Automated News Clip Generation via Robust Video Summarization

Adane Tarekegn, Fazle Rabbi, Lubos Steskal, Bjørnar Tessem

Media Futures

Abstract

This work presents a video news clip generation based on a video summarization framework (BGClip) developed in collaboration with TV2 to enhance newsroom production workflows. The framework automatically processes raw news footage to generate short, editorial-ready clips accompanied by an Edit Decision List (EDL). It employs unsupervised learning with multi-head self-attention for frame selection and deep generative models for clip reconstruction. Using the developed tool, editors can specify clip types and durations, enabling the rapid creation of both short-form highlights and longer segments for broadcast or digital distribution. The proposed model is evaluated using newsroom video archives and two benchmarks (TVSum, SumMe), producing clips that show strong alignment with human-edited content.

Editorial Rules

Video summaries must be customizable to user preferences, particularly in video news production, where broadcasts and social media clips must adhere to strict time limits and maintain content relevance. Motivated by these observations, we incorporate user-defined editorial rules into the final stage of the summarization pipeline, allowing editors to refine the output through an interactive user interface. Rules are model agnostic and defined via a YAML-style schema, allowing users to specify constraints, such as clip type, and duration. In addition, each clip's output can be exported as an Edit Decision List (EDL), which is used as a reference for rendering.

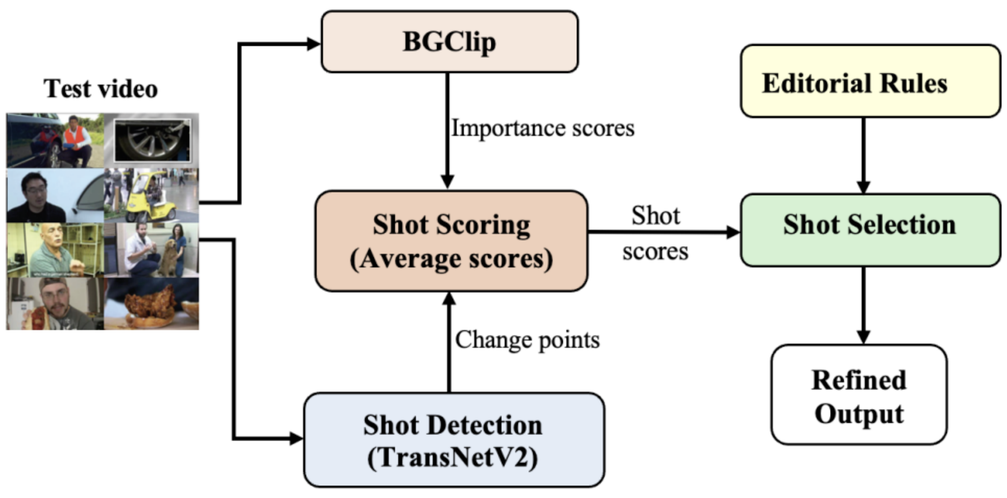

Fig. 1. BGClip model for video clip generation with outputs conditioned on editorial rules. In the clip production process, first BGClip model is applied to the test video to estimate the frame selection probabilities. We use the TransNetV2 model to detect transitions and identify shot boundaries, generating change points. Shot-level scores are then computed by averaging the importance scores of all frames within each shot. After the initial summary is generated, editorial rules are applied as post-processing rules to refine the result.

Methods

Datasets: We use three datasets for training and evaluation: a real-world newsroom dataset to assess editorial compliance, and two benchmarks for comparison with state-of-the-art methods.

Model Components: The framework includes a self-attention for frame selection, Bayesian based VAE for frame reconstruction, GAN, and post-processing steps. It employs multi-head self-attention for video frame importance selection and a Bayesian VAE for frame feature encoding and reconstruction.

Evaluation Approach: F1 score is used as the primary evaluation metric which measures the similarity between model-generated summary and the ground-truth based on their temporal overlap.

Results

Extensive experiments conducted on three datasets demonstrate the effectiveness of our approach for real-world news clip production. On the TVSum dataset, our method achieved an F1 score of 68.43, marking an improvement over the baseline model, and attained a fidelity score of 87–91.5% compared to human-generated clips.

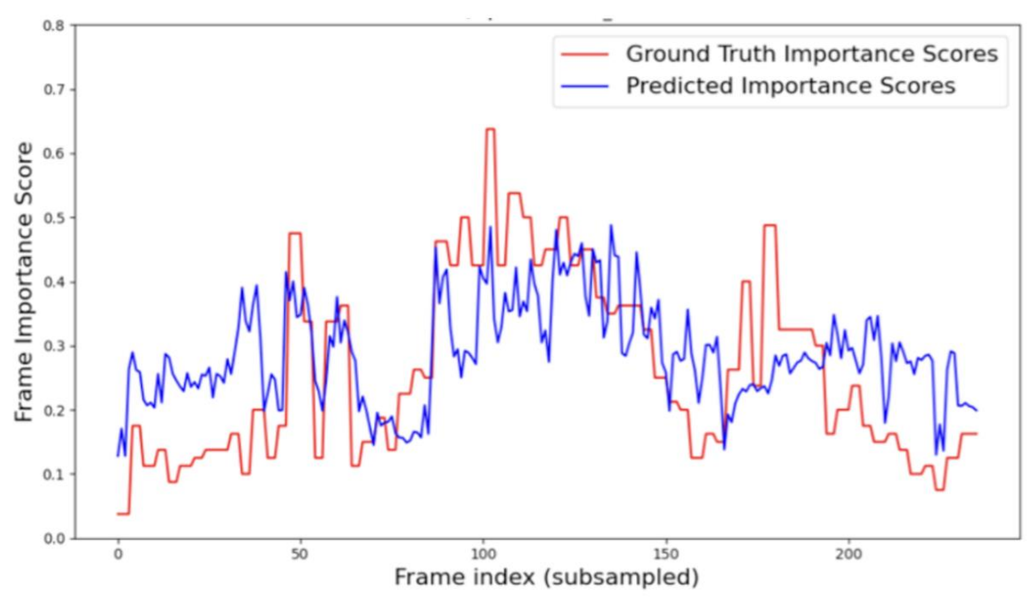

Fig. 2. Correlation between ground truth summaries (red) and model generated summaries (blue) for a video in TVSum.

Conclusion

We propose BGClip, a robust unsupervised video summarization framework for news clip generation, with a post-processing module that adapts summaries to user-defined editorial preferences. We incorporate self-attention mechanisms in the framework for improved feature extraction and frame selection. Unlike existing methods using deterministic models for summary generation, our framework introduces Bayesian inference to better capture data distribution and enhance generalization.

PARTNERS

amedia Schibsted 2 VIZRT NRK Bergens Tidende Faktisk. nb.no Factiverse WOLFTECH NORCE dn media group

HOST

UNIVERSITY OF BERGEN

FUNDED BY

This research is funded by SFI MediaFutures partners and the Research Council of Norway (grant number 309339).

Forskningsrådet

Norwegian Centre for Research-based Innovation