IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 10, May 2025
Impact Factor: 7.67

AI-Powered Sensory Augmentation and Visual Data Processing

Prof. S. S. Mane[1], Shubhankar Madhukar Patil[2], Vedant Manohar Patil[3], Sarthiki Hegade[4], Vishal Nityanand Pawar[5]
Professor, Department of Computer Engineering[1]
Students, Department of Computer Engineering[2-5]
NBN Sinhgad Technical Institute Campus, Pune, India

Abstract: *This paper presents a dual-application system designed to enhance accessibility and image retrieval through the use of Vision-Language models. The first component utilizes BLIP (Bootstrapping Language-Image Pretraining) to generate descriptive tags from images stored in a specified directory. These tags are indexed and stored in a JSON database, allowing users to retrieve relevant images by entering textual queries or keywords. The second component is focused on aiding visually impaired individuals by converting real-time camera feed into descriptive text, enabling auditory perception of visual surroundings via text-to-speech. Both applications leverage the power of multi-modal deep learning models to bridge the gap between vision and language. The system prioritizes lightweight design, real-world usability, and modular implementation. While the solutions presented are in their prototype stages, initial evaluations suggest promising utility for assistive technologies and searchable visual databases.*

Keywords: Artificial Intelligence, Computer Vision, Machine Learning, Sensory Augmentation, Image Processing, Accessibility, CLIP, BLIP