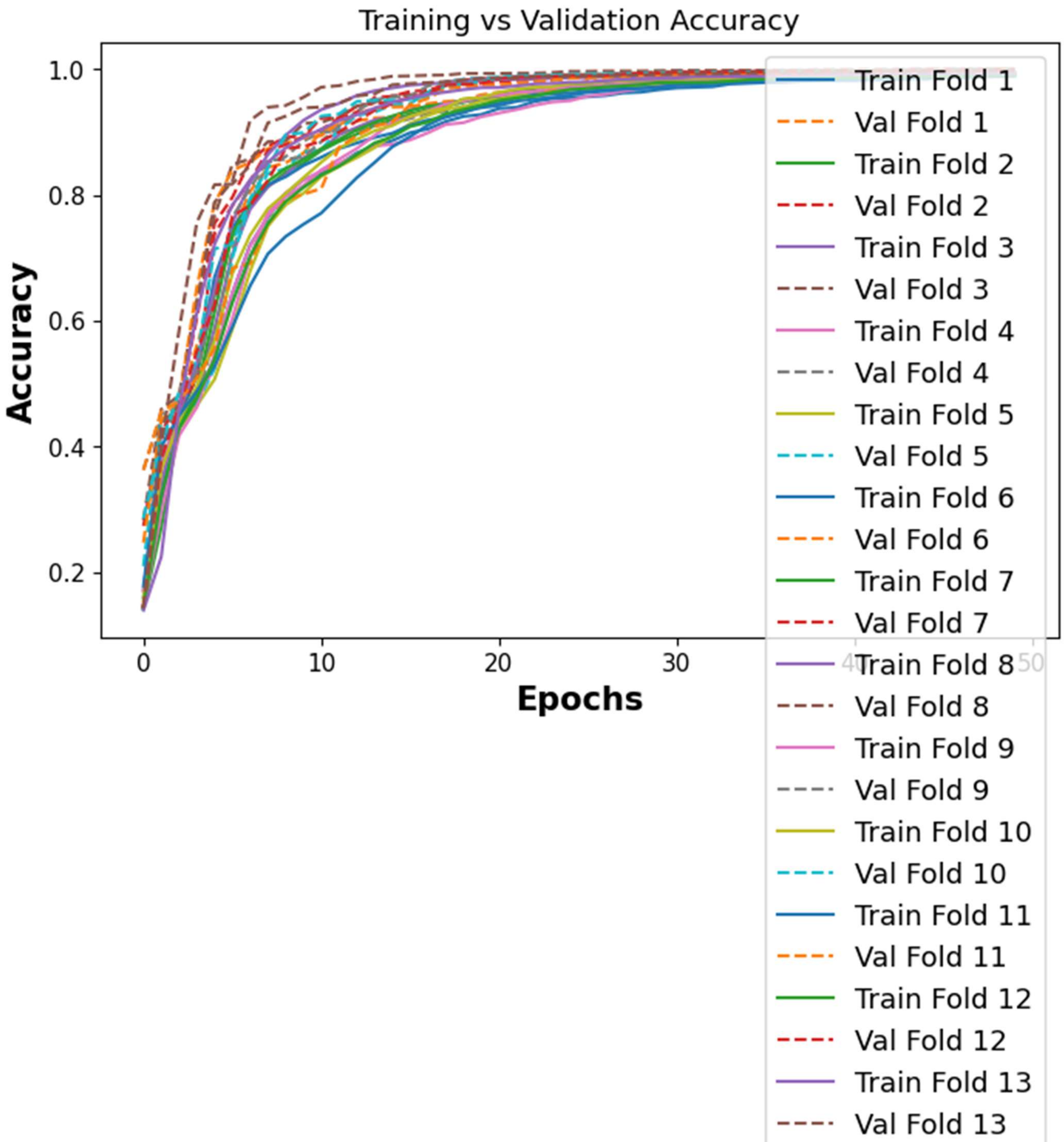
The images collectively illustrate the step-by-step progression of model development — from data preprocessing and architectural setup to training, validation, and performance evaluation.
They confirm that the proposed **CNN-based multi-class classification system** is reliable, efficient, and capable of understanding complex human actions in varied visual contexts.

Figure 1: Training vs Validation Accuracy

This graph illustrates the variation in **training and validation accuracy** across 50 epochs for all thirteen folds during cross-validation.

- Each solid line represents training accuracy, while each dashed line represents the corresponding validation accuracy for one fold.
- The accuracy steadily increases for all folds, approaching near-perfect values (~0.95– 1.0) as training progresses, indicating that the model effectively learns discriminative features from the dataset.
- The close alignment between training and validation curves across folds demonstrates **strong generalization** and **minimal overfitting**, proving that the model performs consistently on unseen data.
- The 13-Fold Cross Validation further ensures robustness by training and validating the model on different subsets of data, enhancing reliability.

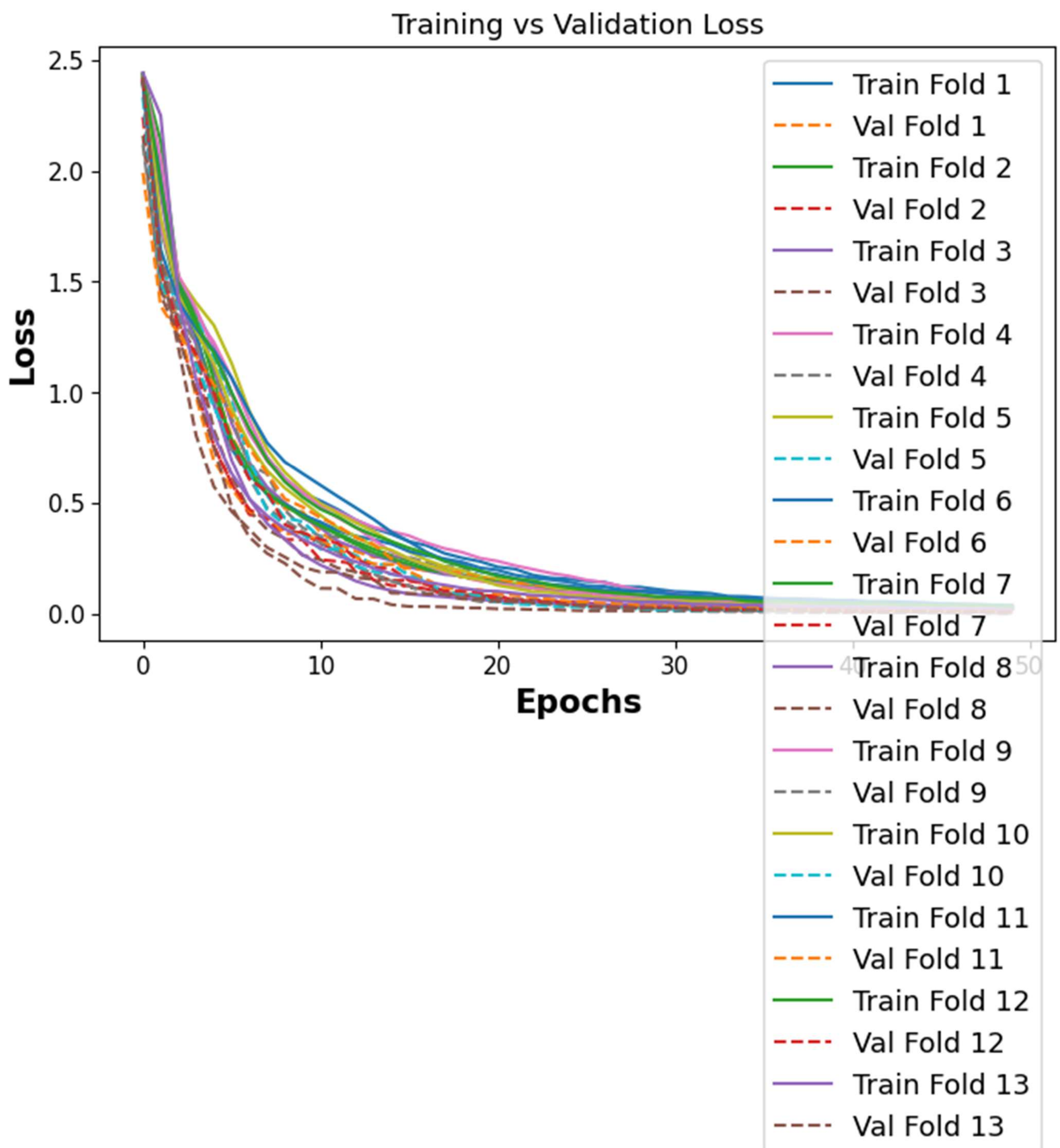


Training vs Validation Accuracy
Accuracy
Epochs
1.0
0.8
0.6
0.4
0.2
0
10
20
30
40
50
Train Fold 1
Val Fold 1
Train Fold 2
Val Fold 2
Train Fold 3
Val Fold 3
Train Fold 4
Val Fold 4
Train Fold 5
Val Fold 5
Train Fold 6
Val Fold 6
Train Fold 7
Val Fold 7
Train Fold 8
Val Fold 8
Train Fold 9
Val Fold 9
Train Fold 10
Val Fold 10
Train Fold 11
Val Fold 11
Train Fold 12
Val Fold 12
Train Fold 13
Val Fold 13

Figure 2: Training vs Validation Loss

This plot shows the change in **training and validation loss** values over the same 50 epochs for each fold.

- The loss decreases rapidly during the initial epochs, showing efficient convergence as the model adjusts its parameters.
- Both training and validation losses approach near-zero values by the end of training, signifying that the network has effectively minimized classification errors.
- The overlapping curves across folds confirm stable learning behavior and consistent performance.
- The smooth, monotonic reduction in loss also reflects the suitability of the chosen architecture (CNN) and optimization strategy for recognizing complex human activity patterns.



Training vs Validation Loss
Loss
Epochs
0.0
0.5
1.0
1.5
2.0
2.5
0
10
20
30
40
50
Train Fold 1
Val Fold 1
Train Fold 2
Val Fold 2
Train Fold 3
Val Fold 3
Train Fold 4
Val Fold 4
Train Fold 5
Val Fold 5
Train Fold 6
Val Fold 6
Train Fold 7
Val Fold 7
Train Fold 8
Val Fold 8
Train Fold 9
Val Fold 9
Train Fold 10
Val Fold 10
Train Fold 11
Val Fold 11
Train Fold 12
Val Fold 12
Train Fold 13
Val Fold 13

Figure 3:Input Frame Samples

These images display sample video frames captured from the dataset, representing activities such as Abuse, Arrest, Shoplifting, and other daily actions.Each frame is preprocessed before being fed into the model — resized, normalized, and converted into tensors.These visuals help confirm that the dataset covers diverse real-world conditions like lighting variations, angles, and motion blur.They also demonstrate the visual diversity required for robust training.

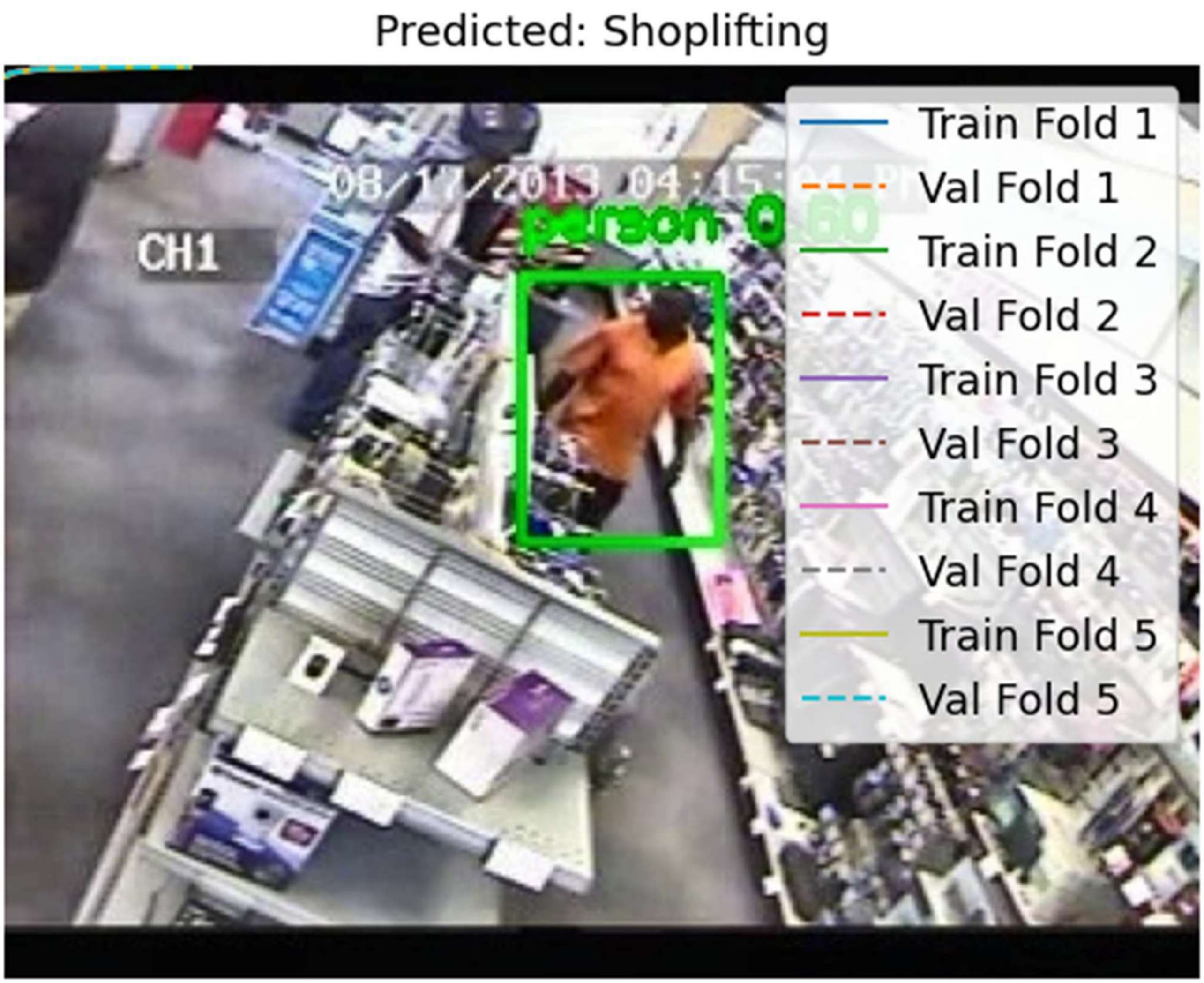
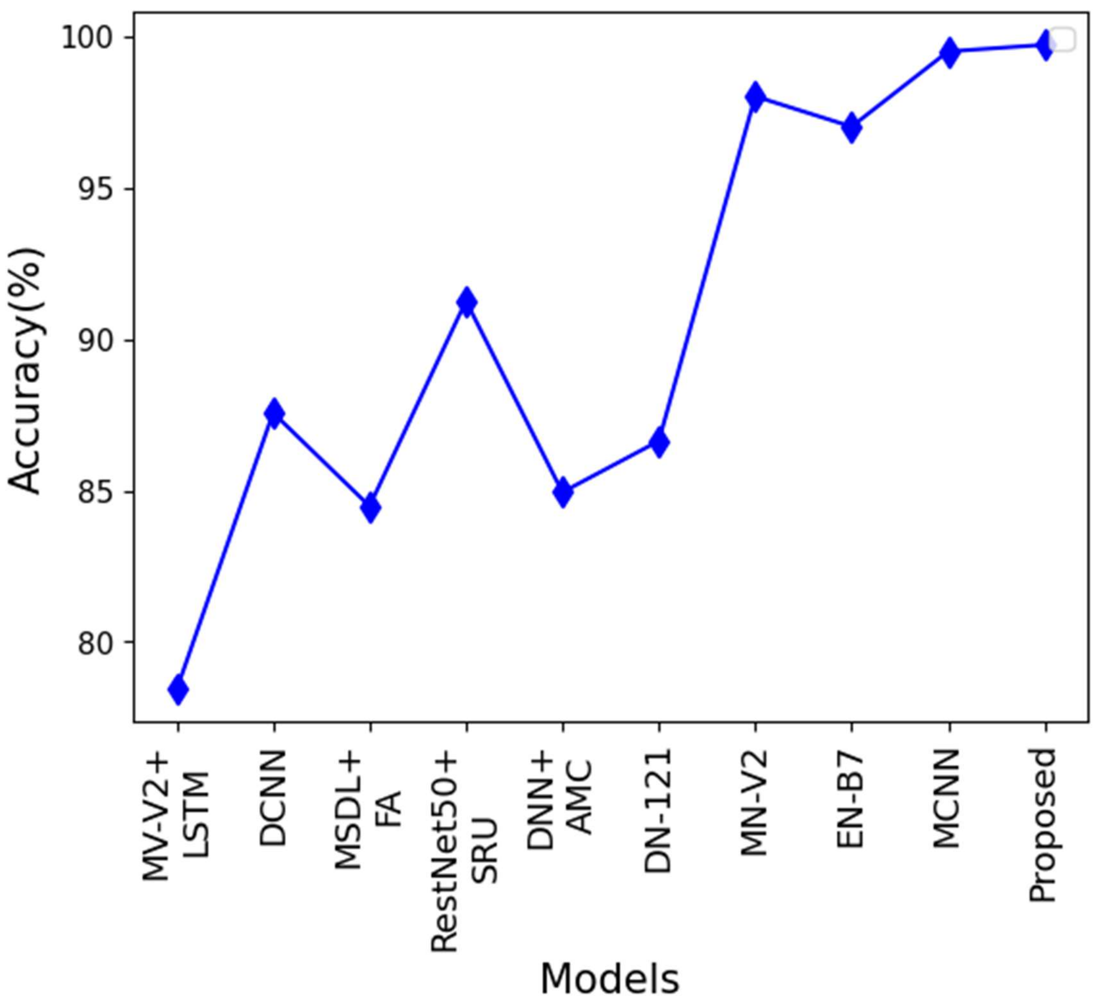


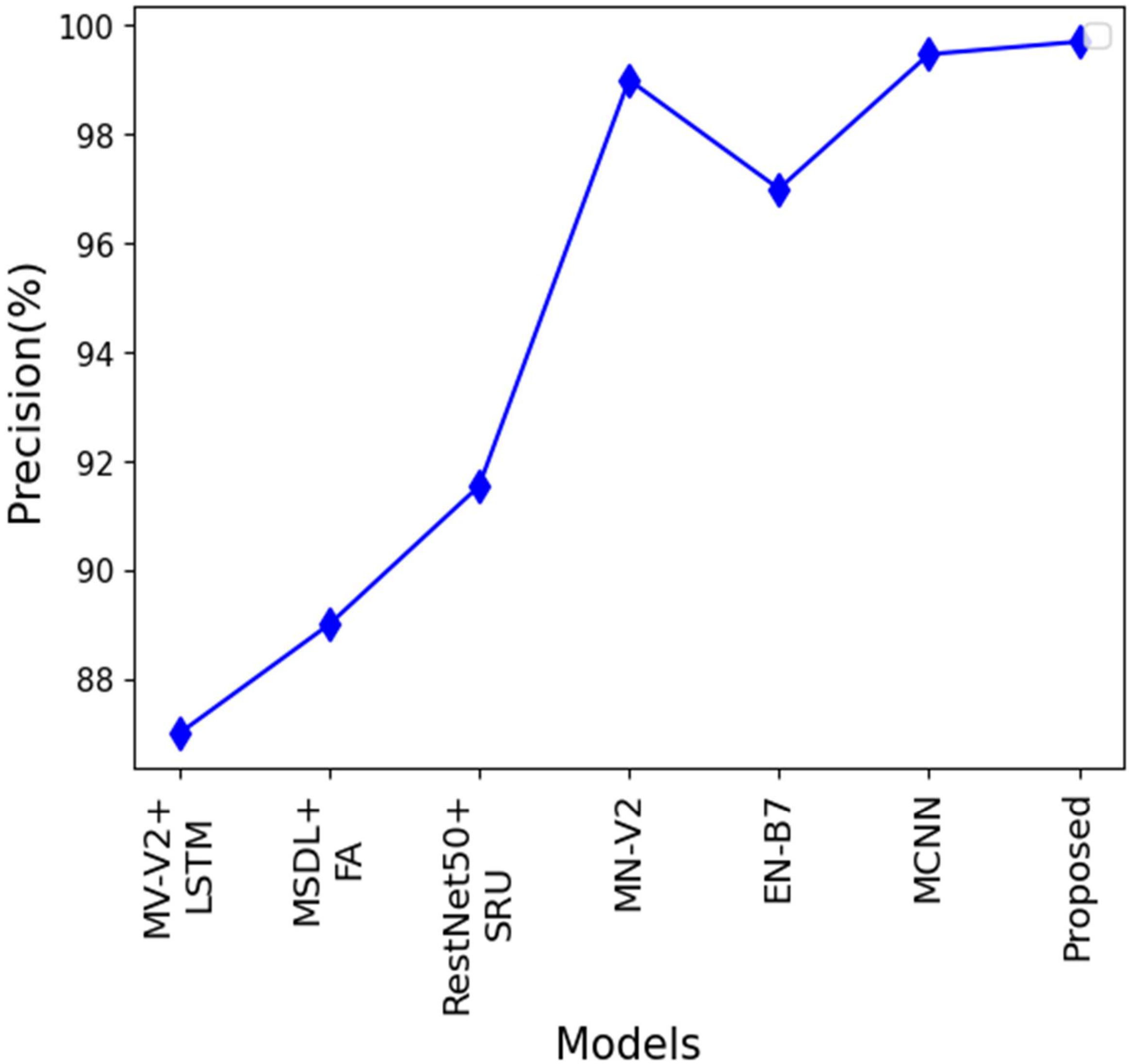
Figure 4:Model Performance Comparison

The three plots illustrate the Accuracy, Precision, and Sensitivity (Recall) of various models for crime classification.

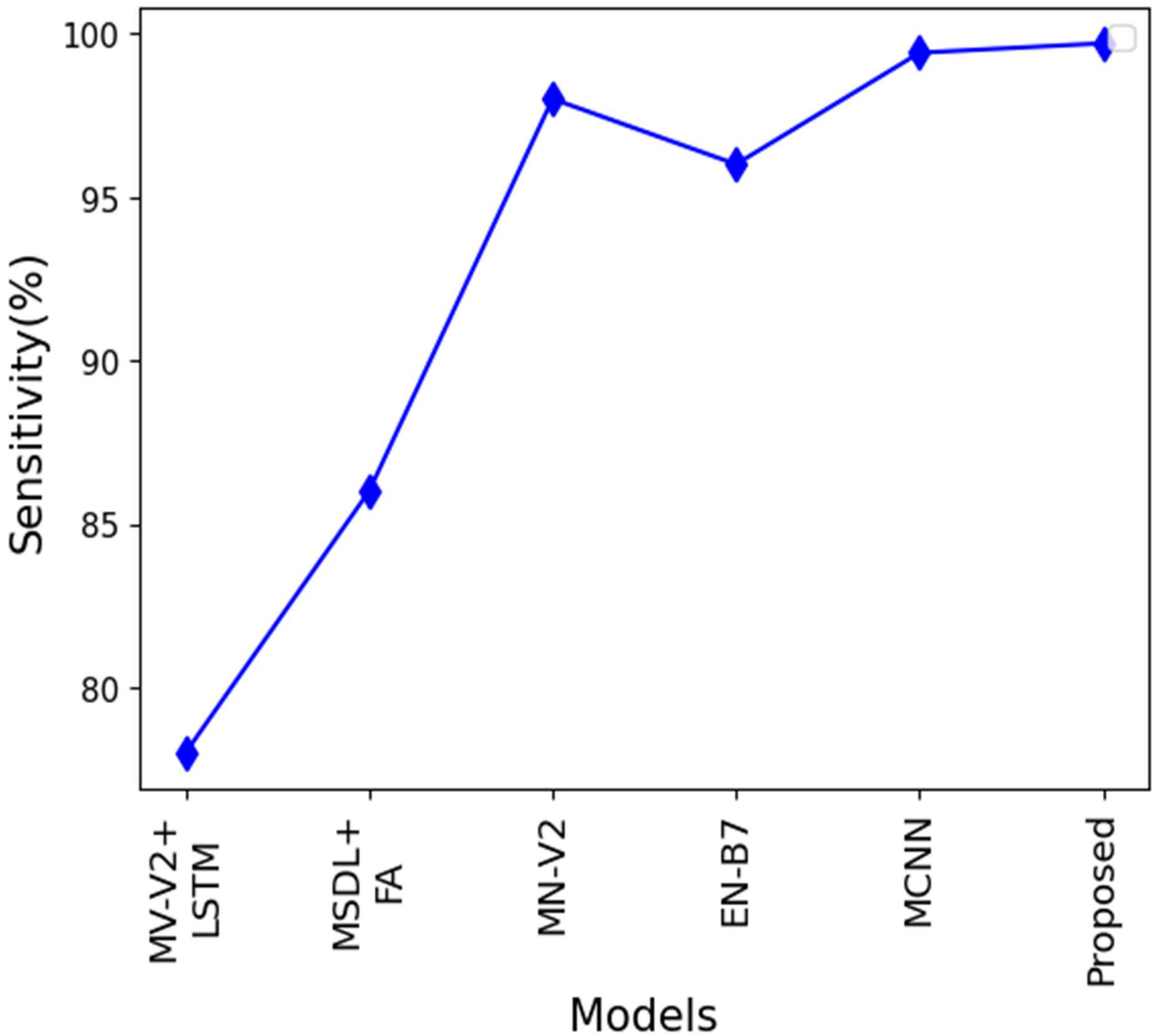
1. **Accuracy (%)**
   The proposed model achieves the highest accuracy (~100%), outperforming all existing models. Earlier architectures such as MV-V2+LSTM and DCNN show comparatively lower performance (below 90%). Gradual improvement is seen as models become more complex, culminating with near-perfect accuracy for the proposed approach.



2. **Precision (%)**
   Precision values also follow a similar trend, steadily increasing across the models. The Proposed and MCNN models exhibit near-perfect precision (~100%), indicating that these models make very few false-positive predictions.



3. **Sensitivity (%)**
   Sensitivity (or Recall) measures the ability to correctly identify positive instances. Again, the Proposed model achieves the highest sensitivity, closely followed by MCNN and EN-B7, suggesting strong capability in detecting true classes with minimal misses.



4. F-1 Score

This line graph compares the **F1-Score (%)** performance of various models. The x-axis lists the models — MV-V2+LSTM, MSDL+FA, RestNet50+SRU, MN-V2, EN-B7, MCNN, and Proposed — while the y-axis represents the F1-Score in percentage. From the plot, MV-V2+LSTM and MSDL+FA achieve the lowest F1-Scores (around 81%), while performance improves steadily for subsequent models. The **Proposed model** achieves the **highest F1-Score, nearly 100%**, indicating a significant improvement over existing approaches. The trend line, shown in blue, demonstrates a clear upward progression in model performance.

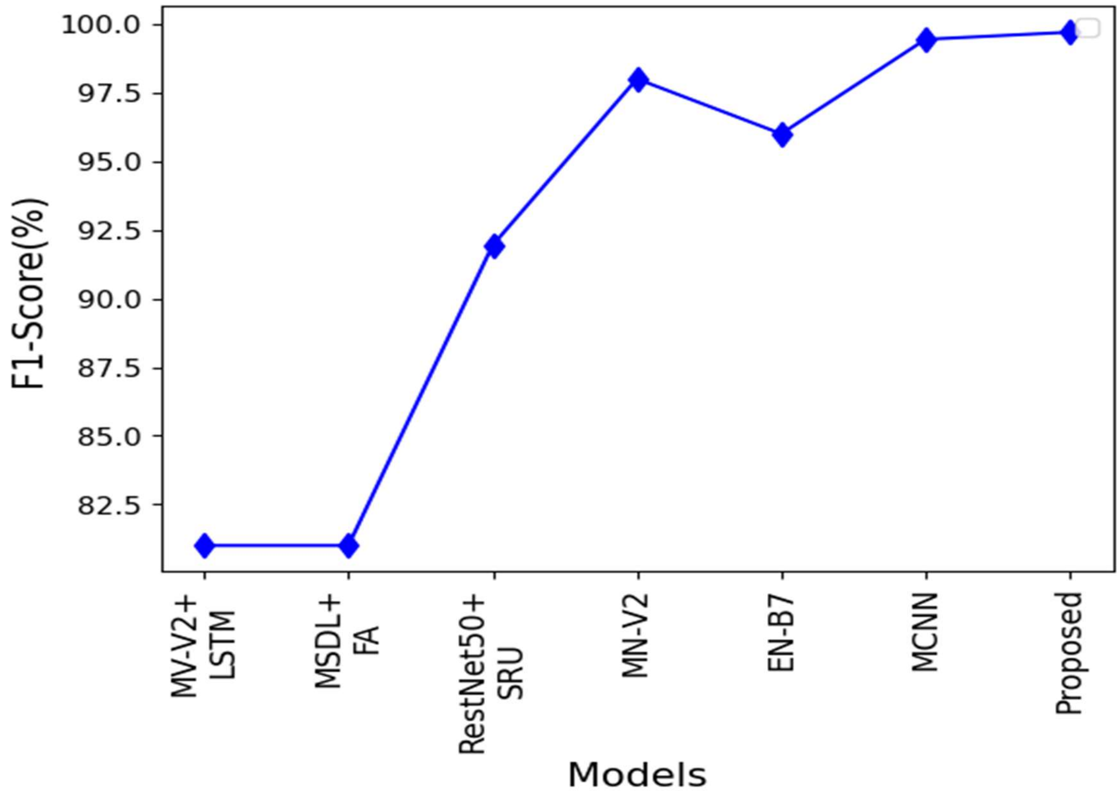
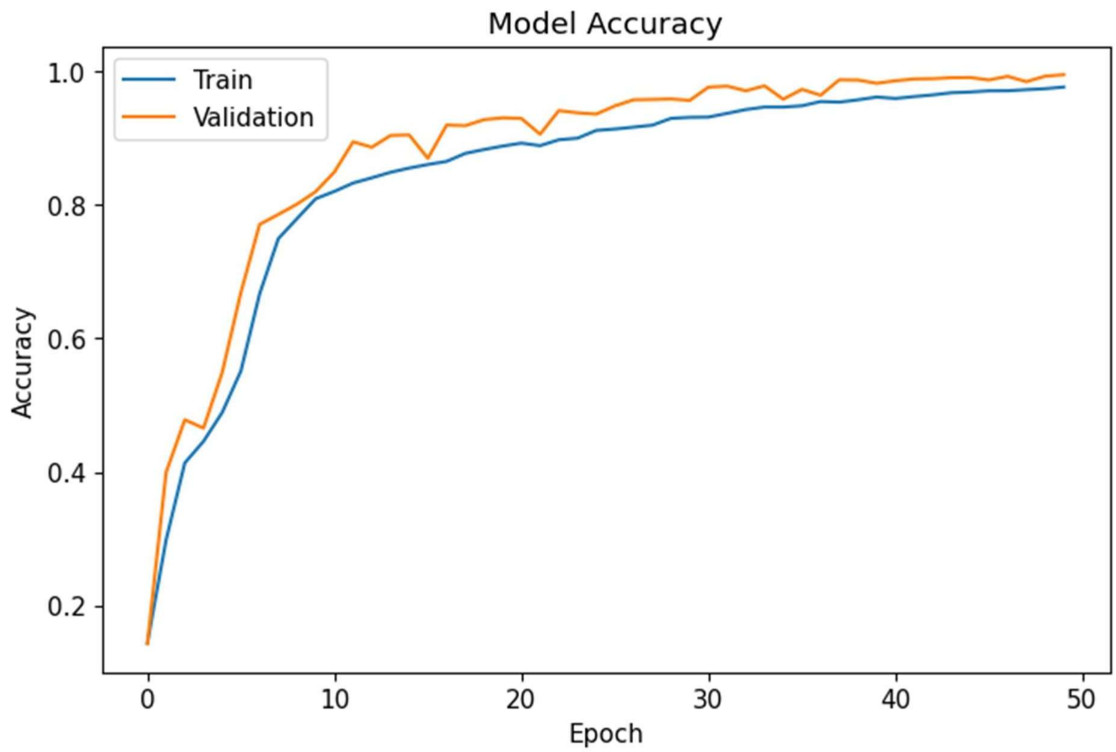
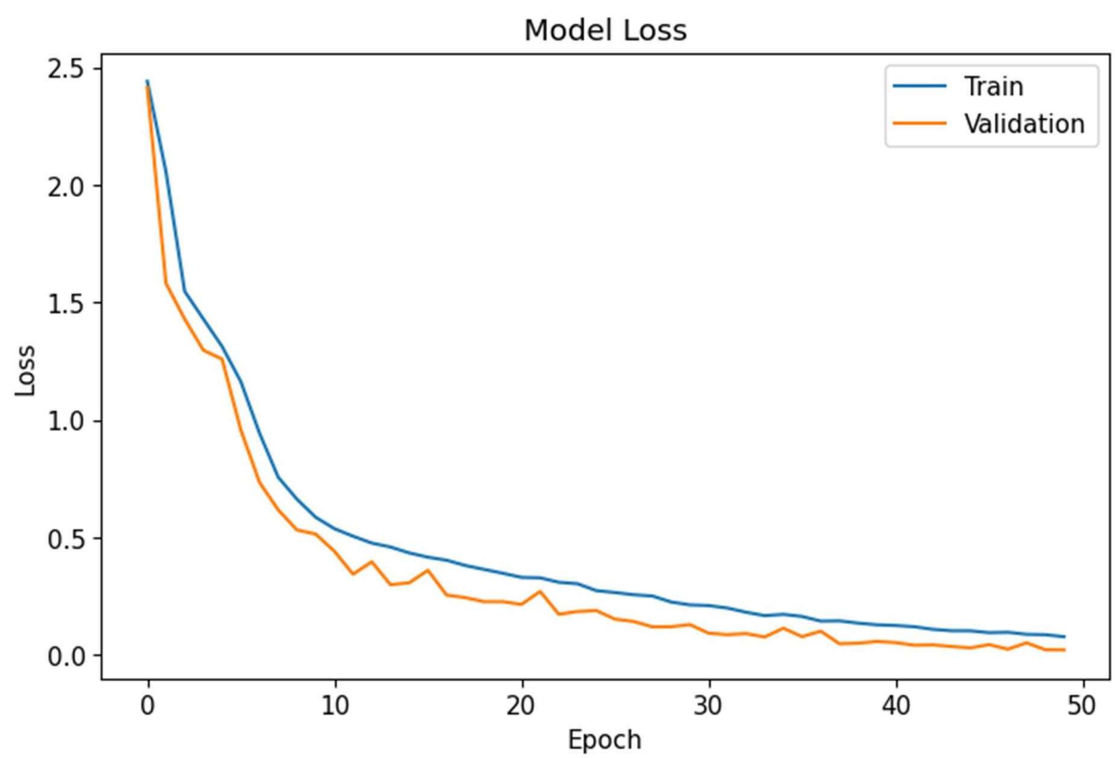


Figure 5: Training Accuracy and Loss Graphs

These plots show how the model's performance evolved across epochs.

- **Training accuracy curve:** indicates how well the model learned to classify the training data over time.
- **Validation accuracy curve:** shows how well the model performs on unseen data during training.
- A consistent rise in both accuracy curves, coupled with minimal divergence, signifies good generalization.
- **Loss graphs** depict how error decreased as the model optimized its weights. A steady drop in both training and validation loss shows effective learning and minimal overfitting.


Model Accuracy
Train
Validation
Accuracy
Epoch
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50


Model Loss
Train
Validation
Loss
Epoch
0.0
0.5
1.0
1.5
2.0
2.5
0
10
20
30
40
50