Research on Deepfake Video Detection Using Deep Learning

Dương Thế Tài
Thai Nguyen University of Information and Communication Technology
Thai Nguyen University, Thai Nguyen, Vietnam
dtc2054802010446@ictu.edu.vn

***Abstract*—This thesis investigates the development of a deep learning-based solution for detecting deepfake videos. The author presents a system combining ResNeXt and LSTM networks to effectively identify deepfake videos by leveraging spatial and temporal features. The model demonstrates high accuracy in detecting inconsistencies such as lip movements and facial expressions, and maintains stability on pre-processed DFDC datasets. However, the model faces challenges such as long training time, high GPU resource requirements, and reduced performance with heavily compressed or high-quality deepfake videos.**

***Index Terms*—Deepfake, Deep Learning, ResNeXt, LSTM, Video Detection, Deepfake Detection**

I. Introduction

Deepfake videos can be difficult to identify from a single frame because manipulation may also involve changes over time. The thesis investigates a model that combines spatial image features with temporal relationships across video frames. It uses the Deepfake Detection Challenge data described in the report and presents a pipeline for preprocessing, training, and evaluation. The aim is to test whether a ResNeXt and LSTM combination can recognize manipulated video material within that experimental setting [1].

II. Methods

The report describes collecting and preprocessing video data, extracting frames, and passing frame images through ResNeXt to obtain spatial features. An LSTM then models relationships between consecutive frames before the final classification decision. The implementation is trained and evaluated on preprocessed DFDC material. The summary follows the source's workflow and does not infer a live monitoring service or independent testing on other datasets. Model resource demands are considered in the source discussion [1].

III. Results

The author reports that the combined model identified manipulated videos in the preprocessed DFDC experiment and remained stable on that material. The discussion attributes this to the combination of spatial information from ResNeXt and temporal information from LSTM. The manuscript describes anomalies such as unsynchronized lip movement and unnatural facial expression. It also notes long training time, substantial GPU requirements, and possible weaker performance for heavily compressed or especially convincing fakes. The summary avoids an unverified numerical accuracy claim [1].

IV. Discussion

The source describes an experimental detector, with its claims limited by the data and preprocessing used in the thesis. Video compression and high-quality manipulation may reduce effectiveness, and training requires considerable computational resources. These constraints matter before applying the method to real-time content verification. The author proposes Transformers, faster inference, larger training data, and later deployment. Those are research directions, not functions demonstrated by the current report [1].

V. Conclusion

The thesis presents a ResNeXt and LSTM approach to detecting deepfake videos in a preprocessed DFDC experiment. Its reported result supports the use of spatial and temporal video features together, while the manuscript identifies training cost and difficult video conditions as limitations. The source does not establish a measured production performance level or a deployed real-time service. Broader data, faster inference, and additional model architectures remain proposed work. The contribution is the documented experimental pipeline and its source-reported observations [1].

Supervisor: Nguyễn Văn Việt, Thai Nguyen University of Information and Communication Technology (nvviet@ictu.edu.vn).

References

[1] Dương Thế Tài, "Research on Deepfake Video Detection Using Deep Learning," internal undergraduate thesis (original in Vietnamese), Thai Nguyen University of Information and Communication Technology, 2026.
[2] Amazon Web Services. "Artificial intelligence (ai) là gì?"
[3] Machine Learning cơ bản. "Giới thiệu về machine learning."
[4] NTTuan8. "Deep learning cơ bản."
[5] Deepfake detection challenge dataset (dfdc).
[6] Pytorch documentation.
[7] "Resnext: Aggregated residual transformations for deep neural networks,"
[8] "Long short-term memory (lstm) networks,"
[9] "Faceforensics++: Learning to detect manipulated facial images,"
[10] "Celeb-df: A large-scale challenging dataset for deepfake forensics,"
[11] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.