Multimodal Characterization of Emotion within Multimedia Space

Authors : Dayo Samuel Banjo, Connice Trimmingham, Niloofar Yousefi, Nitin Agarwal

Abstract : Technological advancement and its omnipresent connection have pushed humans past the boundaries and limitations of a computer screen, physical state, or geographical location. It has provided a depth of avenues that facilitate human-computer interaction that was once inconceivable such as audio and body language detection. Given the complex modularities of emotions, it becomes vital to study human-computer interaction, as it is the commencement of a thorough understanding of the emotional state of users and, in the context of social networks, the producers of multimodal information. This study first acknowledges the accuracy of classification found within multimodal emotion detection systems compared to unimodal solutions. Second, it explores the characterization of multimedia content produced based on their emotions and the coherence of emotion in different modalities by utilizing deep learning models to classify emotion across different modalities.