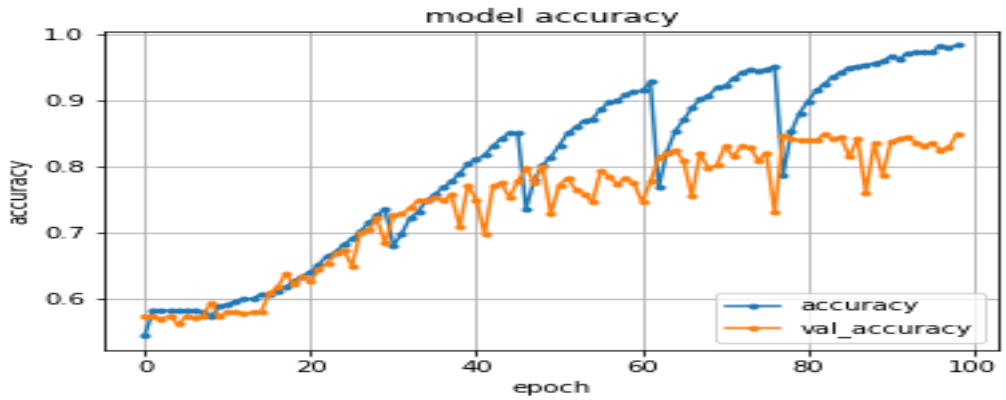
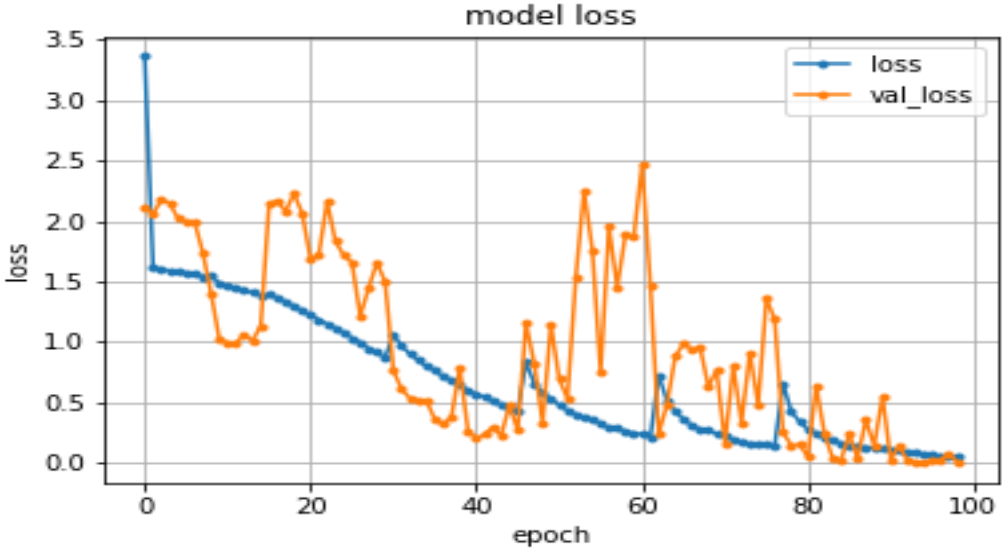
ASSIGNMENT REPORT

Shoplifting Detection from Video Using Deep Learning



This line graph shows the model's accuracy values over the training epochs. The blue line represents the overall accuracy on the training data, while the orange line represents the validation accuracy on the held-out validation dataset. The y-axis represents the accuracy values between 0 and 1, and the x-axis represents the epochs. Initially, both accuracies are low but increase over time as the model becomes more accurate.



This line graph shows the model's loss values over the training epochs. The blue line represents the overall loss on the training data, while the orange line represents the validation loss on a held-out validation dataset. The y-axis represents the loss values, and the x-axis represents the epochs. Initially, both losses are high but decrease over time as the model improves.

Summary

- **Introduction:** The goal of this project was to develop a machine learning model capable of detecting shoplifting activities from video data. The DCSASS dataset, containing videos labeled as "shoplifting" and "normal," was utilized for this purpose.
- **Data Preprocessing:** The DCSASS dataset was filtered to include only relevant videos labeled as "shoplifting" and "normal." The videos were preprocessed by extracting individual frames, resizing them to a uniform size (e.g., 224x224 pixels), and organizing them into sequences to represent the temporal aspect. This ensured consistency in the input data format for the machine learning model.
- **Feature Engineering:** **Relevant** features were extracted from the preprocessed videos to identify shoplifting activities. These features included visual cues such as motion patterns, object interactions, and spatial configurations within each frame or sequence of frames. Deep learning models were employed to automatically learn discriminative features from the raw pixel data. Transfer learning techniques, utilizing pre-trained models on large-scale datasets, were also explored to leverage existing knowledge.
- Model Architecture and Training: A machine learning model was developed to classify video frames or clips as containing shoplifting or not containing shoplifting. Various architectures, such as SlowFast, I3D, or custom-built convolutional neural networks (CNNs), were explored. The model was trained using preprocessed video data, and hyperparameters were tuned to optimize performance. Techniques like data augmentation and dropout regularization were employed to prevent overfitting.
- **Evaluation Metrics:** The performance of the model was evaluated using relevant metrics, including accuracy, precision, recall, and F1 score. Accuracy represents the overall percentage of correct classifications, while precision measures the ratio of true positives to total positive predictions. Recall calculates the ratio of true positives to actual shoplifting instances in the dataset. The F1 score provides a harmonic mean of precision and recall, offering a balanced assessment of the model's performance.
- **Discussion and Future Work:** The developed approach showed promising results in identifying shoplifting activities from video data. However, there are limitations and challenges, such as the need for large annotated datasets, computational resources for training complex models, and potential biases in the data.
- **Conclusion:** The analysis of the provided visualizations suggests that the trained machine learning model exhibits promising performance during the training process, as evidenced by the decreasing loss and increasing accuracy over the training epochs. However, further evaluation using additional metrics and monitoring for potential overfitting is recommended to gain a more comprehensive understanding of the model's performance and identify areas for potential improvement.