Highlight detection in live streams using audience reactions with transformer language models

Maximilian Lukas Gutsche

Abstract

Livestreaming of e-sports events has become very popular in recent years, with millions of people watching livestreams of competitions and commenting synchronously in chat rooms. As an effect of the rise of e-sports, there is a demand for match highlight videos. These videos, which consist of the most exciting moments in a match, help followers of the sport to stay up-to-date with or relive past games. Since their manual creation is time intensive, automatic and semi-automatic approaches for highlight detection in live streams have been devised. In this work, we suggest a novel transformer based approach to highlight detection. We employ the audience reactions found in live stream chat in order to find gripping segments of livestreams. To this end, we suggest an approach which combines contextual transformer embeddings with additional temporal features of the chat. We pre-train a language model for the domain of live stream chat in the game League of Legends and employ it on this task. For training this transformer language model, we collect a corpus from a popular livestreaming platform which contains audience reactions to competitive League of Legends matches. With our new model, we achieve an improvement over the state of the art of 0.01 f-score. We provide a new corpus for the domain and make available our pre-trained language model, which we call TwitchLeagueBert.