

Multimodal Multi-Task Audio-Visual Transformer

This groundbreaking proposal heralds the development of an unprecedented AI system geared toward universal audio processing. At its core lies an ambitious initiative—the creation of an "All-In-One audio (AIO) transformer," leveraging cutting-edge modular transformer architecture, notably the Mixture of Experts (MoE) and Self-Supervised Learning (SSL).

Our aim is to unite visionary minds in the pursuit of a transformative AI solution. We seek collaborators passionate about pushing the boundaries of audio processing and AI. Together, we aspire to emulate the intricate complexities of the human auditory system and develop an adaptable, all-encompassing AIO transformer.

By harnessing the power of MoE and SSL, we envisage a platform capable of transcending traditional limitations, revolutionizing how we approach diverse audio tasks. This proposal invites enthusiastic researchers and experts in AI, audio processing, and transformer architectures to join forces, contributing their unique insights and expertise to sculpt an unparalleled innovation in the realm of AI-driven audio processing.

Join us in this pioneering venture to shape the future of AI-powered audio technology. Together, let's create an AIO transformer that not only echoes the prowess of the human auditory system but redefines the possibilities of AI in audio processing.