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 five 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 5-Fold Cross Validation further ensures robustness by training and validating the model on different subsets of data, enhancing reliability.

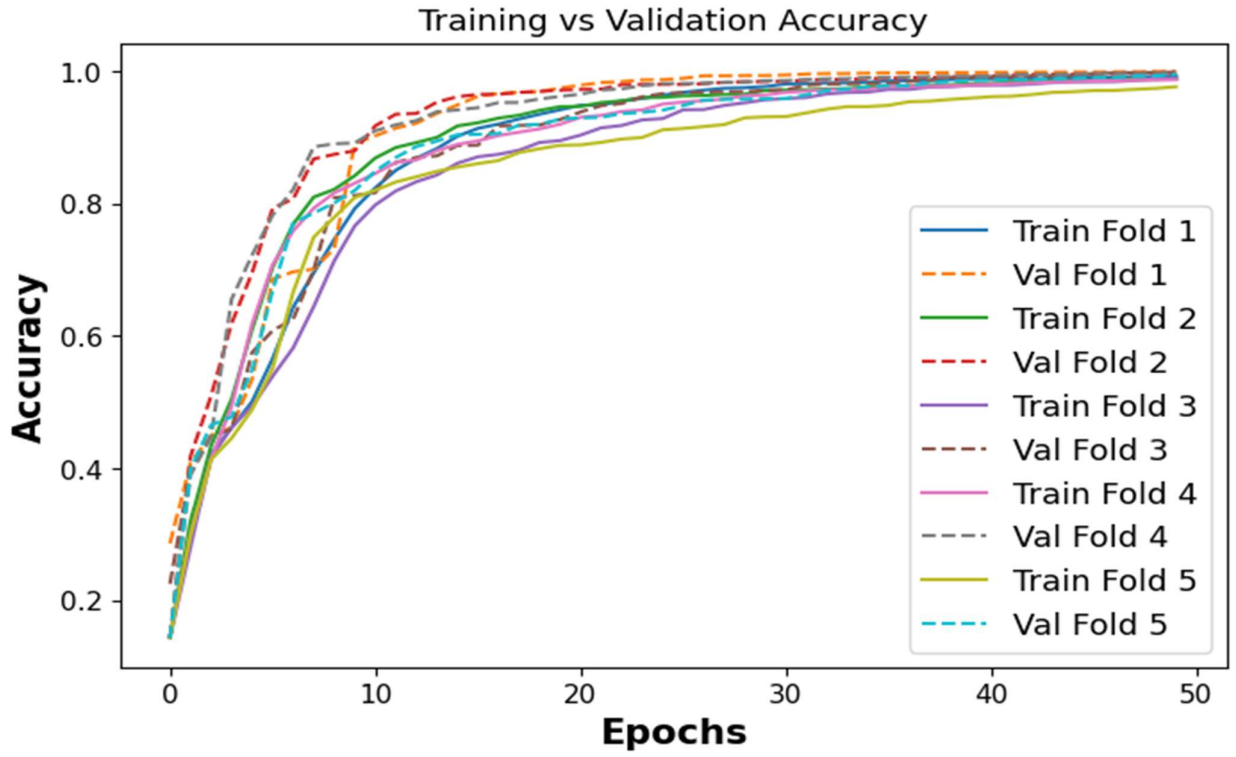


Figure 2: Training vs Validation Loss

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.

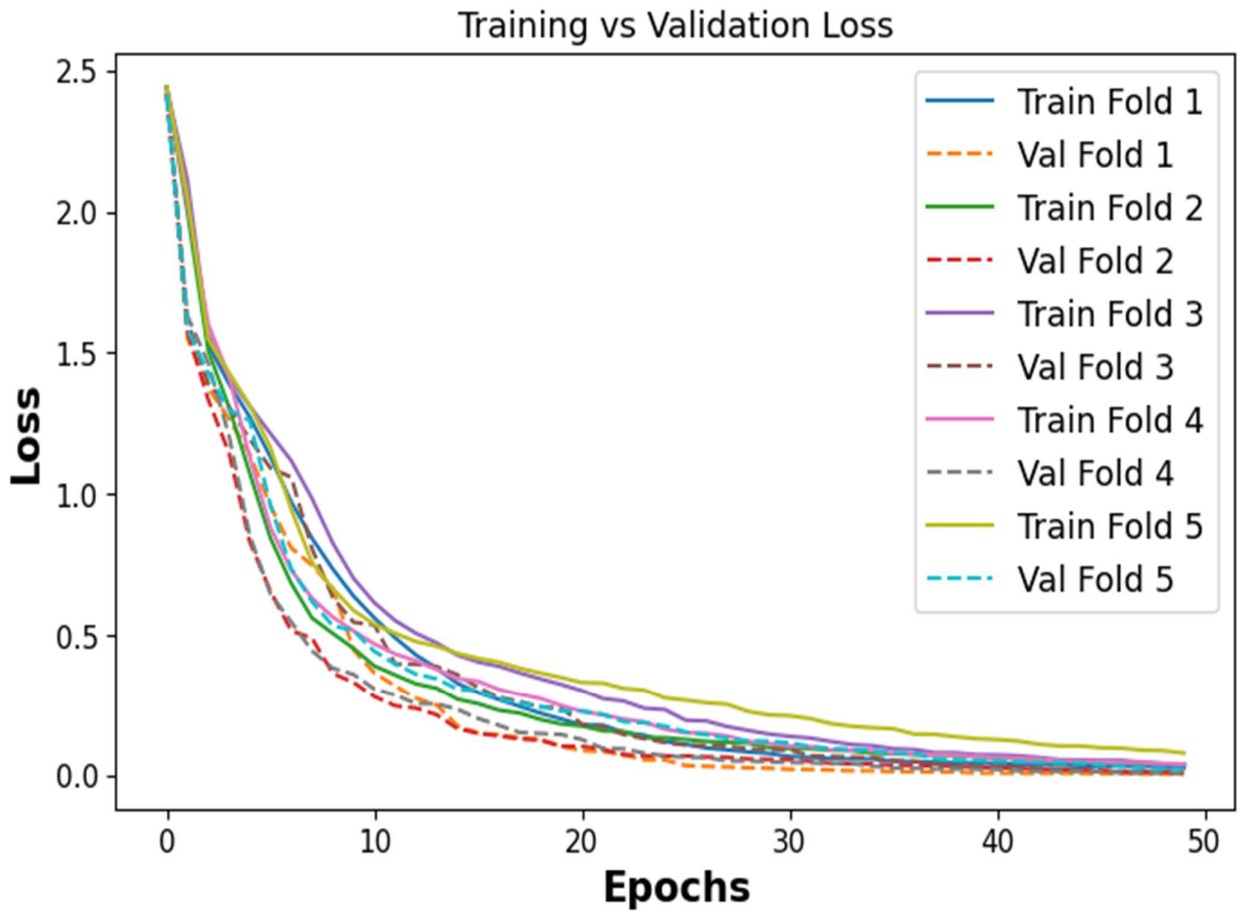


Figure 3:Input Frame Samples

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.

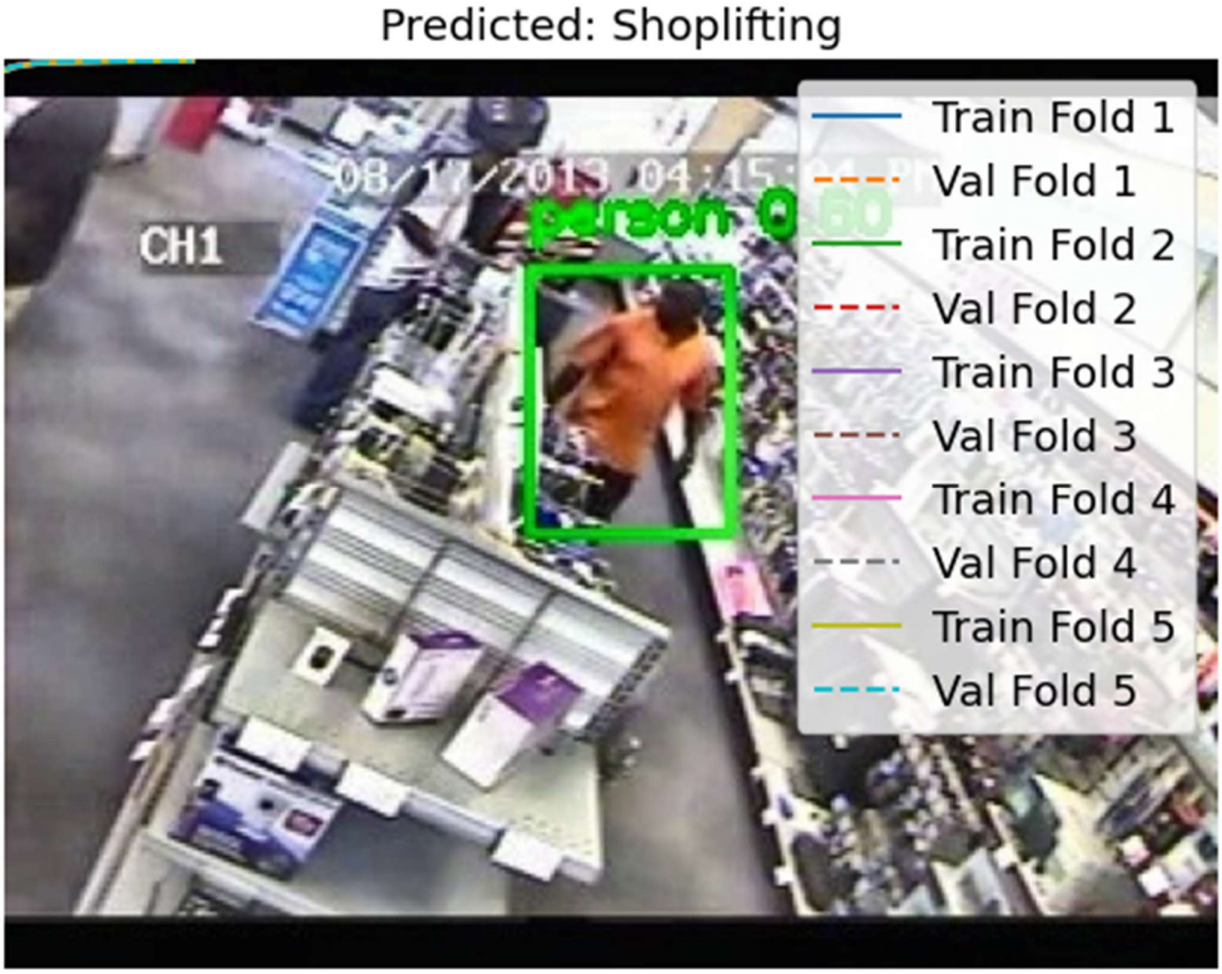


Figure 4:Confusion Matrix

This image displays the model's classification performance across different activity classes.

- Each cell shows the number of predictions per true label vs. predicted label.
- A strong **diagonal line** indicates high accuracy — most activities were correctly classified.
- Minor off-diagonal entries represent small misclassifications, often between visually similar behaviors (e.g., *Abuse* vs. *Arrest* due to similar human poses or environments).

- The matrix confirms that the model captures class distinctions effectively while maintaining robustness.

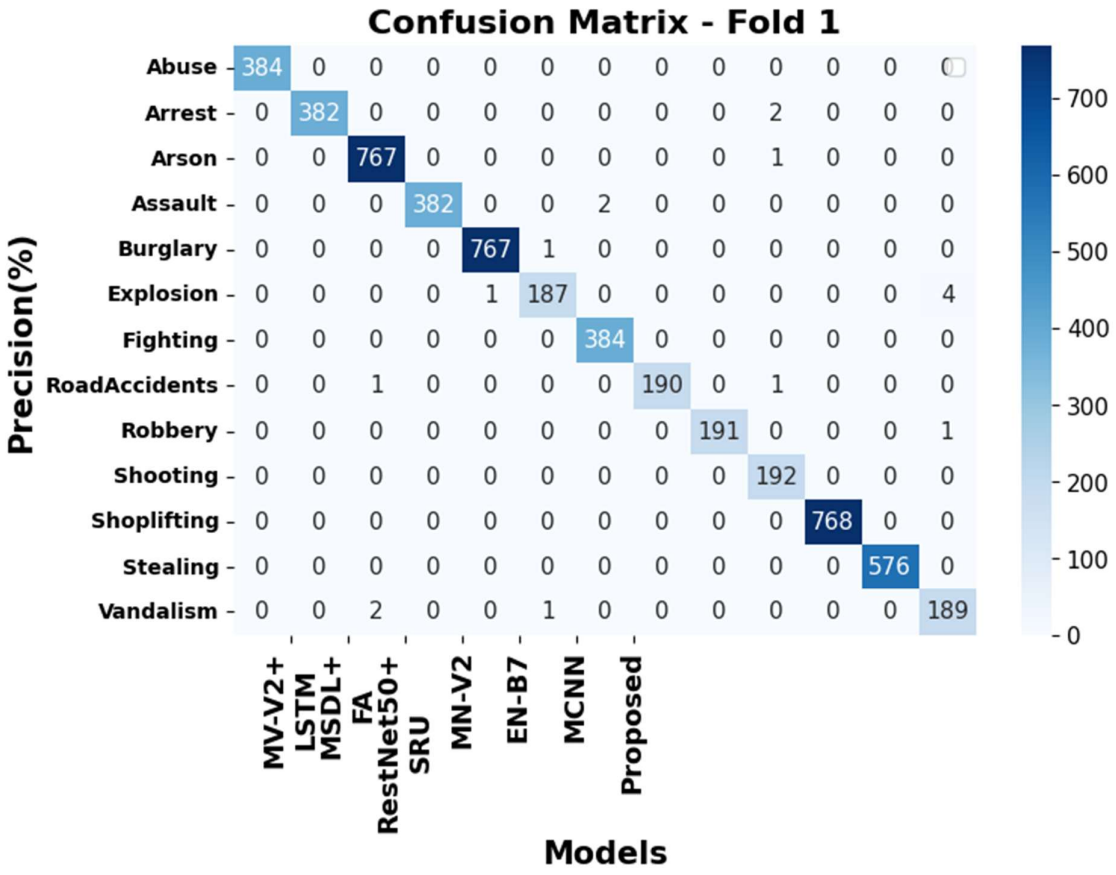
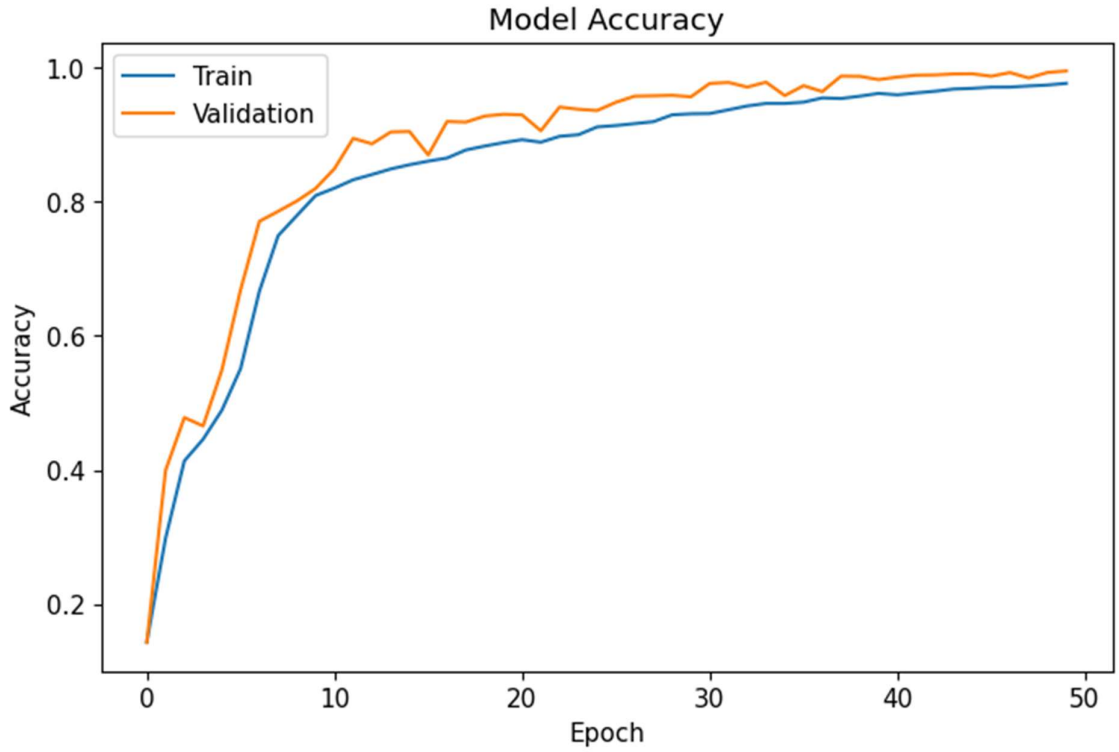
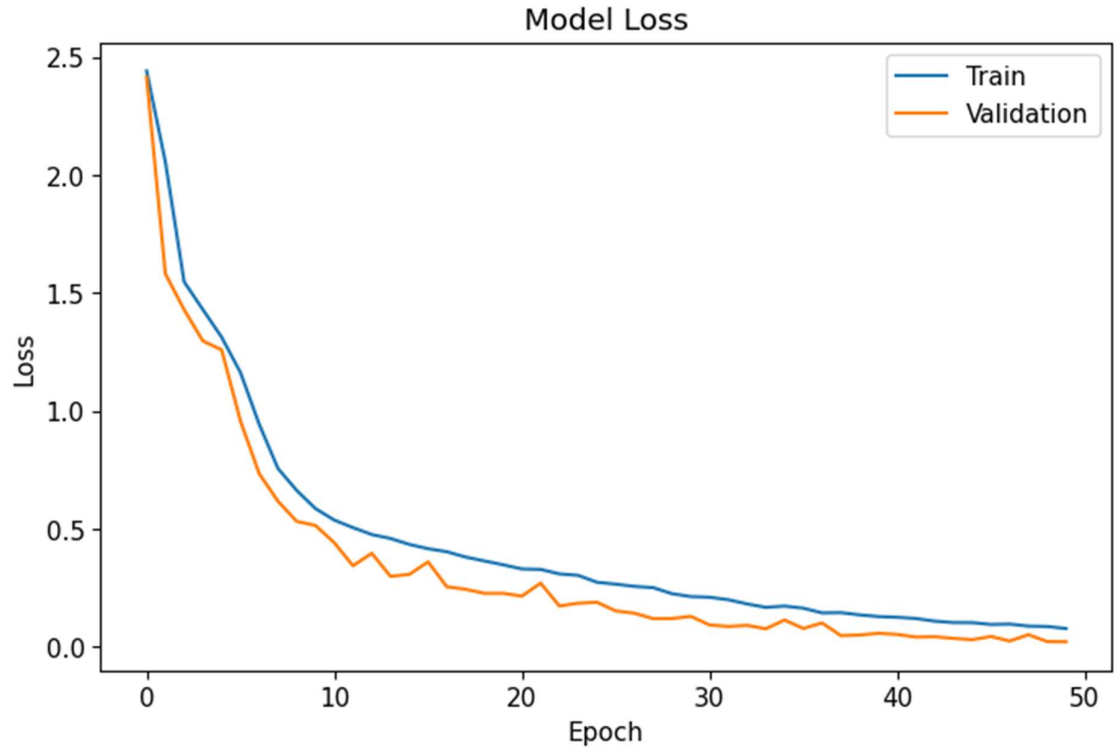


Figure 5: Training Accuracy and Loss Graphs

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
1.0
0.8
0.6
0.4
0.2
0
10
20
30
40
50


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