

Multimodal Emotion Recognition from Audio and Video

Nithyasri S¹, Hemavarthini B², and Dr. Bharathi N. Gopalsamy³

Computer Science and Engineering^{1,2,3}

SRM Institute of Science and Technology, Vadapalani Chennai, India

Abstract: *As humans we want to interact with a machine as we would with a person, in a way that it understands us, advises us, and looks after us with no human supervision. Current systems lack empathy and user understanding in spite of very effective logical reasoning. By predicting the emotions of the users, we are able to identify their needs and cater to them as best as possible. Emotion recognition in video and audio has many potential applications, including conversational agents, recommendation systems as well as systems for smart homes, mental illness care, virtual reality games, remote physical training, education and car-hailing services. The aim of the project is to develop an automatic emotion detection system based on voice and facial expression. We propose a model that highlights contextual, multimodal information for emotion detection and recognition. If systems can understand emotions and respond accordingly to behavioural patterns, we can anticipate artificial agents becoming our cognitive-consulting partners in our daily lives. Additionally, the project can be expanded to make interactions more natural and better suited to handle complex situations.*

Keywords: Multimodal Emotion Recognition, Computer Vision, Deep Learning, Audio and Video, Open CV Librosa