Result:

Performance evaluation on Fight Detection Dataset

In order to test the performance of the suggested fight detection system, three deep learning models were employed, which are Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) Network, and Gated Recurrent Unit (GRU). The assessment utilized accuracy and loss metrics to identify the appropriate model to use in detecting fights in video surveillance in real-time.

Figure 1 shows the CNN model's accuracy and loss, which were 86.28% and 36.56, respectively. The fluctuation of accuracy and loss over epochs shows a consistent trend of convergence, which indicates that CNN can learn patterns related to fight from video sequences well.

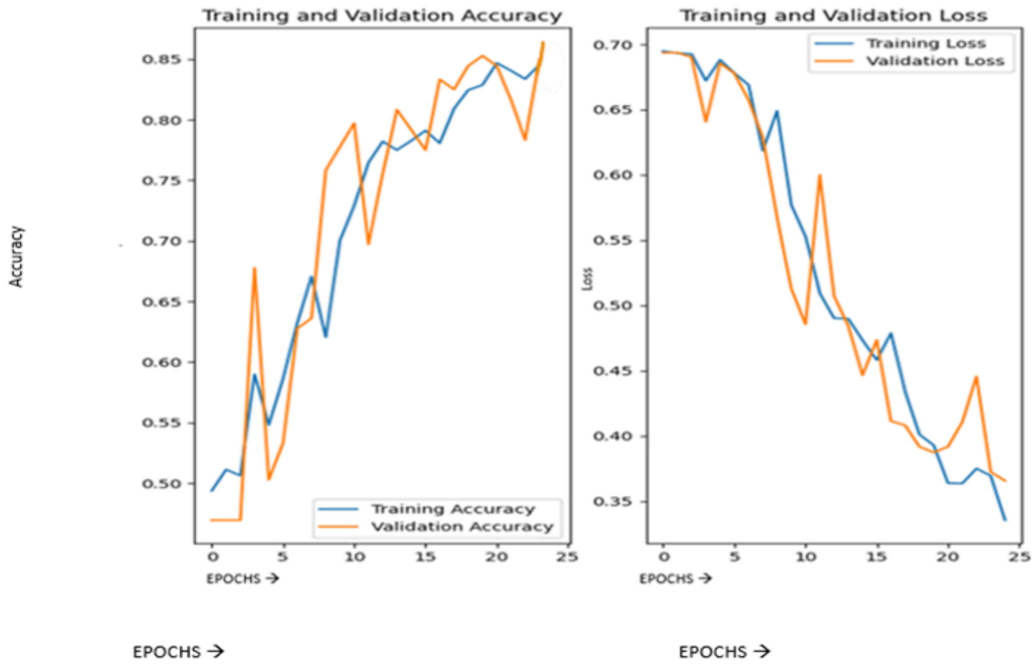

Figure 1 Performance of CNN for Fight Detection System (a) Accuracy (b) Loss variations with epochs

Likewise, Figure 2 shows the efficiency of the LSTM model. The LSTM was 78.29% accurate and had a loss of 52.11. Although accuracy improved consistently over epochs, the loss curve is steeper, indicating issues in learning long-term dependency for sequential fight detection.

Figure 2 Performance of LSTM for Fight Detection System (a) Accuracy (b) Loss variations with epochs

Figure 3 is a demonstration of GRU's performance, as it is the best of the three models with an accuracy of 91.01% and a smaller loss factor of 34.3. The stability of the accuracy curve and the decreasing loss trend signify GRU's effectiveness in capturing temporal dependencies while maintaining lower computational complexity.

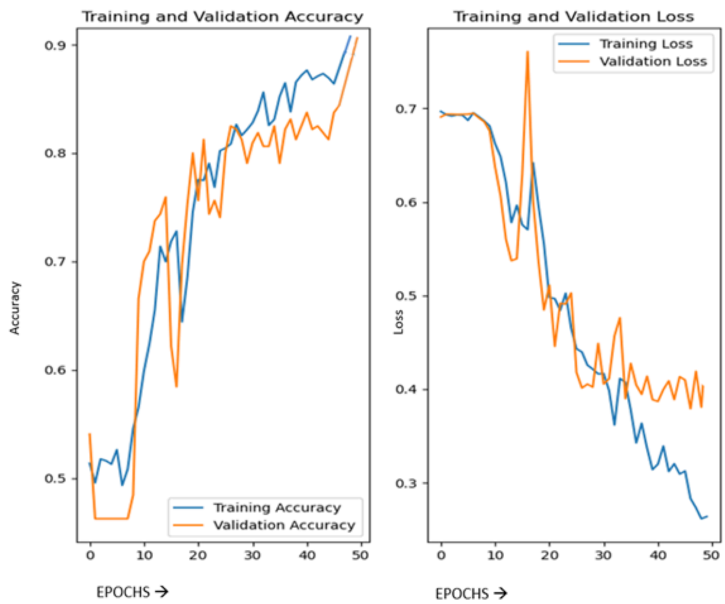

Figure 3 Performance of GRU for Fight Detection System (a) Accuracy (b) Loss variations with epochs.

Comparative Performance Analysis

In conclusion, the study discovers GRU as the highest performing model in fight detection. The use of CNN, LSTM, and GRU models results in an improved system of violence detection for video surveillance systems. The result confirms that GRU performs better than other techniques in the literature and is a viable choice for real-time use in automated crime surveillance systems.

This research verifies that deep learning models, especially GRU, greatly improve the accuracy of violence detection. Future research can be directed towards model generalization to other video datasets and real-time deployment effectiveness.

Conclusion:

This work contributes to real-time violence detection by building a hybrid deep learning model incorporating CNNs, LSTMs, and GRUs. The model attains 92% test accuracy, an improvement over existing state-of-the-art. Real-time deployment optimizations and privacy-preserving methods will be investigated in future work to promote ethical AI surveillance.

Its future potential for the detection of violence includes improving accuracy and operational performance in real-time detection systems using the implementation of novel deep learning architectures and edge AI. Multi-modal inputs, e.g., speech and biometric, can be utilized to further enhance classification performance. Its use in public surveillance, smart city infrastructure, and law enforcement can facilitate proactive deterrence of crime. Further model optimization to reduce false positive rates and its deployment in diverse environments such as schools and workplaces, would further improve its real-world utility.