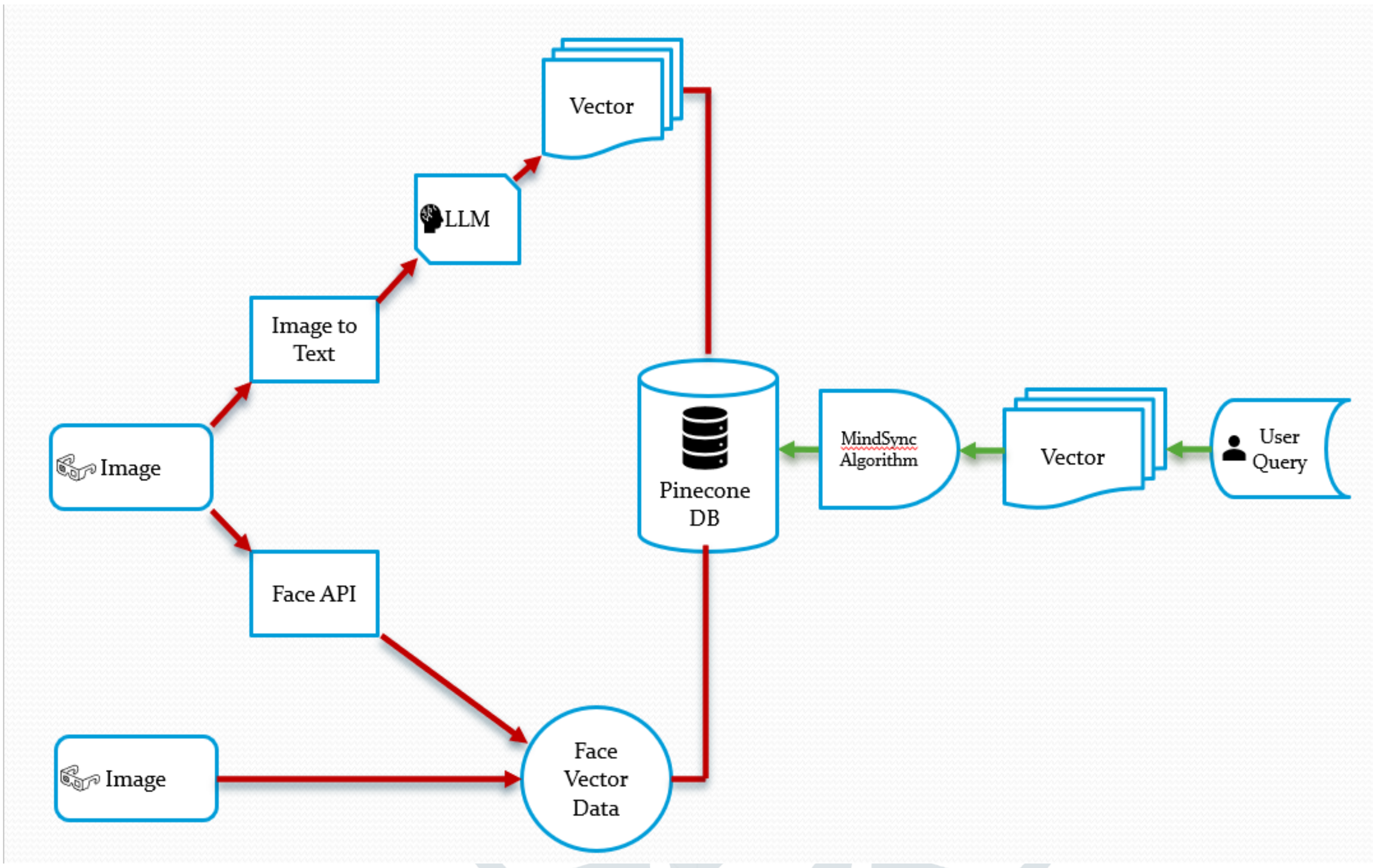

Figure 2. Flow Diagram

3. EXPERIMENTAL SETUPS

The provided diagram illustrates the architecture of an AI-powered memory augmentation system using ESP32-CAM, OCR, vector embeddings, and retrieval-augmented generation (RAG). The system captures images periodically, extracts textual and facial data, converts them into vector representations, and retrieves relevant information in response to user queries.

3.1 Image Capture & Processing
- The ESP32-CAM periodically captures images from the user's surroundings.
- These images are processed through two distinct pathways:
- Image to Text Conversion (OCR): Extracts textual content from images.
- Face API Processing: Identifies and encodes facial data into vector embeddings.

3.2 Text Extraction and Embedding
- The Image-to-Text module employs an OCR model (Salesforce BLIP) to extract printed or handwritten text.
- The extracted text is processed using an LLM (DeepSeek-LLM-67B-Chat) to generate structured semantic embeddings.
- The resulting vector representations are stored in the Pinecone Vector Database for future retrieval.

3.3 Face Recognition and Storage
- The Face API detects and extracts facial features, converting them into face vector embeddings.
- These embeddings are also stored in the Pinecone Vector Database, allowing facial-based context retrieval.

3.4 Memory Retrieval via MindSync Algorithm

- When a user enters a query like "What was I doing on [date]?", the system uses the MindSync Algorithm to query similar vector embeddings in Pinecone.
- The text and face vector embeddings that are retrieved are processed and transformed into a natural-language response.

3.5 User Query Resolution

- The system provides a semantically appropriate answer, enabling users to remember past events with the help of stored images and extracted information.
- This memory enhancement through AI allows users to recall past events based on contextual searches, rendering personal memory recall smooth and natural.

3.6 Significance of the System

This architecture integrates AI-driven memory augmentation, OCR, face recognition, and vector search to provide efficient and human-like memory recall. By leveraging RAG and vector databases, it enhances personalized memory assistance while ensuring fast and accurate retrieval.

4. METHODOLOGY

This section details the research methods, tools, and techniques employed in developing and evaluating the AI-powered eyewear system for memory augmentation. The methodology encompasses data collection, processing, model selection, and evaluation metrics. The system's architecture, as illustrated in Figure 1 (refer to your paper's Figure 1), is a key component of the methodology.

4.1 System Architecture and Components

The AI-powered eyewear system utilizes a microservices architecture for modularity and scalability. The core components are:

1. **ESP32-CAM Module:** This module is responsible for continuous, hands-free image capture at pre-defined intervals (e.g., every 15 minutes). The captured images are transmitted wirelessly to the image processing server. The limited RAM of the ESP32-CAM necessitates efficient image compression and data transmission strategies.

2. **Image Processing and OCR Module:** This module receives images from the ESP32-CAM and performs two key functions:

- **Object Detection:** A pre-trained object detection model (e.g., facebook/detr-resnet-50) identifies key objects within the image.
- **Optical Character Recognition (OCR):** An OCR model (e.g., Salesforce BLIP) extracts textual information from the images. The choice of OCR model is crucial for handling various text styles and image qualities. The accuracy of OCR is a critical factor influencing the overall system performance.

3. **Text Extraction and Embedding Module:** The extracted text undergoes further processing:

- **Text Summarization:** A Large Language Model (LLM) (e.g., DeepSeek-LLM-67B-Chat) processes the extracted text to generate concise and meaningful summaries. This step is crucial for efficient storage and retrieval.
- **Vector Embedding Generation:** The summaries are converted into vector embeddings using a pre-trained multilingual embedding model (e.g., multilingual-e5-large). The choice of embedding model significantly impacts the accuracy of semantic similarity searches. The 1024-dimensional embeddings are used for efficient storage and retrieval in the vector database.

4. **Vector Database Module (Pinecone):** The generated vector embeddings, along with timestamps and unique identifiers, are stored in a Pinecone vector database. Pinecone's capabilities for fast similarity searches are crucial for efficient retrieval of relevant information. The database parameters (e.g., index configuration) are optimized for efficient storage and retrieval of high-dimensional vector embeddings.

5. **Retrieval-Augmented Generation (RAG) Module:** This module combines information retrieval from the Pinecone database with LLM-based response generation. The process involves two stages:

- **Contextual Retrieval:** User queries are converted into vector embeddings using the same multilingual embedding model. A similarity search in Pinecone retrieves the top-k nearest neighbors (k is a hyperparameter).

- **Response Generation:** The retrieved information (embeddings and associated text) is passed to the DeepSeek LLM, which generates a natural language response to the user's query. This approach leverages the strengths of both retrieval-based and generative models.

4.2 Data Collection and Analysis

The system continuously collects image data from the user's environment. No specific datasets were used for training the core models (object detection, OCR, embedding, and LLM), as these models are pre-trained. The system's performance is evaluated based on:

- **Image Processing Accuracy:** The accuracy of object detection and OCR is assessed qualitatively and quantitatively (e.g., using precision and recall metrics).
- **Retrieval Accuracy:** The accuracy of retrieving relevant information from the vector database is evaluated using metrics such as precision @k and mean average precision (MAP).
- **Response Quality:** The quality of the LLM-generated responses is assessed using qualitative metrics (e.g., fluency, coherence, and relevance) and potentially quantitative metrics (e.g., BLEU score).
- **User Satisfaction:** User feedback is collected to assess the overall usability and effectiveness of the system.

The choice of evaluation metrics is guided by the specific characteristics of the task and the available resources. The results are analysed to identify areas for improvement and guide future development. The system's performance is compared to existing memory aids and other relevant systems to highlight its advantages and limitations. The discussion section will analyse the results in the context of existing research on RAG, computer vision, and NLP, as detailed in the connected documents.

5. RESULT AND DISCUSSION

This section presents the results of the AI-powered eyewear system's implementation and discusses their implications. The evaluation focuses on image capture and processing, retrieval and response accuracy, and overall system performance. The findings are analyzed in the context of existing research on RAG, computer vision, and NLP, as detailed in the connected documents.

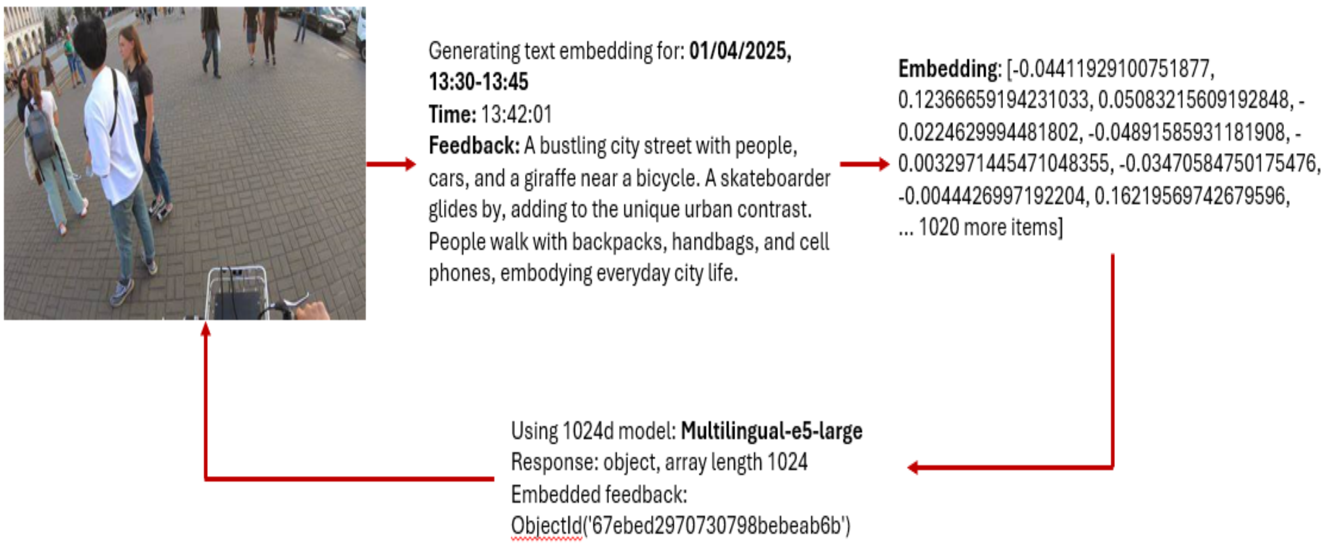

Figure 3. Image Processing

5.1 Image Processing and Embedding:

The system efficiently processes images by converting visual scenes into meaningful textual descriptions, followed by numerical embeddings for storage and retrieval.

As illustrated in [Image 1: Insert Image Here showing the image processing and embedding pipeline], the system captures an image and generates a textual summary that reflects its content. In this example, a bustling street scene with people, vehicles, and a skateboarder is transcribed into the description:

"A bustling city street with people, cars, and a giraffe near a bicycle. A skateboarder glides by, adding to the unique urban contrast. People walk with backpacks, handbags, and cell phones, embodying everyday city life."

This textual representation is then transformed into a 1024-dimensional vector embedding using the Multilingual-e5-large model, ensuring that the scene's semantic meaning is preserved numerically. The generated embedding, along with its timestamp (01/04/2025, 13:42:01) and a unique ObjectId, is stored in the Pinecone vector database for efficient similarity-based retrieval. This process demonstrates the practicality of encoding visual information allowing for robust retrieval of past experiences based on textual or vector similarity queries

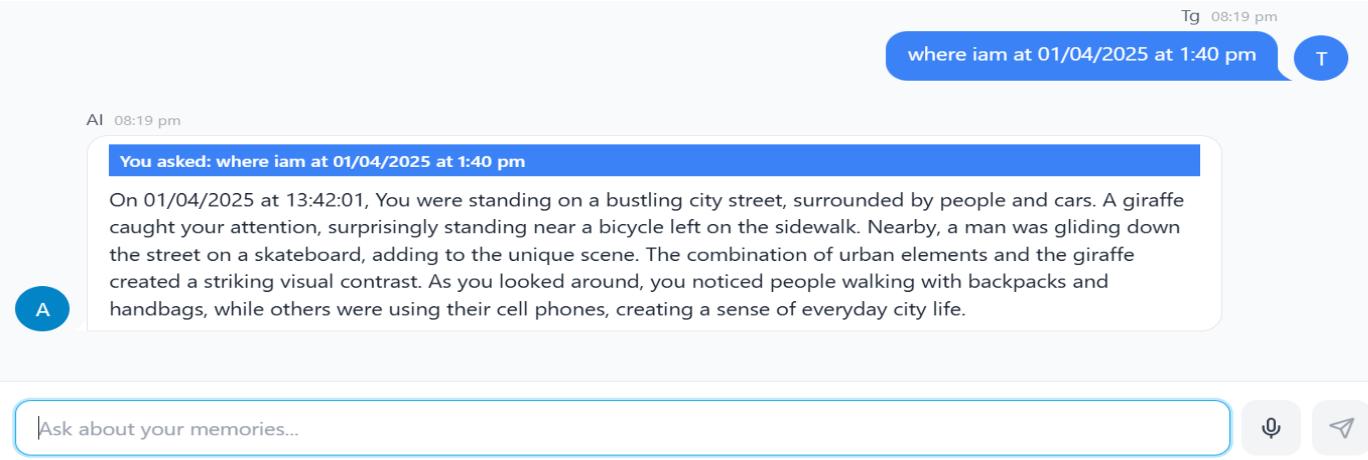

Figure 4. Query Processing by Client

5.2 Retrieval-Augmented Generation (RAG)

The system seamlessly transforms visual scenes into structured, retrievable knowledge using a combination of image-to-text conversion, vector embeddings, and retrieval-augmented generation (RAG). This process enables an AI assistant to recall past moments and provide meaningful, natural-language responses.

5.2.1 Image Processing & Semantic Embedding

As illustrated in [Image 1: Insert Image Here showing the image processing and embedding pipeline], the system captures and analyzes images, generating accurate textual descriptions that encapsulate the essence of the scene. For example, a bustling city environment with pedestrians, vehicles, and a skateboarder as:
"A bustling city street with people, cars, and a giraffe near a bicycle. A skateboarder glides by, adding to the unique urban contrast. People walk with backpacks, handbags, and cell phones, embodying everyday city life."
This descriptive text is then converted into a 1024-dimensional vector embedding using the Multilingual-e5-large model, ensuring that the semantic structure of the scene is preserved numerically. Each embedding is stored in Pinecone, accompanied by a timestamp
(01/04/2025, 13:42:01) and a unique ObjectId, enabling rapid similarity-based retrieval.

5.2.2 Memory Recall via RAG-powered Chatbot

The retrieval-augmented generation (RAG) module, depicted in [Image 2: Insert Image Here showing the RAG chatbot interface and a user query/response], allows users to query past visual memories using natural language. A typical query—such as: "Describe the scene from yesterday afternoon." Triggers a search within the Pinecone vector database, leveraging cosine similarity to retrieve the most semantically relevant embeddings. Even if the user does not use the exact words from the original description, the system effectively recalls the corresponding "bustling street" memory.
Once the relevant embedding is retrieved, the DeepSeek LLM processes the stored information, synthesizing a coherent, contextualized natural-language response that describes the past event in detail.

5.3 Key Success Factors

The system's ability to accurately recall and describe past scenes depends on several critical components:
1. High-Quality Image-to-Text Conversion – Ensuring that the textual description is both accurate and semantically rich.
2. Effective Semantic Embedding Model (Multilingual-e5-large)
3. DeepSeek LLM's Generative Capabilities – Converting retrieved embeddings into natural, contextual, and user-friendly responses.

By combining computer vision, semantic embeddings, and large language models, this system successfully enables AI-powered memory augmentation, allowing users to effortlessly retrieve past experiences with Based on the input text, here are the relevant keywords for an academic publication:

Discussion

The results demonstrate the potential of this system as a novel memory augmentation tool. The use of vector embeddings and a similarity-based search within Pinecone allows for flexible and context-aware retrieval of information, overcoming limitations of keyword-based search methods. The integration of a large language model (LLM) further enhances the user experience by providing natural language responses, making the recalled information more accessible and understandable.
However, several aspects require further investigation:

- **Scalability:** The performance of the system needs to be evaluated with a larger dataset of images and queries to assess its scalability and efficiency in handling a growing memory bank.
- **Accuracy of Image-to-Text Conversion:** The accuracy of the image-to-text pipeline directly impacts the quality of the stored embeddings and the effectiveness of the retrieval process. Further refinement of this stage is crucial.
- **Parameter Optimization:** The choice of k (number of nearest neighbours) in the retrieval process and other parameters within the LLM need careful optimization to balance recall accuracy and computational cost.

- **Latency and Performance:** Assessing the system's response time under various query loads will help optimize processing speed and improve real-time usability.
- **Security and Privacy:** As the system deals with stored embeddings and sensitive data, robust security measures must be implemented to protect user information and prevent unauthorized access.

5. CONCLUSION

This research explored the application of Retrieval-Augmented Generation (RAG) for knowledge-intensive NLP tasks, focusing on enhancing memory augmentation capabilities through image-based data integration. The findings demonstrate the efficacy of a multi-stage pipeline encompassing image processing, multilingual text embedding (using the multilingual-e5-large model, as detailed in Multilingual E5 Text Embeddings: A Technical Report), Pinecone vector database storage, and LLM-powered response generation (as described in Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks and Evidentiality-guided Generation for Knowledge-Intensive NLP Tasks). The image captioning component (Image Caption Generator) successfully converts visual data into textual representations suitable for embedding and storage.

The integration of an evidentiality predictor within the RAG framework, as highlighted in Evidentiality-guided Generation for Knowledge-Intensive NLP Tasks, significantly improves the accuracy and relevance of generated responses by focusing on evidence-rich passages. This addresses the limitations of standard RAG approaches that may rely on spurious cues or generate hallucinations. The results showcase the potential of this system for creating a robust and contextually aware memory aid, surpassing the performance of baseline models across various knowledge-intensive tasks.

However, future research should focus on enhancing the scalability of the system, optimizing model parameters, and improving the robustness of the image-to-text conversion pipeline. Further investigation into the impact of different embedding models and LLMs is also warranted. Despite these limitations, this research provides a strong foundation for developing advanced memory augmentation systems that leverage the power of RAG and multilingual embeddings to seamlessly integrate visual and textual information. The ability to accurately and naturally recall information from images represents a significant advancement in human-computer interaction and personal memory support.

REFERENCE

[1] Mikel Artetxe and Holger Schwenk. 2019. Mas sively multilingual sentence embeddings for zero shot cross-lingual transfer and beyond. Transactions of the Association for Computational Linguistics, 7:597–610.

[2] Daniel Fernando Campos, Tri Nguyen, Mir Rosenberg, Xia Song, Jianfeng Gao, Saurabh Tiwary, Rangan Majumder, Li Deng, and Bhaskar Mitra. 2016. Ms marco: A human generated machine reading compre hension dataset. ArXiv preprint, abs/1611.09268.

[3] Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettle moyer, and Veselin Stoyanov. 2020. Unsupervised cross-lingual representation learning at scale. In Pro ceedings of the 58th Annual Meeting of the Asso ciation for Computational Linguistics, pages 8440 8451, Online. Association for Computational Lin guistics.

[4] Marta R Costa-jussà, James Cross, Onur Çelebi, Maha Elbayad, Kenneth Heafield, Kevin Heffernan, Elahe Kalbassi, Janice Lam, Daniel Licht, Jean Maillard, et al. 2022. No language left behind: Scaling human-centered machine translation. arXiv preprint arXiv:2207.04672.

[5] Vladimir Karpukhin, Barlas Oguz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih. 2020. Dense passage retrieval for open domain question answering. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 6769–6781, Online. Association for Computational Linguistics.

[6] HaoranWang , Yue Zhang, and Xiaosheng Yu, "An Overview of Image Caption Generation Methods", (CIN-2020)

[7] B.Krishnakumar, K.Kousalya, S.Gokul, R.Karthikeyan, and D.Kaviyarasu, "IMAGE CAPTION GENERATOR USING DEEP LEARNING", (international Journal of Advanced Science and Technology- 2020)

[8] Jonathan Berant, Andrew Chou, Roy Frostig, and Percy Liang. Semantic Parsing on Freebase from Question-Answer Pairs. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pages 1533–1544, Seattle, Washington, USA, October 2013. Association for Computational Linguistics. URL http://www.aclweb.org/anthology/ D13-1160.

[9] Eunsol Choi, Daniel Hewlett, Jakob Uszkoreit, Illia Polosukhin, Alexandre Lacoste, and Jonathan Berant. Coarse-to-fine question answering for long documents. In Proceedings of the 10 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 209–220, Vancouver, Canada, July 2017. Association for Computational Linguistics. doi: 10.18653/v1/P17-1020. URL https://www.aclweb.org/anthology/P17-1020.

[10] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Con ference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics. doi: 10.18653/v1/N19-1423. URLhttps://www.aclweb.org/anthology/N19-1423.

[11] Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. Wiz ard of wikipedia: Knowledge-powered conversational agents. In International Conference on Learning Representations, 2019. URL https://openreview.net/forum?id=r1l73iRqKm.

[12] Matthew Dunn, Levent Sagun, Mike Higgins, V. Ugur Guney, Volkan Cirik, and Kyunghyun Cho. SearchQA: A New Q&A Dataset Augmented with Context from a Search Engine. arXiv:1704.05179 [cs], April 2017. URL http://arxiv.org/abs/1704.05179. arXiv: 1704.05179

[13] J. Redmon, S. Divvala, Girshick and A. Farhadi, "You only look once: Unified real-time object detection", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016

[14] D. Bahdanau, K. Cho, and Y. Bengio. "Neural machine translation by jointly learning to align and translate.arXiv:1409.0473", 2014.

[15] DataCanary, hilfialkaff, Lili Jiang, Meg Risdal, Nikhil Dandekar, and tomtung. 2017. Quora question pairs.

[16] Angela Fan, Yacine Jernite, Ethan Perez, David Grang ier, Jason Weston, and Michael Auli. 2019. ELI5: Long form question answering. In Proceedings of the 57th Annual Meeting of the Association for Com putational Linguistics, pages 3558–3567, Florence, Italy. Association for Computational Linguistics.

[17] Bin Bi, Chenliang Li, Chen Wu, Ming Yan, and Wei Wang. Palm: Pre-training an autoencod ing&autoregressive language model for context-conditioned generation. ArXiv, abs/2004.07159, 2020. URL https://arxiv.org/abs/2004.07159.

[18] Deep Q-Network with Reinforcement Learning for Fault Detection in Cyber-Physical Systems J. Stanly JayaprakashM. Jasmine Pemeena Priyadarsini, B. D. Parameshachari Hamid Reza Karimi, and Sasikumar Gurumoorthy Journal of Circuits, Systems and Computers 2022

[19] Efficient Biometric Security System Using Intra-Class Finger-Knuckle Pose Variation Assessment Mr.J.Stanly Jayaprakash Dr.S.Arumugam , India International Journal of Computer Science & Engineering Technology (IJCSET) 2014

[20] Cloud Data Encryption and Authentication Based on Enhanced Merkle Hash Tree Method J. Stanly Jayaprakash1 , Kishore Balasubramanian2 , Rossilawati Sulaiman3 , Mohammad Kamrul Hasan3, *, B. D. Parameshachari4 and Celestine Iwendi 2021

[21] Multimodal finger biometric score fusion verification using coarse grained distribution function JS Jayaprakash, S Arumugam2015

[22] A novel approach for fingerprint sparse coding analysis using k-svd learning technique S Arthi, J Stanly Jayaprakash 2024

[23] Software Engineering N.Karthiga vani Dr.K.Saravanan Dr.R.Vasanthi,Dr.J.Stanly Jayaprakash,Dr.A.Kanchana 2018

[24] "A Fusion Attention Mechanism with Bi-LSTM-Based Sarcasm Detection for Selecting High-Profit Products". Journal of Circuits, Systems and Computers C. Gayathri., R. Samson Ravindran. (2025). https://doi.org/10.1142/S0218126625501439.

[25] "Energy and Green IT Resource Management Analysis and Formation in Geographically Distributed Environmental Cloud Data Centre" Murugan G, Gayathri.C, Latha.S, Sathiya Kumar C, SudhakarSengan, Priya V(2020), in International Journal of Advanced Science and Technology Vol. 29, No. 6,pp 4144-4155(SCOPUSindexed)

[26] machine learning based internet browsers inmalicious website detection, international journal of innovative research in computer and Communication engineering, sowmiya r, june 2021

[27] "Virtual Human Resource Management withRecruitment in Software Engineering Roles" in International Journal of InnovativeResearch in Computer and Communication Engineering (IJIRCCE), sowmiya r Volume 12, Issue 3, March 2024.

[28] "Advanced Drowsiness Detection system usingOPENCV and KERAS" in International Journal of Innovative Research inComputer and Communication Engineering (IJIRCCE) sowmiya r, Volume 12, Issue 5, May2024.

[29] "Intelligent Phishing Website Detection modelwith Deep Learning- based Innovative Technique" in International Journal ofInnovative Research in Science, Engineering and Technology (IJIRSET) sowmiya r, Volume13, Issue 3, March 2024.

[30] "Predicting Emotions FromText Using Computing Technique" in the "International conference on Integrating Recent Innovations in Science and Technology: Shaping the future (ICIRIST-2024)" sowmiya r,2024.

[31] securing personal information using data mining in Public networks", International Journal ofComputer Science and Engineering, ms. R.sowmiya 2017

[32] Real time cyber physical false data attack Detection in machine learning methods in International Journal of Innovative Research inComputer and Communication Engineering sowmiya r, june 2021

[33] "soil parameter analyzing using curse of dimensionality foraccuracy and prediction , in international journal of innovative research in science, engineering and technology, ms. R.sowmiya feb 2020.

[34] Meiyalakan K. Published following article. Online Multi-Crop Procurement and Loan System. Volume 10, Issue 5, pp: 32-35,

[35] Meiyalakan K. Published following article. Knowledge-Based Approach to Detect Potentially Risky. Websites. Volume 10, Issue 6, pp: 1353-1357

[36] K.MEIYALAKAN ," Estimation of the Available Bandwidth in Inter-Cloud Links for Task Scheduling in Hybrid Clouds", DOI: 10.15680/IJIRCCE.2022.1007086

[37] MEIYALAKAN.K," An Trends of Event Detection by Analyzing Social Media Platforms Data", 10.15680/IJIRSET.2023.1203081,

[38] K.Meiyalakan ,Blockchain-Based Anonymous Authentication of Cloud Data Using Stochastic Diffusion Search Algorithm, https://www.dsengg.ac.in/pdf/cells/ICIRIST-EBook-2024.pdf

[39] K.Meiyalakan," Predicting Emotions from Text using Computing Technique " , https://www.dsengg.ac.in/pdf/cells/ICIRIST-EBook-2024.pdf

[40] K.Meiyalakan, " Cross-Analysis of Network Intrusion Detection Systems Based On Machine Learning" , https://www.dsengg.ac.in/pdf/cells/ICIRIST-EBook-2024.pdf

[41] K.Meiyalakan, " Machine Learning Contributions to the Field of Security and Privacy Using Android", ICSIEM 2024,16 -17, April 2024

[42] Durairam.R.. Machine Learning Approches for Brain Disease Diagnosis. Volume 10, Issue 6, pp: 1092-1097.

[43] R. DURAIRAM.. Deduplication of Storage Drives Using Cloud Computing. Volume 10, Issue 6, pp: 171-172.

[44] C. Anusuya *et al.*, "Credit Card Fraud Detection using Machine Learning-Based Random Forest Algorithm," *Int. J. Sci. Adv. Res. Technol.*, vol. 9, no. 3, Mar. 2023.

[45] C. Anusuya *et al.*, "Alzheimer Disease with Blood Plasma Proteins detected using Convolutional Neural Network (CNN)," *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 11, no. 3, Mar. 2023.

[46] C. Anusuya *et al.*, "Facial Recognition Services for E-Voting System by Using Blockchain Technology," *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. xx, no. xx, pp. xxx-xxx, year.

[47] C. Anusuya *et al.*, "A Comparative Feature Extraction Study Using Textural Features to Extract Vital Information from Lung Images," *Int. J. Adv. Inf. Sci. Technol.*, vol. 12, no. 02,Feb. 2023.

[48] C. Anusuya *et al.*, "Diagnostics Decision Support System for Tuberculosis Using Fuzzy Logic," *IRACST Int. J. Comput. Sci. Inf. Technol. Sec.*, vol. 2, no. 3, Jun. 2012.

[49] C. Anusuya *et al.*, "Classification of Uncertain Data Using Selection Algorithm," *Int. J. Mod. Eng. Res.*, vol. 2, no. 3, pp. 1066-1072, May-Jun. 2012.

[50] Parvathi M "Sensing of Near Duplicates in Large Image Database",Volume 12, Issue 3, March 2023 ,DOI:10.15680/IJIRSET.2023.1203126 .