

## Dynamic Video Thumbnail Generation using Deep Learning

Hari K.C.

<sup>1</sup>Tribhuvan University, Institute of Engineering Pashchimanchal Campus, Nepal

Email : harikc@wrc.edu.np

### INTRODUCTION & AIM

Rapid growth of online video content has transformed the way people consume digital media. Video platforms such as YouTube, Netflix, Rumble, and Crackle collectively host billions of videos, catering to diverse audiences and content preferences. One of the primary tools for video content discovery is video thumbnail that serves as a preview image to attract potential viewers. Static thumbnails are replaced with a dynamically generated short clip thumbnail extracted from key moments in a video using BiLSTM and attention techniques. AI driven dynamic previews offer an alternative approach by automatically generating an informative and contextually relevant representation of the video.

The objectives of this study are

- Designing Video summarization framework for appropriate content discovery with dynamic video thumbnails.
- Evaluating user engagement by analyzing the impact of dynamic thumbnails using Click Through Rate(CTR).

### METHOD

The high level architecture for Video thumbnail generation:

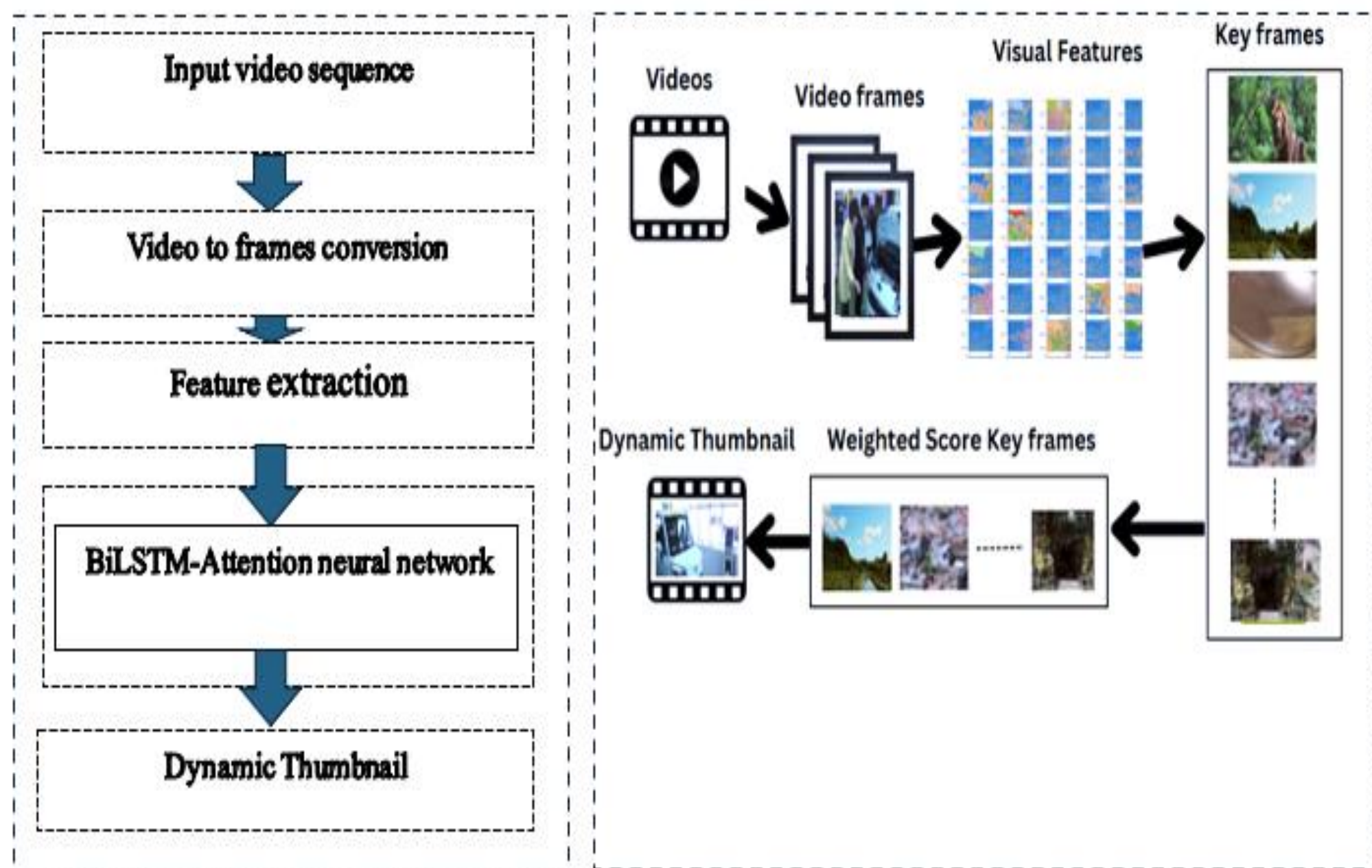


Figure 1: High level architecture

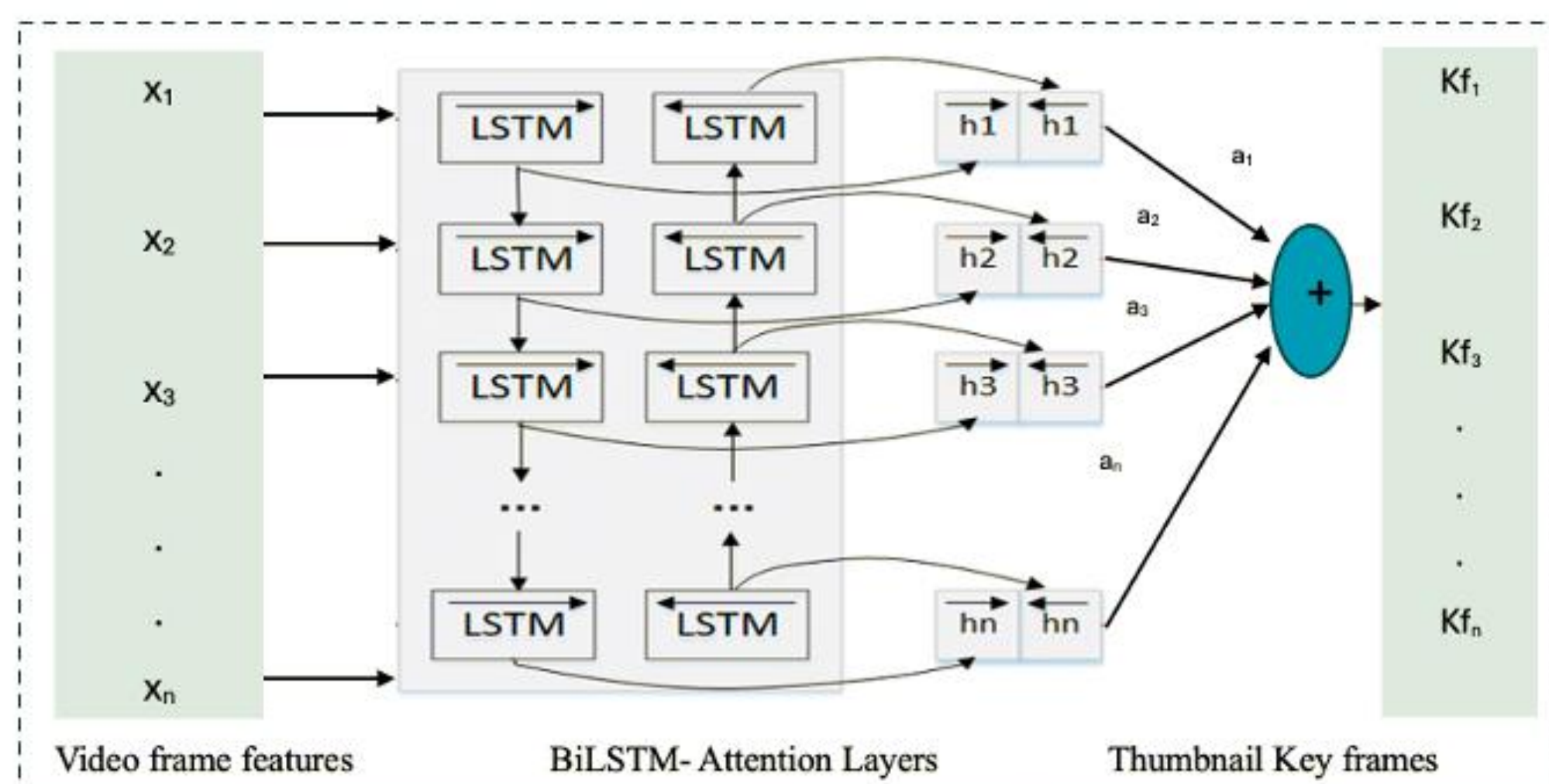


Figure 2: BiLSTM-Attention based framework

### RESULTS & DISCUSSION

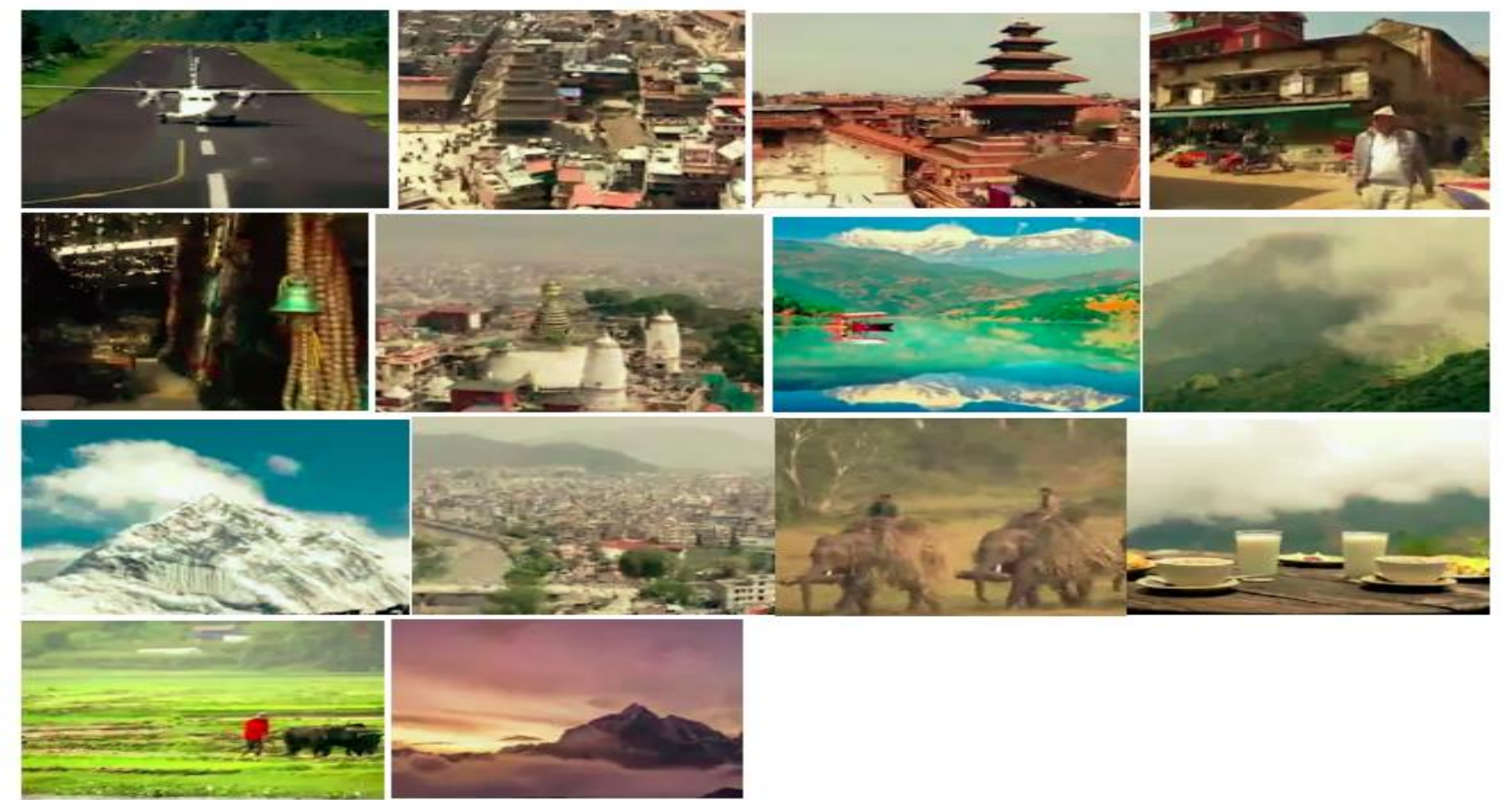


Figure 3: Keyframes for Thumbnail of sample video

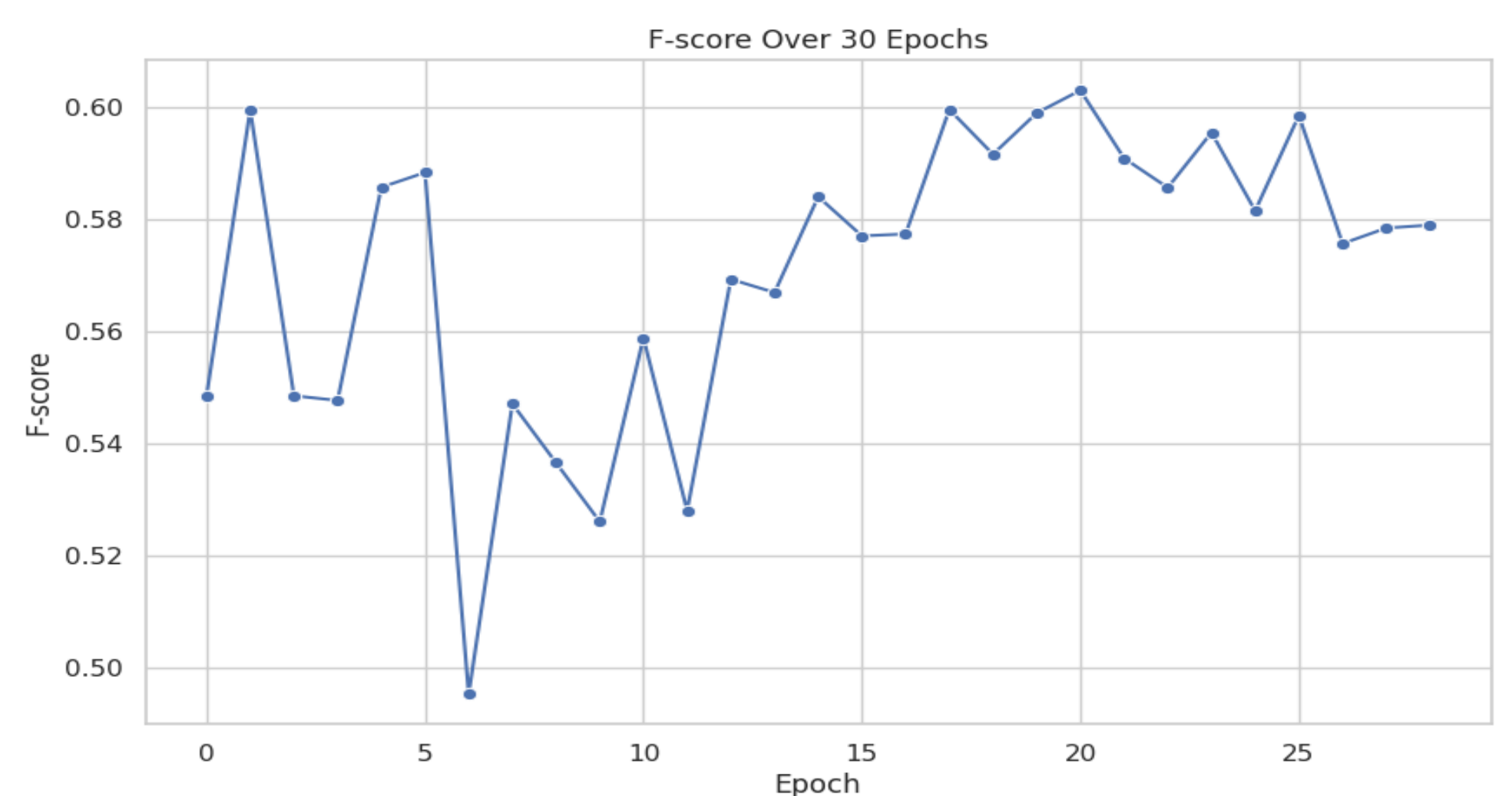


Figure 4: F-Score curve

Table 1: Click Through Rate comparison

Static Thumbnail				Dynamic Thumbnail			
Video	Number of clicks	Number of impressions	CTR(%)	Video	Number of clicks	Number of impressions	CTR(%)
Video 1	20	200	10	Video 1	40	200	20
Video 2	5	200	2.5	Video 2	25	200	12.5
Video 3	16	200	8	Video 3	55	200	27.5
Video 4	8	200	4	Video 4	43	200	21.5
Video 5	12	200	6	Video 5	56	200	28
Video 6	5	200	2.5	Video 6	61	200	30.5
Video 7	15	200	7.5	Video 7	22	200	11
Video 8	20	200	10	Video 8	38	200	19
Video 9	12	200	6	Video 9	34	200	17
Video 10	5	200	2.5	Video 10	56	200	28
Video 11	8	200	4	Video 11	33	200	16.5
Video 12	20	200	10	Video 12	61	200	30.5
Video 13	3	200	1.5	Video 13	23	200	11.5
Video 14	1	200	0.5	Video 14	44	200	22
Video 15	6	200	3	Video 15	32	200	16

This shows the CTR for dynamic thumbnails is higher compared to static thumbnails.

### CONCLUSION

By utilizing convolutional functions for feature extraction and a BiLSTM network with an attention mechanism, our model effectively identifies and extracts key video frames, replacing static thumbnails with informative and engaging dynamic previews. A significant improvement in user engagement can be seen as measured by higher Click Through Rates (CTR) compared to traditional thumbnail based previews.