Enhancing Cognitive Capabilities in Artificial Intelligence Systems through Multi-Modal Learning

Faith Ugbeshe, Adeniyi Akanni, Olajide Adegunwa

Computer Science Department, Caleb University, Imota, Lagos.

Contact email: ugbeshefaith@gmail.com

ABSTRACT

This paper surveys existing multi-modal learning techniques to enhance the cognitive capabilities of AI systems. Multi-modal learning, which integrates data from multiple sources such as text, images, and audio, mimics human cognitive processes and enhances AI's ability to reason and make decisions across different modalities. By comparing multi-modal models to unimodal counterparts, this paper highlights performance improvements in tasks such as natural language processing, image recognition, and audio processing. Additionally, challenges such as data availability and computational complexity are explored, with suggestions for future research, including lightweight and unsupervised multi-modal learning models.

Keywords: Artificial Intelligence, Multi-Modal Learning, Cognitive Capabilities, Deep Learning, Cross-Modal Reasoning, Image Recognition, Natural Language Understanding, Audio Processing, Data Fusion, Unsupervised Learning.