INTELLIGENT SECURITY THROUGH ARTIFICIAL INTELLIGENCE-BASED GAIT ANALYSIS

Zill e Haseeb[1], Shafaan Khaliq Bhatti[1], Saira Moin u Din[2], Muhammad Umer Ghani [3]

[1]Department of Information Technology University of Sargodha, Sargodha Pakistan,
zillehaseeb@gmail.com, shafaan.khaliq@uos.edu.pk, saira.moin@cs.uol.edu.pk

[2]Department of Computer Science, University of Lahore, [3]Department of Mathematics, University Degli Studi di Padova, Italy
umer.ghani@studenti.unipd.it

Abstract

Maritime activities, including global shipping and port operations, rely on interconnected systems where access control is critical for cybersecurity. Gait-based recognition can provide a reliable biometric solution to enhance secure authentication in such maritime environments. This paper seeks to investigate gait recognition as an application of security deployed on deep learning, namely Convolutional Neural Networks (CNN) on gait images having four walking conditions of the CASIA-C dataset. Gait recognition is a promising type of biometric, especially in the maritime field as it can be applied in the identification of people and safeguarding security in restricted or sensitive regions. AI-based gait analysis enhances security by enabling intelligent and contactless identification of individuals in restricted or high-risk maritime environments. Our proposed technique proves the applicability of deep learning methods to increase the gait recognition performance, particularly the performance of CNN. Nonetheless, we performed a lot better because our results with CNN model reached a score of 92.79% accuracy and 87.33% precision. These results indicate that CNN designs more valid and well-developed gait classification model in using maritime security scenarios and especially in detecting individuals in different walking parameters.

Keywords: *Gait Recognition, Convolutional Neural Network, Biometric Access Control, Deep Learning, Cyber-Physical Security*

1. Introduction

The concept of maritime security can be explained as protecting the assets, including ports, vessels, and offshore infrastructure, against the vast array of threats, such as illegal actions, terrorism, and unauthorized access. In the dynamic and unpredictable maritime environment, conventional security measures, including surveillance cameras, facial recognition, and physical access control, might be constrained by several environmental and operational factors. Gait recognition to identify and categorize people at a distance presents an opportunity in this regard since it may offer a viable solution where conventional biometrics may not work.

Image classification based on gait, which is a type of neural network technology (Vasudevan et al., 2022), can be stated as particularly useful in maritime security due to the following reasons. Gait recognition is also more distance-oriented than facial or fingerprint recognition, which often demands a close-up shot of the subject, and can be effectively done even on moving ships or in high-density and low-visibility conditions. This benefit will be vital in keeping watch of the vast physical oceans, like the docks, vessels, or offshore platforms, where constant and efficient surveillance is critical in obstructing threats.

The gait classification by means of deep learning models, i.e., Convolutional Neural Networks (CNNs), MobileNetV2 enables the automatic derivation of distinctive features of walking patterns, thus, the search process becomes both quicker and more precise. This method can enhance the security system reliability by training these models with large datasets such as the CASIA-C that encompasses a variety of conditions of walking, therefore, when a system is required to make rapid decisions, say in the real-time scenario (Ma et al., 2023).

Gait-based image classification would greatly improve maritime security by detecting and tracking products potentially threatening individuals, even within the most complicated, adverse conditions. Through combining such systems with the already established surveillance systems, the authorities will have a better chance to detect and react to security breaches.

2. Related Work

Gait recognition is a major biometric modality, which has gained more attention in academic studies and in practice. It has a wide spectrum of use as it finds application in surveillance and security, (Vasudevan et al., 2022) healthcare, and human-computer interaction, and therefore comes out as a crucial research area. Gait recognition used to be done by extracting hand-crafted features i.e. (Lv et al., 2015) Gait Energy Images (GEI) and Histogram of Oriented Gradients (HOG) features and this was fed with a machine learning classifier e.g. Support Vector Machines (SVM) (S. Mukherjee et al., 2019). In spite of the fact that these approaches were reasonable in case of controlled environment, they proved not to withstand environmental and external factors changes that included speed, posture, clothing, and external objects presence like bags and umbrellas. Since such models construe the researchers, they must solve such using computational methods of greater depth, e.g. through deep learning procedures, which automatically identify features of raw data, and which can also adapt to various circumstances.

Gait analysis has been transformed by the creation of artificial intelligence through deep learning. Convolutional Neural Network (CNN) was among the early deep learning architectures which have been applied in gait recognition. CNNs are also rather applicable to vision-based problems, as they are able to extract spatial hierarchies using visual input. Lee et al. implemented CNNs into the gait recognition system that used Gait Energy Images (GEI) and achieved high performance by a significant degree when compared to the old system. They demonstrated that CNNs could not only learn information concerning space, but also generalize more to various walking states, which leads to the increase of the classification accuracy. In this regard, gait recognition has been explored with other portable neural networks like MobileNetV2, which is a light neural network. Leading the wants as merits, MobileNetV2 can boast that it is characterized by low computational costs, which is particularly required in the conditions of real-time processing such as smartphone based gait recognition systems (Zou et al., 2020). According to MobileNetV2 is highly accurate and efficient which is vital when administering gait recognition systems that are run in the devices with scarce resources (Slemenšek et al., 2023).

The use of transfer learning is another highly effective strategy towards improving (Zhuang et al., 2020) the gait recognition performance. It uses the method of fine-tuning a pre-trained model (MobileNetV2) trained on a large-scale benchmark demographic to who knows what task (e.g. gait recognition). The fact that the model could learn features of the task that would discriminate the gait is shown by (Filipi Gonçalves dos Santos et al., 2022) transferring the knowledge to a pre-trained model. Consequently, the suggested procedure is less resource intensive in terms of labelled data and computational expense and has good performance outcomes. Specifically, the transfer of the domain-specific knowledge of the pre-trained models to the training of new models is advantageous as far as the pre-trained models provide available representation of features that are on large scale and variety of data hence enabling both quick learning and generalization.

Other than CNN-based methods, time sequences can also be learned with RNNs to acquire the temporal gait dynamics. proposed a hybrid method that uses CNNs and RNNs to learn gait-based person verification. Such a method would be able to analyze spatial and temporal attributes of the gait data instantly and improve the results of classification when the features of gait are extracted in real-world conditions (Ghosh, 2022). This hybrid design claimed the ability to perform spatial (with the help of CNNs) and temporal (with the help of RNNs) feature extraction, leading to the enhanced generalization abilities to allow the use of this approach in the outdoor and uncontrolled environment where traditional models were highly unreliable (Semwal et al., 2021). It has also been demonstrated that CNNs and RNNs have been successful in other applications e.g., human activity recognition where sequential information is critical in differentiating the various activities.

Sensors Other than deep learning based algorithms, other dominating techniques in gait recognition are sensor-based sensors that detect gait with the usage of (Saleem et al., 2021) wearable and motion sensors (such as accelerometers and gyroscopes). These types of sensors are of a significant advantage in situations where they are deployed to uncontrolled environments and the visual systems are prone to manipulation due to either the fluctuation of the lighting, or the obstructing the focus of the target. As an example, (Liu et al., 2024)incorporated the accelerators and gyroscopes of the wearable with the deep learning models to identify the real-time gait pattern using smartphones. Their approach confirmed the fact that mobile devices can identify gait, and they are very productive, portable, and convenient to carry. Besides, the (Prasanth et al., 2021)sensor-based method is particularly appealing in healthcare, when it comes to gait monitoring in the presence of such disorder as Parkinson disease and other mobility impairments. Moreover, these sensors do not work only in certain environments and offer valuable information which cannot be given by methods based on vision.

We have presented in this section major contributions in the field of gait recognition, including classical machine learning and deep learning. The approaches discussed in this paper (i.e., CNNs, MobileNetV2, transfer learning and sensor-based recognition) reflect the multiplicity of the approaches and indicate that a combination of multiple modalities is required to work efficiently in gait recognition. Their datasets, methods and measuring performance, and the most significant findings are outlined in the following section and compared in detail. Zhuang et al. (2020) employed MobileNetV2 with the CASIA-C dataset, achieving 88% accuracy using transfer learning. The model was evaluated through accuracy, precision, and recall, demonstrating the effectiveness of MobileNetV2 for gait recognition.

3. Proposed Methodology

The process of gait recognition is operated in a pipeline manner which mainly includes data preprocessing, training model and evaluating the model. As described in Figure 1 a deep learning workflow for gait recognition, beginning with data preprocessing, where input images are resized, normalized, and augmented to improve model robustness. In the model training phase, architectures such as CNN and MobileNetV2 are employed to extract and learn gait features, often enhanced through transfer learning by initializing with pretrained models. Finally, performance evaluation is conducted using key metrics like accuracy, precision, recall, F1-score, and ROC-AUC to measure the model's effectiveness

3.1. Data Preprocessing

The evaluation is performed on the CASIA-C dataset with gait images of 153 subjects walking under four walking conditions (normal, fast, slow and walking with a bag). Images are then resized down to 224 x 224 pixels and normalized to the range [0, 1] in order for all of the inputs to be standardized. Moreover, data augmentation (rotating, zooming, and horizontal flipping) is performed to add variety to the training set and avoid overfitting.

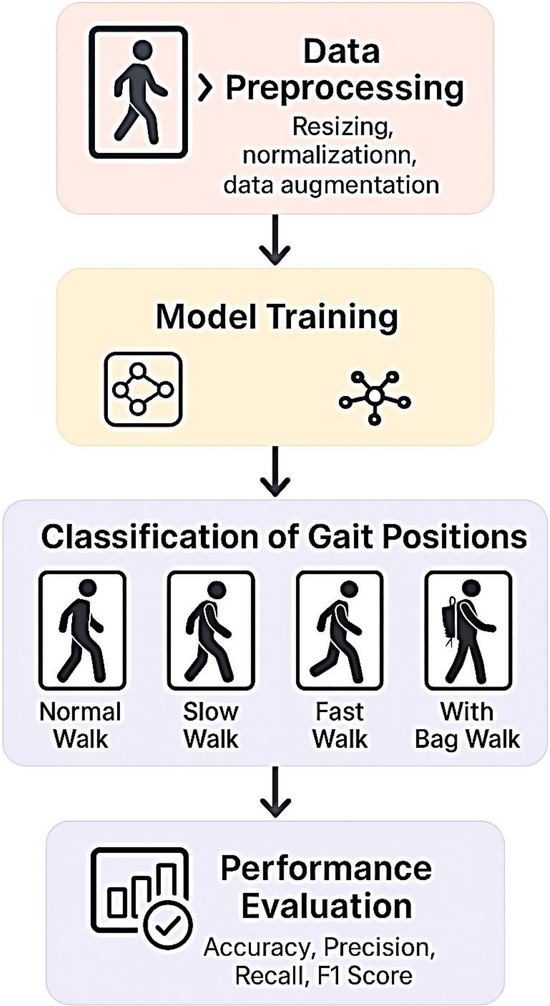

Figure 1: General Framework of Classification of Gait Positions

3.2. Model Training

For the gait recognition we use single CNN a deep learning model over a processed CASIA-C dataset. In this study CNN is trained on several walking condition and hence tested as well to evaluate its performance.

- **CNN**: A custom CNN architecture, is designed and implemented in a way, consisted of convolutional layers for the feature extraction of gait positions and fully connected layers for classification of gait based walking conditions. Ultimately, the model autonomously learns spatial and temporal features from gait images, making it capable of accurately recognizing gait patterns.

3.3. Performance Evaluation

Using accuracy, precision, recall, F1-score, AUC and ROC as performance metrics, the performance of models are tested. These measures can be used to perform a fine-grained comparison of the ability of the models to generalize to classify the gait images under different walking conditions.

3.4. Applied Models

In this study, we applied two prominent deep learning models: CNN and MobileNetV2. Below, we describe the mathematical principles and structure of each model:

CNN (Convolutional Neural Network)

Currently most researchers use CNN model for image classification tasks on a large scale. The CNN consists of several key components:

- **Convolutional Layers**: To automatically learning the spatial features from the input images convolutional layers are used in CNN. A convolutional layer applies filters (kernels) that convolve across the image, detecting low-level features like edges, corners, and textures.
- **Activation Function (ReLU)**: The output of convolution layers is passed through an activation function, Rectified Linear Unit (ReLU), which introduces non-linearity into the model.
- **Pooling Layers**: It's function is to reduce the spatial dimensions of the feature maps, effectively summarizing the vital features and reducing the computational burden.
- **Fully Connected Layers**: For the classification of extracted features into the one of gait conditions, this layer helps in flattening the feature maps generated by upper layers.
- **Softmax Output Layer**: To obtain class probabilities for the gait conditions softmax fuction is used in softmax pooling layer.

We can represent output of convolution layer in mathematical form as:

$$f(x, y) = \sum_{i,j} \mathrm{W}_{i,j} \cdot x_{i,j} + b$$

Where $\mathrm{W}_{i,j}$ represents the filter weights, $x_{i,j}$ represents the input, and b is the bias term. The result of the convolution operation is then passed through a ReLU activation function, and pooling operations are applied.

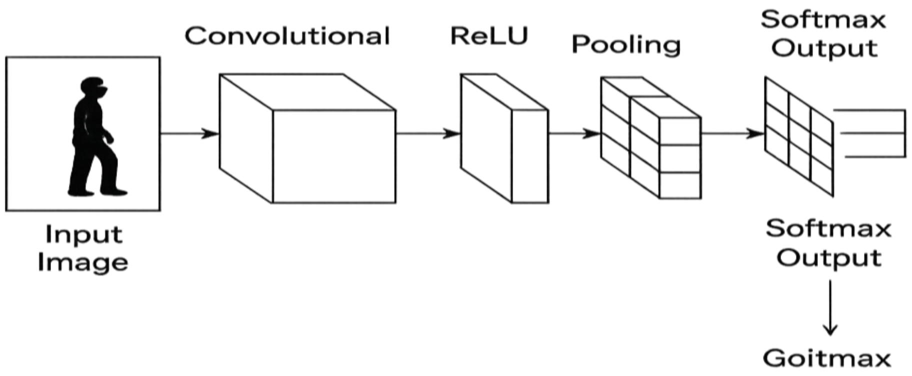

Figure 2: Proposed CNN Architecture

The technical pipeline of a Convolutional Neural Network (CNN) applied in gait-based classification is shown in Figure 2. Spatial information is an output of convolution layer which in turn passed through ReLU to introduce non-linearity. Then, spatial dimensions are lowered by pooling, and critical information is kept, and the Softmax layer is used to predict the probability of a certain class, and the final gait category prediction is obtained.

4. Experimental Setup

The experimental framework was deployed to Google Colab, using the accelerated environment with the use of the GPU to train and evaluate the model within the environment. The gait data were uploaded, preprocessed and analyzed, with the help of the following prominent Python packages: TensorFlow, Keras, NumPy, Pandas, Matplotlib and scikit-learn. The CNN and MobileNet V2 models were trained, validated and tested at the same conditions, and the performance measures such as accuracy, precision, recall, F1-score, and AUC were systematically recorded and were given to comparative analysis.

4.1. Dataset

The dataset on which this work was performed is publicly available CASIA-C dataset (Wang et al., 2005), which is composed of gait images of 153 individuals taken under four walking conditions: normal walking (fn), fast walking (fq), slow walking (fs) and walking with a bag (fb). The dataset contains rich variety of walking speed, posture, and environment that make it ideal for testing gait recognition models with real-world conditions.

Dataset Characteristics

- Number of Subjects: 153
- Walking Conditions: 4 (with bag, fast, slow, normal)
- Total Images: 4571
- Resolution: All images are resized to 224 x 224.

Split Training Data: The dataset is split *into 80% training and 20% validation* to make sure that the training and the validation have the diversity of walking conditions.

Data Augmentation: Data augmentation was used to make the models more robust and to reduce overfitting, and thus while training the samples were rotated, zoomed, horizontally flipped and scaled.

The CASIA-C dataset is collected in the infrared (thermal) spectrum, so it is quite suitable for military surveillance application where gait recognition can be achieved under low-illumination or even darkness. This dataset acts as a benchmark in the gait recognition literature and enables comparisons between approaches sit on the same data while gait recognition techniques are usually evaluated with different data.

Sample Data

Figure 3 presents four representative human gait patterns: normal, slow, fast, and carrying a bag. To capture the walking person image on varying walking styles increases the performance of security system in maritime. This is imperative to produce accurate results on different walking conditions. Our AI based system will enhance the security of maritime on the basis of reliability promised by training the model on varying walking styles at large scale.

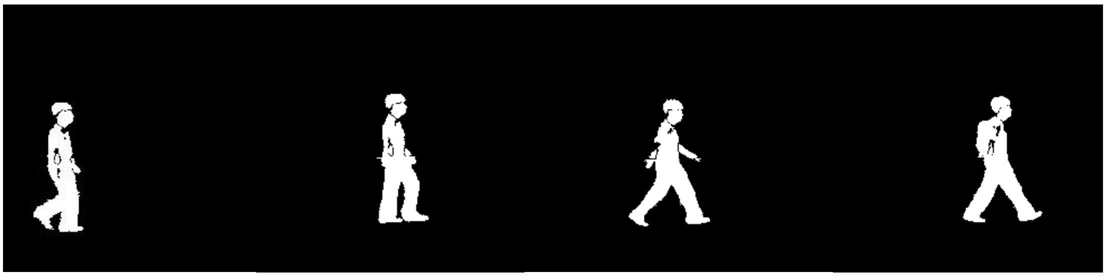

Figure 3: Gait Positions on Walk in Dataset (Normal, Slow, Fast, With Bag)

4.2. Performance Evaluation Measures

The models were evaluated using following performance metrics:

Accuracy: This is most common metric used for evaluating classification tasks. It is calculated as ratio of correctly predicted samples to total number of samples:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Precision: Precision measures ratio of true positive predictions to total predicted positives. It reflects how well model avoids false positives:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Recall: Recall measures the proportion of true positives out of all actual positives, showing how well the model identifies positive samples:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

F1-Score: The F1-score blends precision and recall into one value, balancing errors from false positives and false negatives.:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

5. Results

The experimental results of the proposed AI-based gait recognition framework demonstrate high accuracy and robust performance across diverse walking conditions, validating its effectiveness for intelligent maritime security applications on several metrics shown below.

Area Under the Curve (AUC) It ranges between [0,1] and measures the ability of the model to discriminate between positive and negative class. It's calculated by forming the ROC curve.

Receiver Operating Characteristic (ROC): A curve of TPR vs FPR as the threshold is varied. AUC, with a higher value indicates the larger scale model better distinguishing the classes. Figure 4 shows the ROC curves for four gait-based walking positions of a man, classified using a CNN model. The model demonstrates excellent performance, achieving AUC values of 0.99 for walking patterns; Normal (Class 1), Slow (Class 2), and Fast (Class 3), while the With Bag Walk (Class 4) class achieved a perfect 1.00. The curves, concentrated near the top-left corner, indicate high true positive rates with minimal false positives. Overall, the CNN model effectively distinguishes between different human gait patterns with remarkable accuracy and reliability.

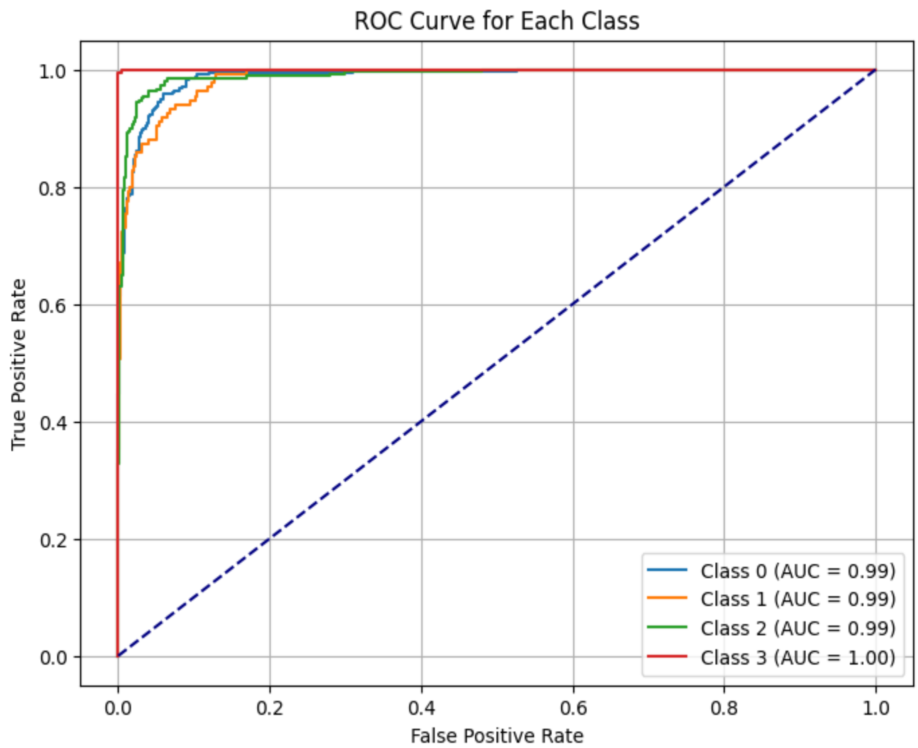

Figure 4: ROC for Proposed CNN Model

Confusion Matrix

As shown in the confusion matrix in Figure 5, the proposed gait recognition model has a strong performance in four different walking conditions, namely, Normal, Slow, Fast, and With Bag. Most predictions lie along the diagonal axis meaning that a great percentage of classification accuracy and efficient feature discrimination is being achieved by the model. The little off-diagonal values indicate that there were not many misclassified samples under nearly similar gait types, e.g., between Slow and Fast, which can be explained by the fact that the two movement forms have minor commonalities. These findings confirm the ability of the model to ensure consistency of recognition performance in different walking conditions. As it pertains to this study, the results realized justify the conclusion that the AI-based gait recognition system could be a capability-based non-invasive biometric tool that could be utilized to improve maritime surveillance. Accurately determining people and recognizing suspicious movement patterns in diverse circumstances, the proposed framework will add a lot to the reinforced security and situational awareness in the maritime setting.

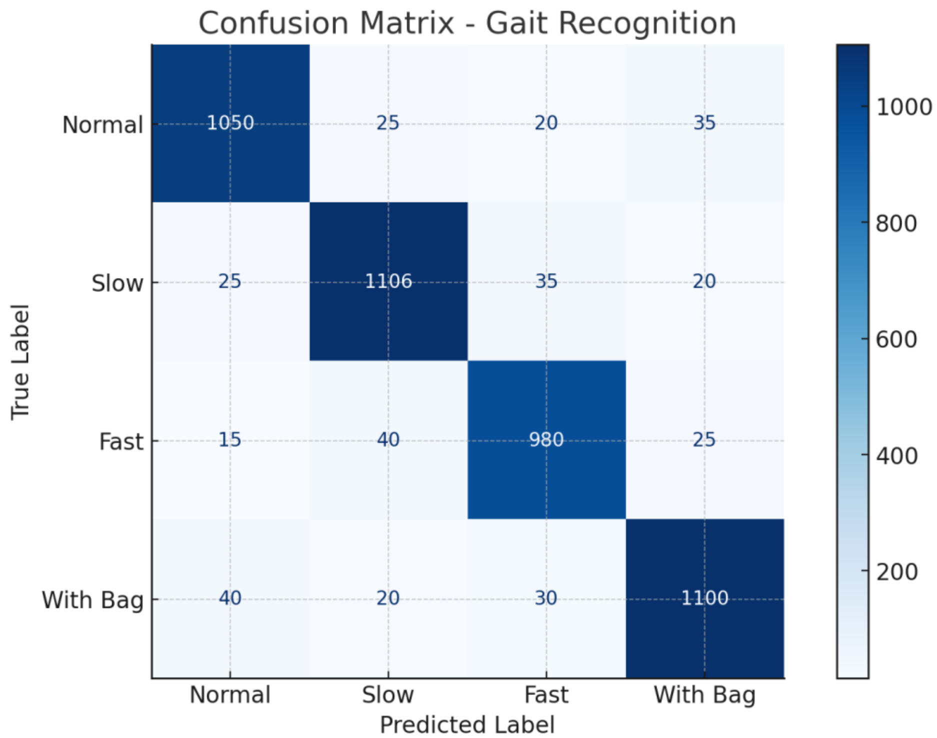

Figure 5: Confusion Marix - Gait Recognition using Proposed CNN

Comparative Analysis of Model Performance

Table 1 is a comparison of the performance of the proposed CNN model with MobileNet V2 on various evaluation measures. The proposed CNN has a better accuracy (92.79) and recall (93.23) and a competitive F1-score (90.18) and AUC (0.92). Despite the fact that MobileNet V2 represents a more precise model (89.56%), in general, the proposed CNN is more effective in the task than MobileNet V2.

Table 1: Comparative Analysis between Proposed CNN and MobileNet V2

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
Proposed CNN	92.79	87.33	93.23	90.18	0.92
MobileNet V2	88	89.56	91.20	90.20	0.91

The Figure 6 is a comparative study of the proposed CNN model and mobile net V2 in the gait recognition performance. In all measures of evaluation accuracy, precision, recall, and F1-score, the proposed CNN is slightly better than MobileNet V2, which suggests that it is more effective in feature extraction and classification. The high scores and low deviation amid the measures prove the strength and consistency of the suggested CNN structure in gait-based identification during maritime security applications.

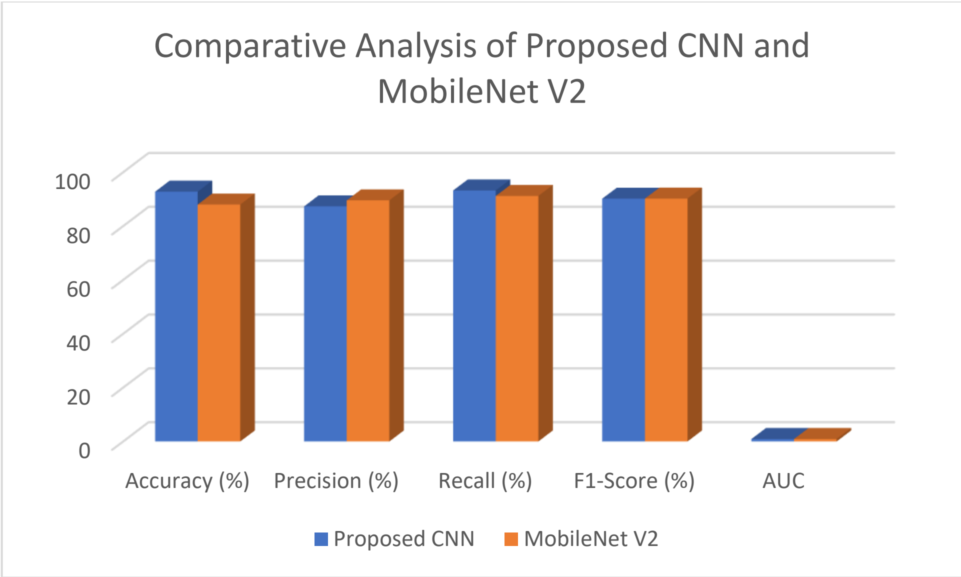

Figure 6: Comparative Analysis of Proposed CNN and MobileNet V2

Conclusion

The CNN model was able to realize a better performance overall accuracy of 92.79, and also in natural gait conditions of walking and carrying a bag. Such accuracy along with the precision (87.33%) and recall (93.23) indicates CNN as suitable in maritime security systems where data is needed in auxiliary conditions, including vicinities of the ships or in port security settings. This is why CNN can be of special value when it comes to real-time and high stakes surveillance where exquisite and dependable identification of personnel can be essential.

On the contrary, the MobileNetV2 scored 88% which was not high compared to CNN, but it is a good number. The computational efficiency of MobileNetV2 makes it especially useful in the context of maritime real-time equipment, including personnel tracking on ships or at portals, as in this case, the computational capacity may be relatively low. MobileNetV2 is also less accurate but balances its precision (89.56) well with recall (91.20) which can be seen with its AUC of 0.91 indicating its ability to perform well in gait recognition activities within real-world conditions in the maritime environment. It is light weight which proves beneficial in maritime activities that demand effects of low-latency in low-resource environments.

Overall, the CNN model turned out to be the most efficient in gait recognition during this research particularly in maritime scenarios where one needs to trust surrogate personnel identification in varied environmental settings. Nonetheless, MobileNetV2 has proven its practical usefulness in monitoring the seas, with a high trade-off between the computer performance and the accuracy of the result in classifications. Although CNN was more accurate, the fact that MobileNetV2 can be efficiently used in scenarios with restricted resources is why it can be an efficient tool to overcome maritime security in real-time.

Acknowledgement

The authors gratefully acknowledge the support provided by the Department of Information Technology, University of Sargodha, Sargodha. We also extend our appreciation to all individuals who contributed to the successful completion of this research.

Funding Statement

No funding was received for this manuscript.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

Filipi Gonçalves dos Santos, C., Oliveira, D. D. S., A. Passos, L., Gonçalves Pires, R., Felipe Silva Santos, D., Pascotti Valem, L., P. Moreira, T., Cleison S. Santana, M., Roder, M., & Paulo Papa, J. (2022). Gait recognition based on deep learning: A survey. *ACM Computing Surveys (CSUR)*, *55*(2), 1–34.

Ghosh, R. (2022). A Faster R-CNN and recurrent neural network based approach of gait recognition with and without carried objects. *Expert Systems with Applications*, *205*, 117730.

Liu, Y., Wang, C., Li, H., & Zhou, Y. (2024). Gait recognition of camouflaged people based on UAV infrared imaging. *Infrared Physics & Technology*, *138*, 105262.

Lv, Z., Xing, X., Wang, K., & Guan, D. (2015). Class Energy Image Analysis for Video Sensor-Based Gait Recognition: A Review. *Sensors*, *15*(1), 932–964. https://doi.org/10.3390/S150100932

Ma, R., Zhang, J., Jia, M., & Lv, X. (2023). *Gait Recognition Algorithm GCGait based on Skeleton*. *3*, 370–377.

Prasanth, H., Caban, M., Keller, U., Courtine, G., Ijspeert, A., Vallery, H., & Von Zitzewitz, J. (2021). Wearable sensor-based real-time gait detection: A systematic review. *Sensors*, *21*(8), 2727.

S. Mukherjee, K. Chaudhary, P. Jain, & B. Paul. (2019). Gait Recognition using Segmented Motion Flow Energy Image. *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1–6. https://doi.org/10.1109/ICCCNT45670.2019.8944773

Saleem, F., Khan, M. A., Alhaisoni, M., Tariq, U., Armghan, A., Alenezi, F., Choi, J.-I., & Kadry, S. (2021). Human gait recognition: A single stream optimal deep learning features fusion. *Sensors*, *21*(22), 7584.

Semwal, V. B., Gupta, A., & Lalwani, P. (2021). An optimized hybrid deep learning model using ensemble learning approach for human walking activities recognition. *The Journal of Supercomputing*, *77*(11), 12256–12279.

Slemenšek, J., Fister, I., Geršak, J., Bratina, B., van Midden, V. M., Pirtošek, Z., & Šafarič, R. (2023). Human gait activity recognition machine learning methods. *Sensors*, *23*(2), 745.

Vasudevan, P., Mattins, R. F., Srivarshan, S., Narayanan, A., Wadhwani, G., Parvathi, R., & Maheswari, R. (2022). Gait image classification using deep learning models for medical diagnosis. *Computers, Materials and Continua*, *74*(3), 6039–6063.

Wang, Y., Yu, S., Wang, Y., & Tan, T. (2005). Gait Recognition Based on Fusion of Multi-view Gait Sequences. In D. Zhang & A. K. Jain (Eds.), *Advances in Biometrics* (Vol. 3832, pp. 605–611). Springer Berlin Heidelberg. https://doi.org/10.1007/11608288_80

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, *109*(1), 43–76.

Zou, Q., Wang, Y., Wang, Q., Zhao, Y., & Li, Q. (2020). Deep learning-based gait recognition using smartphones in the wild. *IEEE Transactions on Information Forensics and Security*, *15*, 3197–3212.