[conference]IEEEtran [utf8]inputenc graphicx amsmath amsfonts amssymb cite url booktabs multirow array

Multimodal Data Augmentation for Alzheimer's Disease Detection using Generative Models in Latent Space

Hriddhit Datta1, Rahul Das1, Jyothika Seru1, Reddi Pallavi1, Soham Kelaskar1 1Department of Computer Science

document

abstract Early detection of Alzheimer's Disease (AD) is crucial for patient care and treatment planning. This work presents a novel approach for multimodal data augmentation in the latent embedding space to improve AD detection performance. We employ three generative modelsConditional Variational Autoencoders (VAE), Normalizing Flows, and Conditional Generative Adversarial Networks (GAN)to synthesize multimodal embeddings combining audio and text features from the ADReSS-IS2020 dataset. Our pipeline extracts embeddings using state-of-the-art models (Whisper, Wav2Vec2 for audio; ClinicalBERT, BioBERT for text) and fuses them using concatenation and cross-attention mechanisms. Experimental results demonstrate that latent space augmentation can improve classification accuracy by up to X% across different model configurations, with Conditional GANs showing the most promising results for embedding quality and downstream task performance.