

Author Name (Year)	Journal Name	Dataset Used	Image type	Subjects	Data Preprocess	Data Processing	Methodology (in)	Parameters Used	One Big Drawback
Hsu et al. (2021)	IEEE/ACM T. Au	LibriSpeech (960	Speech (Continu	Audio recordings	MFCC extraction	Masked Predictic	Self-supervised l	CNN layers [512	Relies on computationally expensive, two-phase training (offline clustering then BERT training).
Zhang et al. (202	Experimental Ge	Clinical dataset fi	Video, Image, Sp	160 participants	Feature extractio	Interpretable Dec	Zero-shot feature	Tree Depth; Min	Performance constrained by the single-center dataset, limiting generalizability and robust bias/fairness analysis.
Burke et al. (202	Aging, Neuropsy	Pitt Corpus (Dem	Spontaneous Spi	142 participants	Manual transcrip	Random Forest;	Classification of /	Standard Randor	Lack of biomarker confirmation for AD diagnosis, which restricts correlation with underlying biological pathology.
Alkenani et al. (2	Journal of Biome	CTPC (Spoken);	Written and Spok	CTPC: 236 AD, 2	RobustScaler no	Heterogeneous E	Optimized featur	10-fold CV;	Impressive performance is restricted to relatively small, constrained, and English-centric datasets (CTPC, ADBC).
Li et al. (2025)	npj Digital Medici	U.S. Veterans He	Longitudinal Uns	1000 AD cases,	LLaMa 3.1 8B D	Multi-Agent LLM	Ultra-early risk fo	LLaMa 3 70B; LL	Maximum accuracy of 0.53 for 10-year prediction, reflecting the low signal-to-noise ratio in messy, unstructured EHRs.