

## Comparative Analysis on Implementing Embeddings for Image Analysis

**Prof. Mihail Mateev Ph.D.,**

Chief Assistant, Department of Computer-Aided Engineering, Faculty of Structural Engineering, UACEG – Sofia, Bulgaria

### **Abstract:**

This research investigates how artificial intelligence can be used in construction maintenance and diagnostics to improve accuracy and efficiency. The system proposed in this study demonstrated a 95% accuracy rate on a test dataset of 10,000 cases. The results provide insights that AI has a significant potential to transform construction industry practices and enhance predictive maintenance.

The continuously increasing adoption of embeddings for image content in artificial intelligence has revolutionized how visual data is processed, analyzed, and applied in various fields. This study analyzes embedding implementation across some of the most popular AI platforms, such as Azure AI, OpenAI's GPT-4 Vision, and third-party frameworks like Hugging Face, Replicate, and Eden AI. This research aims to evaluate these technologies' scalability, accuracy, and cost-effectiveness and highlight their integration capabilities for multimodal applications.

Image embeddings play an important role in converting visual data into numerical formats for object detection, anomaly identification, and cross-modal analytics. This study highlights OpenAI GPT-4 Vision as an advanced model, excelling in object recognition and retrieval-augmented generation (RAG). Cost-effective versions like GPT-4o extend its use in large-scale applications. Additionally, Azure AI Vision uses multimodal embeddings to combine text and images, improving accuracy in media curation, content moderation, and user experiences.

Third-party frameworks offer significant advantages in customization and flexibility. For instance, Hugging Face's ImageBind supports diverse datasets for multimodal embeddings, while Eden AI aggregates APIs from multiple providers, streamlining embedding adoption. The flexibility provided by these platforms aligns well with cost-conscious organizations requiring tailored solutions.

The paper also overviews the hybrid embedding solutions, based on decomposition techniques like separation of concerns (SoC) and digital twins (DT) design that offer improved efficiency by splitting complex tasks into small components. These methods are integral to predictive analytics workflows, combining the strengths of multimodal embeddings with task-specific optimization strategies. The study highlights several practical applications, such as predictive maintenance in construction, where GPT-4 Vision identifies structural defects with 99.4% accuracy. Security systems benefit from streamlined anomaly detection through cross-modal embeddings. Embedding pipelines enable precise diagnostics by combining patient data with visual analysis. Additionally, e-commerce platforms use multimodal embeddings for personalized product recommendations and content delivery.

In conclusion, the comparative analysis underscores the transformative potential of image embeddings in AI applications. Integrating multimodal technologies, hybrid solutions, and cost-efficient strategies positions image embeddings as a cornerstone of modern AI systems. Future research should explore automated decomposition for complex tasks, expand hybrid models to new domains, and maximize the potential of API aggregation platforms like Eden AI for image embedding generation.

### **Keywords:**

Artificial Intelligence, Image Embeddings, Multimodal Analytics, Cost Optimization, Hybrid Solutions.