

TikTok Video Classification Based on Emotions Using Convolutional Neural Networks

Nagodavithana S.D.^{1*}, Rupasingha R.A.H.M.² and Kumara B.T.G.S.³

¹Department of Computing and Information Systems, Faculty of Computing,
Sabaragamuwa University of Sri Lanka, Sri Lanka

²Department of Economics and Statistics, Faculty of Social Sciences and Languages, Sabaragamuwa
University of Sri Lanka, Sri Lanka

³Department of Data Science, Faculty of Computing,
Sabaragamuwa University of Sri Lanka, Sri Lanka *sajikadilshan2000@gmail.com

Emotion recognition is a product of AI and computer vision, which enables systems to decipher human emotions during a range of uses. As TikTok grows, expressions made by users in the short videos offer valuable information that can be used in real life to study emotions. Detection of emotions in these videos is, however, a challenge because of changes in lighting, angles, and user behavior. Emotion-based TikTok video classification can be essential to enhance content recommendations, interaction, and emotion-driven social media analytics. To overcome this challenge, as an objective of this research, a Convolutional Neural Network (CNN)-based approach is prepared to categorize TikTok videos based on their emotions. This study is novel in its comparative evaluation of CNN-based emotion recognition models on real TikTok videos, an area largely ignored by existing methods. A sample of 4,000 TikTok videos were gathered, containing the main four emotions, namely happiness, sadness, anger, and surprise. Videos were processed during preprocessing, which involved compression, frame conversion with Python and OpenCV, and normalization and data augmentation. The features were extracted, and models were trained and tested using the three CNN architectures, namely, MobileNet, VGGNet, and ResNet. The data was divided using a percentage split, as well as 5-fold and 10-fold cross-validation, with different epochs and batch sizes to achieve the best model performance. Accuracy, precision, recall, F-score, and error rate were used as a measure of model performance. ResNet was the most accurate with 96% accuracy, which is far better than VGGNet (71%) and MobileNet (58%). These results prove the usefulness of CNN architectures to classify emotions in TikTok videos and how they can be applied in social media analytics and affective computing. Future research will focus on multimodal audio-text methods for real-time emotion analysis on social media.

Keywords: *Computer vision, Convolutional neural networks, Emotion recognition, Social media analytics, TikTok*