AI-Based Analysis of Traffic Accident Footage: Predicting Accident Types and Liability Ratios Using Deep Learning

Harang Kim and Hyun Min Song*

Dankook University, Yongin, Gyeonggi, South Korea
{harangkim, hyunminsong}@dankook.ac.kr

Abstract

Traffic accidents necessitate prompt, accurate assessments, often impeded by conventional methods. This study introduces an efficient AI framework for analyzing accident videos, employing SegFormer and the Video Swin Transformer. Our approach streamlines the evaluation process, extracting critical incident details from optimized video data, enhancing the accuracy of real-world traffic accident analysis, and significantly curtailing assessment duration.

Keywords: AI, post-accident analysis, CNFs,

1 Objective

Our research introduces an AI-based method to analyze traffic accident videos, merging image segmentation with video classification for accurate evaluations. This technique offers a faster and more efficient solution for insurance assessments, addressing the need for prompt and precise post-accident analysis.

2 Data Collection and Preprocessing

Figure 1: Accident Video Samples

In this research, we used real-world traffic accident videos from South Korea, provided by AI Hub [1]. This dataset includes various video types such as vehicle black box recordings, road CCTV footage, and third-party black box videos. Each video shows 10 seconds around the accident at 15 fps and 1920 × 1080 resolution. We preprocessed the data for analysis efficiency. The videos were resized to 1024 × 512 pixels for uniformity and reduced computational load. From the original 150 frames per video, we extracted a key subset of 32 frames to focus on the most vital moments of each accident, balancing comprehensive coverage with data processing efficiency.

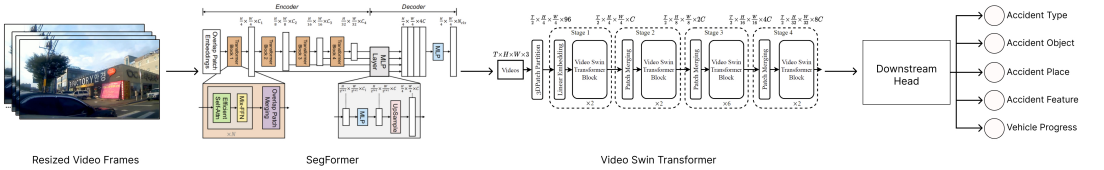

Figure 2: AI Model Structure

3 Traffic Accident Video Analysis Model

Our AI framework uses a two-tiered strategy for traffic accident video analysis, focusing on both detailed and temporal insights. The first step involves NVIDIA's SegFormer model [2], pre-trained on the cityscapes dataset, to extract detailed information from individual video frames. This process forms the basis for in-depth data to be used in subsequent analyses. The next phase employs the Video Swin Transformer (VST) [3] for temporal analysis, focusing on inter-frame dynamics. The VST extracts essential temporal details, critical for decoding the events' sequence and interactions. It integrates the data from individual frames into a unified video representation, capturing the incidents' overall context and nuances. The system then applies a tailored classification head to this rich video representation. This stage predicts key accident characteristics, including the accident place, entities involved, progresses of entities, and liability ratios. These predictions are vital for follow-up procedures and legal considerations. Overall, our AI model provides a streamlined approach to traffic accident analysis. It combines granular data from individual frames with a broader video perspective, offering predictions that align with the complexities of real-world scenarios.

4 Conclusion

Our research presents a pioneering AI approach to traffic accident scrutiny, offering detailed, accelerated analyses by intelligently processing incident videos. The framework's efficacy in pinpointing accident specifics not only promises swifter, more precise assessments but also indicates its potential for broader application. Continued refinement and expanded scenario training will further its readiness for widespread adoption in traffic-related adjudications.

Acknowledgement:
This work was supported by the Technology Innovation Program (P0023522, HRD Program for Industrial Innovation) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea)

References

[1] AI hub. Traffic accident video data. online, 2021. https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=597.

[2] Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M Alvarez, and Ping Luo. Segformer: Simple and efficient design for semantic segmentation with transformers. *Advances in Neural Information Processing Systems*, 34:12077–12090, 2021.

The 7th International Conference on Mobile Internet Security (MobiSec'23), Dec. 19-21, 2023, Okinawa, Japan, Article No.P-21
*Corresponding author: Dankook University, Yongin, 16890, Republic of Korea, Email: hyunminsong@dankook.ac.kr

[3] Ze Liu, Jia Ning, Yue Cao, Yixuan Wei, Zheng Zhang, Stephen Lin, and Han Hu. Video swin transformer. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3202–3211, 2022.