



National University of Computer and Emerging Sciences, Karachi  
FAST School of Computing, Spring 2025



## **Deep Learning for Perception Project Report**

**Project Title: Person Recognition Using Gait Matching**

**Group Members:**

**Mohammad Shahmeer Ul Haq 22K-4643 (Section B)**

**Armughan Ather Siddiqui 22K-4416 (Section C)**

**Roohan Ahmed 22K-4611 (Section C)**

### **1. Project Objective:**

The objective of this project is to develop a deep learning-based system for person recognition using gait patterns — a biometric that offers the advantage of being captured at a distance without requiring physical contact or user cooperation. We implement a custom hybrid architecture combining Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to extract both spatial and temporal features from sequences of silhouette frames. The model is trained and evaluated on the CASIA-B dataset under multiple walking conditions, including normal walking, walking with a bag, and walking in different clothing. In addition to recognizing known individuals, the system is adapted for **open-set recognition** by enabling the detection of **unknown individuals** during inference. Through this, we aim to investigate both the effectiveness and generalizability of gait-based recognition, particularly in handling variations in appearance as well as real-world scenarios where unseen individuals may appear.

## 2. Datasets Used:

We use the CASIA-B dataset, a benchmark dataset widely adopted in gait recognition research. It contains silhouette images of 124 individuals captured from 11 different camera angles under three types of walking conditions:

- **NM (Normal Walking):** Individuals walking without any external variation.
- **BG (Walking with a Bag):** Individuals carrying a bag, introducing changes to the gait silhouette.
- **CL (Walking with Different Clothing):** Individuals wearing different outfits, affecting the gait appearance.

Each subject has multiple sequences per condition, organized in a structured directory format. The dataset provides a comprehensive basis to test the robustness of gait recognition models against various real-world challenges.

In our project, we utilize data from 25 individuals selected from the full dataset to build and evaluate our gait recognition model.

**Dataset Source:** <https://www.kaggle.com/datasets/trnquanghuyn/casia-b/data>

## 3. Model(s) Used:

The model used for gait recognition is a hybrid deep learning architecture combining Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. This architecture was designed to effectively capture both spatial and temporal features from gait silhouettes.

## CNN + LSTM-based Sequence Model

- **Input Format:** The model takes input sequences of 16 grayscale silhouette frames, each resized to **64×44 pixels**.
- **Spatial Feature Extraction (TimeDistributed CNN):**
  - Each frame is passed through three convolutional blocks using TimeDistributed layers, enabling consistent CNN-based feature extraction across all time steps.
  - Each block consists of a Conv2D layer followed by BatchNormalization, ReLU activation, and MaxPooling.
  - After three such blocks, the feature map is flattened to a 1D feature vector for each frame, resulting in a feature sequence of shape (16, 2560).
- **Temporal Modeling (LSTM):**
  - The sequence of flattened feature vectors is processed by a two-layer LSTM stack.
  - The first LSTM layer outputs a sequence of 128-dimensional features for each time step.
  - The second LSTM layer outputs a fixed-length 64-dimensional embedding that summarizes the temporal dynamics of the full sequence.
- **Classification Head:**
  - A Dense(64) layer followed by Dropout is applied to the LSTM output.
  - A final Dense(20) layer with softmax activation is used for classifying among 20 known individuals.
- **Total Parameters:** Approximately 1.45 million, indicating a moderately complex yet efficient model suitable for sequence-based visual recognition.

This model was trained to recognize known individuals but is further adapted for open-set recognition by modifying the testing procedure to include unknown

individuals not present during training. The learned embeddings from the LSTM can be compared using cosine similarity or threshold-based rejection methods.

## 4. Results:

```
Starting Gait Recognition System with K-Fold Cross Validation
Loading and preprocessing data for K-Fold CV...
Processing known subjects sequences: 100%|██████████| 2190/2190 [00:29<00:00, 75.29it/s]
Processing unknown subjects sequences: 100%|██████████| 550/550 [00:07<00:00, 70.02it/s]
Loaded known subjects: X_all_known shape: (10336, 16, 64, 44), y_all_known shape: (10336,)
Loaded unknown subjects: X_all_unknown shape: (2804, 16, 64, 44)

--- Fold 1/5 ---
Training model...
Epoch 1/10
1034/1034 364s 415ms/step - accuracy: 0.0848 - loss: 2.9090 - val_accuracy: 0.3172 - val_loss: 2.0422 - learning_rate: 5.0000e-04
Epoch 2/10
1034/1034 424s 410ms/step - accuracy: 0.3497 - loss: 1.9637 - val_accuracy: 0.6983 - val_loss: 1.0934 - learning_rate: 5.0000e-04
Epoch 3/10
1034/1034 412s 398ms/step - accuracy: 0.6345 - loss: 1.1368 - val_accuracy: 0.7916 - val_loss: 0.6833 - learning_rate: 5.0000e-04
Epoch 4/10
1034/1034 384s 371ms/step - accuracy: 0.7995 - loss: 0.6574 - val_accuracy: 0.8951 - val_loss: 0.3487 - learning_rate: 5.0000e-04
Epoch 5/10
1034/1034 374s 362ms/step - accuracy: 0.8897 - loss: 0.4000 - val_accuracy: 0.9110 - val_loss: 0.2905 - learning_rate: 5.0000e-04
Epoch 6/10
1034/1034 374s 362ms/step - accuracy: 0.9078 - loss: 0.3387 - val_accuracy: 0.9420 - val_loss: 0.1730 - learning_rate: 5.0000e-04
Epoch 7/10
1034/1034 376s 364ms/step - accuracy: 0.9410 - loss: 0.2161 - val_accuracy: 0.9594 - val_loss: 0.1513 - learning_rate: 5.0000e-04
Epoch 8/10
1034/1034 377s 364ms/step - accuracy: 0.9383 - loss: 0.2167 - val_accuracy: 0.9188 - val_loss: 0.2715 - learning_rate: 5.0000e-04
Epoch 9/10
1034/1034 378s 366ms/step - accuracy: 0.9634 - loss: 0.1570 - val_accuracy: 0.9531 - val_loss: 0.1389 - learning_rate: 5.0000e-04
Epoch 10/10
1034/1034 380s 368ms/step - accuracy: 0.9634 - loss: 0.1398 - val_accuracy: 0.9715 - val_loss: 0.1080 - learning_rate: 5.0000e-04
Calculating unknown threshold...
```

```
Evaluating Fold 1:
Evaluating model...

Classification Report (Overall):
      precision    recall  f1-score   support

 Person_1       0.71     0.83     0.77       101
 Person_2       0.77     0.66     0.71       117
 Person_3       0.94     0.79     0.86       112
 Person_4       0.87     0.72     0.79       119
 Person_5       1.00     0.98     0.99        50
 Person_6       0.94     0.41     0.57        79
 Person_7       1.00     0.74     0.85       107
 Person_8       0.96     0.97     0.96       124
 Person_9       0.84     0.87     0.85       108
 Person_10      0.96     0.43     0.59       112
 Person_11      0.98     0.92     0.95       115
 Person_12      0.48     0.85     0.61       110
 Person_13      0.98     0.81     0.89       122
 Person_14      0.63     0.68     0.65        96
 Person_15      1.00     0.57     0.73       100
 Person_16      0.48     0.77     0.59       113
 Person_17      1.00     0.58     0.74       113
 Person_18      1.00     0.94     0.97        89
 Person_19      0.99     0.88     0.93        90
 Person_20      1.00     0.84     0.91        91
   Unknown      0.83     0.88     0.86      2804
```

```

    accuracy          0.83    4872
  macro avg         0.87    0.77    0.80    4872
 weighted avg         0.85    0.83    0.83    4872

Known vs Unknown Detection Report:
      precision    recall  f1-score   support

  Actually Known         0.82         0.76         0.79         2068
  Actually Unknown        0.83         0.88         0.86         2804

    accuracy          0.83    4872
  macro avg         0.83    0.82    0.82    4872
 weighted avg         0.83    0.83    0.83    4872

Fold 1 Accuracy: 0.8286
```

```

--- Fold 2/5 ---
Training model...
Epoch 1/10
1034/1034 395s 373ms/step - accuracy: 0.0714 - loss: 2.9445 - val_accuracy: 0.2806 - val_loss: 2.1165 - learning_rate: 5.0000e-04
Epoch 2/10
1034/1034 398s 385ms/step - accuracy: 0.3099 - loss: 2.0619 - val_accuracy: 0.6120 - val_loss: 1.2169 - learning_rate: 5.0000e-04
Epoch 3/10
1034/1034 390s 378ms/step - accuracy: 0.5873 - loss: 1.2559 - val_accuracy: 0.7862 - val_loss: 0.6705 - learning_rate: 5.0000e-04
Epoch 4/10
1034/1034 391s 378ms/step - accuracy: 0.7747 - loss: 0.7223 - val_accuracy: 0.8742 - val_loss: 0.4069 - learning_rate: 5.0000e-04
Epoch 5/10
1034/1034 394s 381ms/step - accuracy: 0.8668 - loss: 0.4704 - val_accuracy: 0.8887 - val_loss: 0.3246 - learning_rate: 5.0000e-04
Epoch 6/10
1034/1034 397s 383ms/step - accuracy: 0.9148 - loss: 0.3169 - val_accuracy: 0.9482 - val_loss: 0.1848 - learning_rate: 5.0000e-04
Epoch 7/10
1034/1034 405s 392ms/step - accuracy: 0.9423 - loss: 0.2193 - val_accuracy: 0.9468 - val_loss: 0.1805 - learning_rate: 5.0000e-04
Epoch 8/10
1034/1034 401s 388ms/step - accuracy: 0.9632 - loss: 0.1496 - val_accuracy: 0.9168 - val_loss: 0.2990 - learning_rate: 5.0000e-04
Epoch 9/10
1034/1034 399s 386ms/step - accuracy: 0.9661 - loss: 0.1458 - val_accuracy: 0.9773 - val_loss: 0.0820 - learning_rate: 5.0000e-04
Epoch 10/10
1034/1034 402s 389ms/step - accuracy: 0.9791 - loss: 0.1056 - val_accuracy: 0.9502 - val_loss: 0.1617 - learning_rate: 5.0000e-04
Calculating unknown threshold...
Unknown detection threshold set to: 0.8741

Evaluating Fold 2:
Evaluating model...
```

```

Classification Report (Overall):
      precision    recall  f1-score   support

  Person_1         0.73         0.49         0.59         106
  Person_2         0.80         0.71         0.75         114
  Person_3         0.99         0.75         0.85         129
  Person_4         0.91         0.70         0.79         107
  Person_5         1.00         0.98         0.99         58
  Person_6         0.90         0.09         0.17         98
  Person_7         1.00         0.80         0.89         97
  Person_8         1.00         1.00         1.00         102
  Person_9         0.81         0.83         0.82         114
  Person_10        0.75         0.68         0.71         119
  Person_11        1.00         0.79         0.88         100
  Person_12        0.29         0.90         0.44         121
  Person_13        0.99         0.87         0.92         112
  Person_14        0.97         0.60         0.74         110
  Person_15        0.98         0.48         0.65         104
  Person_16        0.59         0.90         0.71         106
  Person_17        0.98         0.49         0.66         95
  Person_18        1.00         0.85         0.92         82
  Person_19        0.94         0.65         0.77         89
  Person_20        1.00         0.84         0.91         104
   Unknown         0.80         0.85         0.82         2804

    accuracy          0.79    4871
  macro avg         0.88         0.73         0.76    4871
 weighted avg         0.83         0.79         0.79    4871
```

```
Known vs Unknown Detection Report:
      precision    recall  f1-score   support

  Actually Known      0.78      0.72      0.75      2067
  Actually Unknown    0.80      0.85      0.82      2804

   accuracy
 macro avg      0.79      0.78      0.79      4871
 weighted avg    0.79      0.79      0.79      4871

Fold 2 Accuracy: 0.7920

--- Fold 3/5 ---
Training model...
Epoch 1/10
1034/1034 402s 379ms/step - accuracy: 0.0803 - loss: 2.9315 - val_accuracy: 0.3227 - val_loss: 2.0889 - learning_rate: 5.0000e-04
Epoch 2/10
1034/1034 399s 386ms/step - accuracy: 0.3610 - loss: 1.9170 - val_accuracy: 0.6338 - val_loss: 1.1416 - learning_rate: 5.0000e-04
Epoch 3/10
1034/1034 401s 388ms/step - accuracy: 0.6470 - loss: 1.1020 - val_accuracy: 0.8234 - val_loss: 0.5548 - learning_rate: 5.0000e-04
Epoch 4/10
1034/1034 403s 390ms/step - accuracy: 0.7987 - loss: 0.6681 - val_accuracy: 0.8887 - val_loss: 0.3564 - learning_rate: 5.0000e-04
Epoch 5/10
1034/1034 405s 391ms/step - accuracy: 0.8969 - loss: 0.3934 - val_accuracy: 0.8902 - val_loss: 0.3468 - learning_rate: 5.0000e-04
Epoch 6/10
1034/1034 407s 394ms/step - accuracy: 0.9262 - loss: 0.2776 - val_accuracy: 0.9187 - val_loss: 0.2690 - learning_rate: 5.0000e-04
Epoch 7/10
1034/1034 412s 399ms/step - accuracy: 0.9213 - loss: 0.3080 - val_accuracy: 0.9444 - val_loss: 0.2068 - learning_rate: 5.0000e-04

Epoch 8/10
1034/1034 413s 399ms/step - accuracy: 0.9446 - loss: 0.2014 - val_accuracy: 0.9482 - val_loss: 0.2030 - learning_rate: 5.0000e-04
Epoch 9/10
1034/1034 413s 399ms/step - accuracy: 0.9643 - loss: 0.1384 - val_accuracy: 0.9318 - val_loss: 0.2394 - learning_rate: 5.0000e-04
Epoch 10/10
1034/1034 827s 801ms/step - accuracy: 0.9684 - loss: 0.1268 - val_accuracy: 0.9540 - val_loss: 0.1624 - learning_rate: 5.0000e-04
Calculating unknown threshold...
Unknown detection threshold set to: 1.0308

Evaluating Fold 3:
Evaluating model...

Classification Report (Overall):
      precision    recall  f1-score   support

  Person_1      0.71      0.52      0.60      93
  Person_2      0.71      0.74      0.73     123
  Person_3      0.98      0.87      0.92     103
  Person_4      0.99      0.65      0.78     120
  Person_5      0.98      0.86      0.92      58
  Person_6      0.95      0.23      0.37      92
  Person_7      1.00      0.84      0.91      95
  Person_8      0.99      0.94      0.96      95
  Person_9      0.87      0.78      0.82     110
  Person_10     0.71      0.47      0.57     111
  Person_11     0.99      0.90      0.94     116
  Person_12     0.45      0.96      0.61     121
  Person_13     1.00      0.80      0.89     128
  Person_14     0.62      0.68      0.65     118
  Person_15     1.00      0.47      0.64     115
```

Person_16	0.51	0.90	0.65	92
Person_17	0.99	0.59	0.74	120
Person_18	0.99	0.94	0.96	86
Person_19	0.98	0.73	0.84	82
Person_20	1.00	0.58	0.74	89
Unknown	0.81	0.87	0.84	2804
accuracy			0.81	4871
macro avg	0.87	0.73	0.77	4871
weighted avg	0.83	0.81	0.80	4871

Known vs Unknown Detection Report:

	precision	recall	f1-score	support
Actually Known	0.80	0.72	0.76	2067
Actually Unknown	0.81	0.87	0.84	2804
accuracy			0.81	4871
macro avg	0.81	0.80	0.80	4871
weighted avg	0.81	0.81	0.81	4871

Fold 3 Accuracy: 0.8060

--- Fold 4/5 ---

Training model...

Epoch 1/10  
**1034/1034** 418s 395ms/step - accuracy: 0.0780 - loss: 2.9255 - val\_accuracy: 0.2284 - val\_loss: 2.3888 - learning\_rate: 5.0000e-04

Epoch 2/10  
**1034/1034** 408s 395ms/step - accuracy: 0.2439 - loss: 2.2711 - val\_accuracy: 0.5119 - val\_loss: 1.5057 - learning\_rate: 5.0000e-04

Epoch 3/10  
**1034/1034** 410s 397ms/step - accuracy: 0.5053 - loss: 1.4791 - val\_accuracy: 0.6981 - val\_loss: 0.8636 - learning\_rate: 5.0000e-04

Epoch 4/10  
**1034/1034** 410s 397ms/step - accuracy: 0.7192 - loss: 0.8743 - val\_accuracy: 0.8534 - val\_loss: 0.4311 - learning\_rate: 5.0000e-04

Epoch 5/10  
**1034/1034** 414s 400ms/step - accuracy: 0.8306 - loss: 0.5458 - val\_accuracy: 0.9274 - val\_loss: 0.2271 - learning\_rate: 5.0000e-04

Epoch 6/10  
**1034/1034** 414s 400ms/step - accuracy: 0.8948 - loss: 0.3706 - val\_accuracy: 0.8815 - val\_loss: 0.3873 - learning\_rate: 5.0000e-04

Epoch 7/10  
**1034/1034** 414s 401ms/step - accuracy: 0.9254 - loss: 0.2588 - val\_accuracy: 0.9482 - val\_loss: 0.1904 - learning\_rate: 5.0000e-04

Epoch 8/10  
**1034/1034** 415s 401ms/step - accuracy: 0.9520 - loss: 0.1777 - val\_accuracy: 0.9497 - val\_loss: 0.1704 - learning\_rate: 5.0000e-04

Epoch 9/10  
**1034/1034** 416s 402ms/step - accuracy: 0.9518 - loss: 0.1899 - val\_accuracy: 0.9632 - val\_loss: 0.1112 - learning\_rate: 5.0000e-04

Epoch 10/10  
**1034/1034** 416s 402ms/step - accuracy: 0.9674 - loss: 0.1223 - val\_accuracy: 0.9497 - val\_loss: 0.1848 - learning\_rate: 5.0000e-04

Calculating unknown threshold...

Unknown detection threshold set to: 0.7408

Evaluating Fold 4:  
Evaluating model...

Classification Report (Overall):

	precision	recall	f1-score	support
Person_1	0.88	0.61	0.72	84
Person_2	0.88	0.54	0.67	95
Person_3	0.87	0.59	0.71	103
Person_4	1.00	0.69	0.81	128
Person_5	1.00	0.94	0.97	70

Known vs Unknown	Detection Report:			
	precision	recall	f1-score	support
Actually Known	0.77	0.65	0.70	2067
Actually Unknown	0.77	0.86	0.81	2804
accuracy			0.77	4871
macro avg	0.77	0.75	0.75	4871
weighted avg	0.77	0.77	0.76	4871

```

Fold 4 Accuracy: 0.7658

--- Fold 5/5 ---
Training model...
Epoch 1/10
1034/1034 ██████████ 435s 411ms/step - accuracy
Epoch 2/10
1034/1034 ██████████ 420s 406ms/step - accuracy
Epoch 3/10
1034/1034 ██████████ 420s 406ms/step - accuracy
Epoch 4/10
1034/1034 ██████████ 418s 404ms/step - accuracy
Epoch 5/10
1034/1034 ██████████ 413s 399ms/step - accuracy
Epoch 6/10
1034/1034 ██████████ 407s 394ms/step - accuracy
Epoch 7/10
1034/1034 ██████████ 404s 390ms/step - accuracy
Epoch 8/10
1034/1034 ██████████ 397s 384ms/step - accuracy

```



```

1034/1034 — 399s 386ms/step - accuracy: 0.9487 - loss: 0.2243 - val_accuracy: 0.9453 - val_loss: 0.1820 - learning_rate: 5.0000e-04
Epoch 10/10
1034/1034 — 391s 379ms/step - accuracy: 0.9557 - loss: 0.1615 - val_accuracy: 0.9773 - val_loss: 0.0797 - learning_rate: 5.0000e-04
calculating unknown threshold...
Unknown detection threshold set to: 1.0408

Evaluating Fold 5:
Evaluating model...

Classification Report (Overall):

```

	precision	recall	f1-score	support
Person_1	0.98	0.60	0.75	86
Person_2	0.50	0.79	0.61	121
Person_3	0.96	0.83	0.89	102
Person_4	0.96	0.67	0.79	123
Person_5	1.00	0.85	0.92	48
Person_6	0.87	0.66	0.75	88
Person_7	1.00	0.80	0.89	98
Person_8	1.00	0.97	0.99	113
Person_9	0.91	0.85	0.88	109
Person_10	0.89	0.50	0.64	102
Person_11	1.00	0.87	0.93	113
Person_12	0.43	0.89	0.58	116
Person_13	0.98	0.88	0.93	125
Person_14	0.84	0.78	0.81	132
Person_15	0.98	0.78	0.87	109
Person_16	0.66	0.65	0.66	97
Person_17	0.97	0.66	0.79	107
Person_18	1.00	0.93	0.96	82

```

Person_19    0.99    0.84    0.91    97
Person_20    1.00    0.76    0.86    99
Unknown      0.85    0.89    0.87   2804

    accuracy                0.84   4871
  macro avg    0.89    0.78    0.82   4871
weighted avg    0.86    0.84    0.84   4871

Known vs Unknown Detection Report:

```

	precision	recall	f1-score	support
Actually Known	0.83	0.78	0.81	2067
Actually Unknown	0.85	0.89	0.87	2804
accuracy			0.84	4871
macro avg	0.84	0.83	0.84	4871
weighted avg	0.84	0.84	0.84	4871

```

Fold 5 Accuracy: 0.8409
Saving model from fold 5
Model and parameters saved to gait_model_fold_5

--- K-Fold Cross Validation Results ---
Number of folds: 5
Average Accuracy: 0.8067 (+/- 0.0266)
Average Precision (Weighted): 0.8386 (+/- 0.0166)
Average Recall (Weighted): 0.8067 (+/- 0.0266)
Average F1-score (Weighted): 0.8069 (+/- 0.0281)

```

```

Individual Fold Accuracies:
Fold 1: 0.8286
Fold 2: 0.7920
Fold 3: 0.8060
Fold 4: 0.7658
Fold 5: 0.8409

```

## 5. Discussion on Results:

The gait recognition system, evaluated using a rigorous 5-fold cross-validation methodology, demonstrated promising performance in identifying known individuals and detecting unknown subjects. The K-fold approach provides a more robust estimation of the model's generalization capabilities compared to a single train-test split, offering insights into its stability and average performance across different subsets of the data.

### **Overall System Performance and Model Training:**

The system achieved an average overall accuracy of 80.67% across the 5 folds, with a standard deviation of  $\pm 2.66\%$ . This indicates a reasonably consistent performance, though the accuracy per fold varied, suggesting a balanced precision and recall overall.

### **State of the ART accuracy on CASIA B Dataset:**

The best reported accuracies on the CASIA-B dataset were achieved by **GaitSTR (2024)**, which attained an average accuracy of **94.7% under normal walking conditions, 98.4% when carrying a bag, and 96.2% while wearing a coat**. However, none of these models appeared to address **unknown subject detection**.

In contrast, the study *An Efficient Gait Recognition Method for Known and Unknown Covariate Conditions (2020)* specifically tackled both **known and unknown recognition scenarios**, reporting an average accuracy of **90.32%**.

Our system achieved an average accuracy between **80–84%**, which is commendable given the computational constraints we faced. Due to limited resources and long training times, we were only able to train the model for 10 epochs requiring approximately 7–8 hours. Attempts to go beyond this threshold often led to program crashes. Notably, increasing training from 5 to 10 epochs resulted in an approximate 6% gain in accuracy, suggesting that further training (e.g., up to 30 epochs) could likely yield significantly improved results.

Despite these limitations, the model demonstrated strong learning behaviour. During training, training accuracy consistently increased, and validation accuracy for known subjects also improved significantly. For example, by epoch 10, validation accuracy reached:

- **Fold 1:** 0.9715
- **Fold 2:** 0.9773
- **Fold 5:** 0.9773

These results indicate that our **CNN-LSTM architecture** effectively learned discriminative gait features for known individuals. Additionally, the integration of **EarlyStopping** and **ReduceLROnPlateau** (though not fully activated within 10 epochs) was beneficial in maintaining generalization and reducing the risk of overfitting.

### **Known Subject Identification:**

The primary task of identifying registered individuals showed varied success.

- **Strong Performance for Certain Individuals:** Several subjects (e.g., Person 5, Person 8, Person 18 across different folds) were identified with high precision and recall, often achieving F1-scores close to or at 1.00. This indicates that their gait patterns were distinct and well-learned by the model. For instance, in Fold 1, Person 5 achieved an F1-score of 0.99, and Person 8 achieved 0.96.
- **Challenges with Specific Individuals:** Conversely, the system struggled with certain individuals, as evidenced by lower F1-scores. For example:
  - Person 6 consistently showed low recall (e.g., 0.41 in Fold 1, 0.09 in Fold 2, 0.23 in Fold 3, 0.24 in Fold 4), indicating the model often failed to identify this person correctly.
  - Person 12 often had high recall but low precision (e.g., precision 0.48, recall 0.85 in Fold 1; precision 0.29, recall 0.90 in Fold 2), suggesting that other individuals were frequently misclassified as Person 12.
  - Person 16 also showed similar patterns of high recall but lower precision in some folds (e.g., Fold 1: precision 0.48, recall 0.77; Fold 4: precision 0.29, recall 0.86).
  - The variability in performance across different subjects (macro average F1 around 0.80 vs. weighted average F1 around 0.83-0.85 for the overall classification reports per fold) highlights that some gait signatures are inherently more challenging to distinguish or may have less discriminative features captured by the current model setup.

### **Unknown Subject Detection:**

A critical aspect of this system is its ability to identify individuals not enrolled in the known database. The performance in this area was consistently strong:

- The "Unknown" class generally achieved high F1-scores across the folds, typically ranging from 0.81 to 0.87 (e.g., Fold 1: 0.86, Fold 5: 0.87). This

demonstrates the system's effectiveness in correctly flagging unseen individuals.

- The "Known vs. Unknown Detection Report" further supports this. For instance, in Fold 1, "Actually Known" subjects were identified with an F1-score of 0.79, while "Actually Unknown" subjects achieved an F1-score of 0.86. Similar robust performance was observed in other folds (e.g., Fold 5: "Actually Known" F1 0.81, "Actually Unknown" F1 0.87). This indicates a good balance between correctly identifying known individuals and correctly rejecting unknown ones.
- The `unknown_threshold` for distinguishing known from unknown features varied across folds (e.g., 0.9569 in Fold 1, 0.7408 in Fold 4, 1.0408 in Fold 5). This is expected, as the threshold is dynamically calculated based on the feature distributions of the training and validation data within each specific fold. This variability suggests that the optimal separation boundary can be sensitive to the specific subset of known individuals used for training.

### **Impact of K-Fold Cross-Validation:**

The use of 5-fold cross-validation provides a more reliable and comprehensive assessment of the model's performance.

- It mitigates the bias of a single arbitrary train-test split and gives a better sense of how the model might perform on average on unseen data from the same distribution.
- The standard deviation in performance metrics (e.g.,  $\pm 2.66\%$  for accuracy) quantifies the model's stability across different training datasets.
- The detailed per-fold reports revealed subject-specific inconsistencies that might have been missed in a single evaluation run, offering valuable insights for model improvement.

### **Limitations and Areas for Improvement:**

While the results are encouraging, certain limitations are apparent:

**Computational Constraints and Epochs:** Training was restricted to 10 epochs per fold due to significant computational time (approximately 7-8 hours). While the model showed clear learning progress within these epochs, this limitation prevented exploration of potentially better performance or the observation of the onset of overfitting that might occur with more extensive training. Further optimization of the model or access to greater computational resources would be beneficial to explore longer training durations.

**Per-Subject Variability:** The significant difference in identification accuracy for some known individuals suggests that either their gait features are less unique, the data for these subjects is less representative, or the model struggles with particular gait patterns.