

Article

Performance Evaluation of Various Deep Learning Models in Gait Recognition Using the CASIA-B Dataset

Nakib Aman ¹, Md. Rabiul Islam ¹, Md. Faysal Ahamed ² and Mominul Ahsan ^{3,*}¹ Department of Computer Science and Engineering, Rajshahi University of Engineering and Technology, Rajshahi 6204, Bangladesh; nakibaman@gmail.com (N.A.); rabiul_cse@cse.ruet.ac.bd (M.R.I.)² Department of Electrical & Computer Engineering, Rajshahi University of Engineering and Technology, Rajshahi 6204, Bangladesh; faysal.ahamed@ece.ruet.ac.bd³ Department of Computer Science, University of York, Deramore Lane, York YO10 5GH, UK

* Correspondence: mominul.ahsan2@gmail.com

Abstract: Human gait recognition (HGR) has been employed as a biometric technique for security purposes over the last decade. Various factors, including clothing, carrying items, and walking surfaces, can influence the performance of gait recognition. Additionally, identifying individuals from different viewpoints presents a significant challenge in HGR. Numerous conventional and deep learning techniques have been introduced in the literature for HGR, but traditional methods are not well suited to handling large datasets. This research explores the effectiveness of four deep learning models for gait identification in the CASIA B dataset: the convolutional neural network (CNN), multi-layer perceptron (MLP), self-organizing map (SOMs), and transfer learning with EfficientNet. The selected deep learning techniques offer robust feature extraction and the efficient handling of large datasets, making them ideal in enhancing the accuracy of gait recognition. The collection includes gait sequences from 10 individuals, with a total of 92,596 images that have been reduced to 64×64 pixels for uniformity. A modified model was developed by integrating sequential convolutional layers for detailed spatial feature extraction, followed by dense layers for classification, optimized through rigorous hyperparameter tuning and regularization techniques, resulting in an accuracy of 97.12% for the test set. This work enhances our understanding of deep learning methods in gait analysis, offering significant insights for choosing optimal models in security and surveillance applications.

Keywords: gait recognition; deep learning; convolutional neural network; multi-layer perceptron; self-organizing map; transfer learning; EfficientNet; CASIA B dataset



Citation: Aman, N.; Islam, M.R.; Ahamed, M.F.; Ahsan, M.

Performance Evaluation of Various Deep Learning Models in Gait Recognition Using the CASIA-B Dataset. *Technologies* **2024**, *12*, 264.

<https://doi.org/10.3390/technologies12120264>

Academic Editor: Juan Gabriel Avina-Cervantes

Received: 12 October 2024

Revised: 9 December 2024

Accepted: 12 December 2024

Published: 17 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Gait recognition is a biometric authentication method that has garnered significant interest in fields such as security, surveillance, healthcare, and human–computer interaction. Unlike traditional biometric methods such as fingerprints or iris scans, gait recognition leverages unique walking patterns for identification, offering benefits such as remote identification, resistance to spoofing, and applicability in scenarios where other biometrics are impractical [1].

The challenge in gait recognition lies in accurately identifying individuals based on their walking patterns, which involves extracting and analyzing gait features from video sequences. The initial research involved manual extraction and basic classification algorithms [1], but advancements in computer vision, machine learning, and sensor technologies have allowed gait recognition to evolve into a sophisticated biometric method. Specifically, deep learning techniques, such as convolutional neural networks (CNNs), have revolutionized gait detection by enabling the automatic extraction of distinctive characteristics from gait sequences [2]. The current research on deep learning for gait recognition explores various architectures, including CNNs, multi-layer perceptrons (MLPs), self-organizing maps (SOMs), and transfer learning models such as EfficientNet [3]. Each

architecture offers distinct advantages in terms of accuracy, computational efficiency, and interpretability. However, selecting the optimal model remains challenging because of factors such as dataset properties, feature representation, and training methodologies. Furthermore, the lack of standardized evaluation protocols complicates model comparison, making it difficult for researchers and practitioners to identify the most suitable approach for specific applications [4–6]. In security, gait recognition can be used for surveillance systems that identify individuals from a distance without the need for active participation, offering a non-intrusive alternative to traditional biometrics such as face or fingerprint recognition. In healthcare, it has the potential to assist in monitoring patients with mobility impairments or neurological disorders, providing valuable insights into movement patterns for diagnosis and treatment [1]. Additionally, advancements in gait recognition can enhance human–computer interaction, enabling more intuitive and natural interfaces for devices and systems. The methods explored in this study are highly relevant in advancing these applications by improving the accuracy and efficiency of gait recognition systems in real-world settings. This study addresses the problem of identifying the most effective deep learning model for gait recognition by evaluating the CNN, MLP, SOM, and EfficientNet architectures via the CASIA B dataset [6]. The authors aimed to compare these models based on their performance, strengths, and limitations to provide a clearer understanding of their applicability in gait identification tasks. Additionally, this research highlights the need for more refined models and optimization techniques to increase the robustness and accuracy of gait recognition systems.

This work bridges this research gap by offering a detailed comparative study of deep learning models, thus advancing the understanding of their effectiveness in gait recognition and providing directions for future improvements. Gait recognition, a biometric authentication method, has gained significant interest in several fields, such as security, surveillance, healthcare, and human–computer interfaces. Gait recognition uses the unique walking patterns of people for identification, unlike typical biometric methods such as fingerprints or iris scans. This unobtrusive method has several benefits, such as remote identification capabilities, resistance to spoofing assaults, and effectiveness in situations where other biometric methods may not be suitable. While the primary focus of this study is gait recognition for biometric applications, the findings also have potential implications for the neuroscience community. Gait analysis plays a crucial role in understanding motor control, neurological disorders such as Parkinson’s disease, and rehabilitation processes. The techniques developed in this study could aid in assessing and monitoring gait abnormalities, providing valuable insights into motor function and neurological health.

The research on human gait as a biometric identifier has been ongoing for several decades, originally including the manual extraction of gait data and basic classification algorithms [1]. Advancements in computer vision, machine learning, and sensor technologies have advanced gait recognition, developing it into a complex and effective biometric method. Deep learning methods, namely, convolutional neural networks (CNNs), have transformed gait detection by allowing the automatic extraction of distinctive characteristics from gait sequences recorded by surveillance cameras or sensors.

Researchers have investigated several deep learning architectures for gait detection tasks because of the growing need for strong and effective biometric identification solutions. The designs vary from basic multi-layer perceptrons (MLPs) to intricate models such as self-organizing maps (SOMs) and EfficientNet. Each model has distinct benefits in terms of accuracy, computing efficiency, and interpretability, making it an appropriate choice for various gait identification situations.

The specific research gap addressed in this paper involves the need for a more efficient and accurate deep learning model for human gait recognition, especially in challenging real-world scenarios. Prior studies focused on limited models, often overlooking the impact of environmental factors such as varying view angles, clothing styles, and carrying conditions, which reduce the reliability of gait recognition systems. This paper addresses this gap by evaluating four deep learning models—CNNs, MLP, SOMs, and

EfficientNet—on the CASIA-B dataset to determine their effectiveness in overcoming these real-world challenges. This is important because existing models lack the robustness required for practical applications, and the study's comparative analysis fills this gap by providing insights into the strengths and limitations of each model, contributing to the optimization of gait recognition systems for real-world use.

The innovative approach used in this research involves the comprehensive evaluation of multiple deep learning models, particularly focusing on how convolutional layers in CNNs and feature extraction techniques in EfficientNet can improve gait recognition performance. The use of a custom CNN architecture with additional convolutional layers allows better feature extraction, while EfficientNet introduces a novel approach to transfer learning in gait recognition. Furthermore, the integration of regularization techniques such as dropout layers and the fine-tuning of hyperparameters into the model enhances its robustness. These techniques differ from those of previous studies by optimizing both model accuracy and computational efficiency, aiming to balance real-time performance with high accuracy.

The results of this study could have practical implications for security, surveillance, and biometric systems, where the accurate and real-time identification of individuals based on their gait is crucial. The enhanced CNN model developed in this research could be highly suited to real-world applications, where conditions such as changing view angles, clothing variations, and environmental noise can degrade system performance. By optimizing model accuracy and computational efficiency, this research offers a scalable solution that can be implemented in real-time security systems, law enforcement tools, and healthcare monitoring, making it a valuable contribution to the study of the practical deployment of gait recognition technology.

Problem Statement: This research addresses several challenges in gait recognition, including recognizing gaits under varying conditions, such as changing view angles and walking styles, and carrying items such as a coat or bag, as well as the following:

- Addressing the issue of distinct yet sometimes similar gaits among different subjects can lead to misclassification and reduced system performance.
- Tackling the occasional failure of the two-step process of subject detection followed by classification because of incorrect subject detection and increased computational time.
- Dealing with variations in clothing styles, making it difficult to extract the rich features necessary for accurate classification.
- Managing irrelevant feature information extracted from original frames, which impacts the system accuracy and increases the computational time.

Contributions: The significant contributions of this study are as follows:

- A comparative analysis of four deep learning models for gait recognition;
- Insights into the strengths and weaknesses of each model;
- Recommendations for optimizing deep learning architectures for gait identification;
- An evaluation of model performance on a benchmark dataset is needed to guide future research in selecting appropriate models for practical applications.

This method stands out from prior methods, such as the single-stream deep learning approach highlighted in [2], as it not only enhances the accuracy but also improves the system's robustness across various environmental conditions. Moreover, it addresses the computational inefficiencies found in traditional silhouette-based approaches by optimizing the computational load without sacrificing accuracy. This makes the framework more suitable for real-time applications, representing a significant advancement in the field of human gait recognition. The purpose of this study was to evaluate and compare various deep learning algorithms for human gait recognition (HGR) using the CASIA-B dataset. While we propose a modified CNN architecture, the main contribution lies in determining the most effective algorithm for HGR under different conditions, providing valuable insights into model selection for practical applications. The novelty of this work lies in its systematic evaluation of several models, including CNNs, MLP, SOMs, and EfficientNet,

to determine the most effective algorithm for gait recognition under varying conditions (viewpoint variations, clothing variation, carrying conditions, occlusions, walking speed changes, etc.), offering significant improvements in accuracy and computational efficiency.

Structurally, this work is organized as follows. Section 2 reviews the relevant literature on human gait recognition (HGR), highlighting the evolution of biometric techniques and the challenges associated with gait analysis. Section 3 details the methodology employed, including the design and development of the deep learning models and the preparation of the CASIA B dataset. Section 4 presents the experimental setup and results, offering a comparative analysis of the performance of the CNN, MLP, SOM, and transfer learning models. Section 5 discusses the implications of these findings, their impact on the field of biometric security, and potential applications and concludes this research by summarizing key insights, addressing limitations, and suggesting directions for future work.

2. Recent Advancements

This paper conducts a thorough investigation evaluating CNNs [2], EfficientNet [3], MLP [4], and SOM [5] models for gait identification in the face of these obstacles. To analyze the performance, strengths, and limits of each model, the CASIA-B dataset [6], a commonly utilized benchmark dataset in gait recognition, was used. Our goal was to enhance the understanding of deep-learning-based gait recognition systems and guide the selection of suitable model architectures for practical use by comparing these models in different experimental settings.

Claudio Filipi Gonçalves dos et al. [7] investigated the application of deep learning in gait identification, a technique used to identify individuals by analyzing their walking patterns. This text emphasizes the advantages of deep learning over traditional biometric methods such as fingerprints and iris scans. It also discusses how deep learning may enhance feature extraction and accuracy. This research presents a historical analysis of biometric identification methods, contrasts gait recognition with traditional methods, and delineates the strengths and weaknesses of nine frequently utilized gait recognition datasets.

The text examines current studies that utilize deep learning for gait identification, focusing on its advantages and limitations. The research proposes the necessity of utilizing more varied and authentic datasets that consider elements such as perspectives, surroundings, and attention. It also promotes additional research on addressing the constraints of existing deep learning models in practical environments. This study is a great resource for academics and developers interested in deep-learning-based gait identification. This approach emphasizes the promise and constraints of this technology and calls for more research to guarantee strong performance in real-world scenarios.

A. Saboor et al. [8] focused on gait analysis via wearable sensors and machine learning (ML). The analysis of human movement provides valuable information in several sectors, such as healthcare, security, sports, and fitness. This study emphasizes the increase in research in machine learning, enabling the precise extraction of gait features. It also highlights problems such as small sample sizes, restricted computer capabilities, and issues regarding the energy economy. The report recommends that researchers select suitable algorithms, sensors, and locations on the basis of individual requirements and states that ideal sample sizes help minimize bias and lead to more robust findings. Regarding the future, the emphasis is placed on tackling unresolved issues such as generalizability and ethical concerns in data privacy.

B. Jawed et al. [9] explored the application of gait analysis for reidentification, with a specific focus on the last ten years. The benefits of gait analysis include it serving as a strong identification indicator and being suitable for surveillance applications that do not require collaboration. It recognizes the difficulties posed by current technology, including changes in perspective and fluctuations in illumination. The review classifies existing methods, evaluates the obtained results, and outlines future research paths. Possible applications include surveillance, security, forensics, investigation, and medical and fitness monitoring.

This research suggests that additional breakthroughs in this sector can be achieved by investigating new sensor technologies and overcoming existing restrictions. Human gait recognition (HGR) is a biometric identification technique that uses video recordings of individuals walking to recognize them. HGR does not rely on active participation, such as fingerprints or voice recognition, making it appealing for security and surveillance purposes. HGR has key advantages, such as being nonintrusive, distance-friendly, and disguise-resistant. Challenges arise from fluctuations in accuracy caused by variables such as clothes, perspective, and walking pace, as well as ethical concerns related to involuntary surveillance. Advancements in deep learning methods and 3D sensors have the potential to enhance precision and resilience. HGR provides a distinctive method for identification that has potential uses in security, forensics, and other fields. It is essential that privacy concerns are addressed and algorithms are enhanced to ensure ethical and successful deployment.

Zhang et al. [10] introduced a new approach to cross-view gait recognition via the Koopman operator theory, which captures the dynamic nature of gait patterns. The method uses a convolutional variational autoencoder to extract features from various viewpoints and approximate the Koopman operator, which represents overall gait dynamics. This method offers solid physical interpretability and improved performance, as demonstrated in experiments on the OU-MVLP dataset. Further research is needed to investigate its robustness to clothing variations and environmental changes, compare its performance with other dynamical system theories, and explore its applications beyond gait recognition.

Mutlag et al. [11] analyzed the significance of feature extraction in image processing activities such as diagnosis, classification, and detection. This work offers a detailed discussion of several feature extraction approaches, such as geometric, statistical, texture, and color characteristics. The authors analyze the performance of various features in geometry and texture tasks by conducting a comparative study utilizing face and plant picture datasets. The investigation revealed that the best qualities may differ on the basis of the type of image. This research highlights the importance of feature extraction in image processing, possible enhancements in feature selection, and the relevance of taking picture-specific factors into account in decision-making. The article is a great resource for understanding and choosing feature extraction strategies.

Pandey et al. [12] explored a technique for recognizing individuals by analyzing their walking pattern, utilizing sparse representation. The approach, which uses sparse combinations of established gait patterns, was evaluated on the CASIA-B database with individuals and five distinct viewing angles. The findings indicated an average identification rate of 96.5% for all angles, with the peak rate being 98.1% for a particular angle. The approach is invariant with regard to view changes, resilient, and simple to execute, making it beneficial for surveillance purposes. This study proposes enhancements by specifying the sparse representation method employed, indicating the size of the training dataset, and providing comparisons with alternative gait detection systems.

Pansambal et al. [13] analyzed a model-free method for human gait recognition (HGR), a biometric system that recognizes individuals by their walking style. The focus is on three main aspects: motion-free picture representation, dimensionality reduction, and classification techniques. This report examines publicly accessible gait datasets for research and development, outlines existing research obstacles in model-free human gait recognition (HGR), and suggests future paths for enhancing this field of biometrics. The intended audience comprises academics and developers who are interested in model-free hand gesture recognition methods. Possible enhancements include discussing well-known feature extraction techniques that do not rely on models, emphasizing particular dimensionality reduction approaches, and sharing instances of popular classification algorithms.

Wu et al. [14] presented an in-depth analysis of human identification via cross-view gait recognition. It leverages deep convolutional neural networks (CNNs) to address the challenges associated with varying viewpoints. This study explores multiple CNN architectures and training strategies to increase recognition accuracy. The authors introduce a novel gait representation that captures both spatial and temporal features. Extensive

experiments demonstrate significant improvements over traditional methods, validating the effectiveness of deep learning approaches in handling cross-view variations. This paper also discusses the impact of different network parameters and the importance of large-scale gait datasets. This research contributes to advancing the robustness and applicability of gait-based human identification systems across diverse real-world scenarios.

Chao et al. [15] introduced a novel approach to cross-view gait recognition by treating gait sequences as sets. This method leverages the set-based perspective to address the variations in viewing angles. The GaitSet model uses a deep learning framework to extract comprehensive gait features, which are then aggregated to improve recognition performance. The authors conducted extensive experiments on large-scale datasets, demonstrating that GaitSet significantly outperforms existing methods. The paper also highlights the robustness of the approach in handling diverse scenarios and complex backgrounds. The proposed model is noted for its efficiency and effectiveness in real-world applications, contributing to the advancement of gait recognition technology. This work represents a substantial step forward in the field, emphasizing the importance of innovative feature representation techniques.

Mehmood et al. [16] presented an innovative approach to human gait recognition via pre-trained convolutional neural networks (CNNs). The proposed system is an end-to-end solution that leverages pre-trained CNN models to extract and select relevant gait features. This method enhances recognition accuracy by utilizing the powerful feature extraction capabilities inherent in pre-trained models. The authors conducted extensive experiments to validate their approach, which demonstrated significant improvements in recognition performance compared with traditional methods. Their system is efficient, showing robustness across different datasets and varying conditions. This study highlights the importance of feature selection in improving the overall effectiveness of gait recognition systems. This work contributes to the field by providing a scalable and accurate solution for gait-based human identification.

Several recent studies have explored deep learning models for gait recognition using the CASIA-B dataset. For example, Pandey et al. [12] achieved an accuracy of 96.50% using a multiview sparse representation approach. Mehmood et al. [16] reported 94.26% accuracy with pre-trained CNNs. In comparison, the proposed CNN model achieved an accuracy of 97.12%, demonstrating improved performance due to the optimized architecture and hyperparameter tuning. This enhancement addresses challenges such as varying viewing angles and clothing conditions more effectively, contributing to the robustness and scalability of gait recognition systems. While prior research on the CASIA-B dataset has explored different architectures, such as CNNs and transfer learning models, few studies have performed direct comparisons of these models under varying environmental conditions. This study addresses this gap by evaluating four distinct deep learning architectures, optimizing them for computational efficiency and generalization across diverse gait patterns.

Currently, the research on human gait recognition is making significant strides, especially with the advent of deep learning techniques. However, notable gaps and limitations remain in this domain. Most existing methods, such as those utilizing single-stream deep learning models, often struggle with the variability introduced by different covariate factors, such as clothing changes, carrying conditions, and multiview environments. For example, Saleem F. et al. [17] focused primarily on feature fusion but did not fully address the challenge of recognizing gait across multiple views and varying conditions. Moreover, Asif M. et al. [18] attempted to address these issues, but they often relied on complex preprocessing steps, and their networks were computationally intensive, making them less suitable for real-time applications. These approaches also tend to perform sub-optimally when dealing with large, diverse datasets, limiting their scalability and practical utility.

Deng, M. et al. [19] introduced a novel frontal-view gait recognition method that enhances human gait recognition accuracy by leveraging gait dynamics and deep learning. Unlike most previous works, which rely on lateral-view parameters, the authors propose the use of frontal-view gait features—kinematic, spatial ratio, and area features. To fur-

ther improve recognition, temporal dynamics in human walking are captured, and deep learning techniques are incorporated to optimize performance. The method calculates similarities between the test and training gait dynamics and uses an error-based feature fusion scheme for robustness against walking variations. However, a gap exists in the paper's focus on improving frontal-view recognition, as challenges related to environmental factors in real-world applications, such as occlusions or lighting variations, are not extensively explored.

Hou, S. et al. [20] critically evaluated silhouette-based gait recognition systems, arguing that while they demonstrate high performance in academic settings, there are significant gaps when they are applied to real-world environments. The authors identify issues such as oversimplified evaluation protocols and sensitivity to noise from factors such as occlusions and rotations. They also introduce the Multi-Height Gait (MHG) dataset, which features walking data for different camera heights, clothing variations, and carried objects. This highlights the gap between academic performance and practical application, emphasizing the need for more complex, realistic evaluations. This work addresses the limitations of existing datasets but lacks comprehensive solutions for noise sensitivity and dataset augmentation.

Qin, L et al. [21] proposed a method utilizing two-stream convolutional neural networks (two-stream CNNs) for gait recognition based on multisensor data, such as inertial and pressure sensors. The innovative aspect of this work is the progressive feature fusion (PFF) module, which optimally integrates multistage features, and the feature enhancement channel attention (FECA) module, which highlights crucial data. The attention-based feature fusion (ABFF) module further enhances recognition by allowing flexible feature fusion across different sensor modalities. The results show superior recognition accuracy compared to previous approaches. A gap in this study is its limited discussion on how this multisensor approach could be generalized to uncontrolled, large-scale real-world environments, where sensor placement and variance could affect performance.

Sepas-Moghaddam, A. et al. [22] provided a comprehensive survey of recent developments in gait recognition, focusing on how deep learning has transformed this field since 2015. The authors propose a new taxonomy to categorize state-of-the-art methods based on body representation, temporal representation, feature representation, and neural architecture. They discuss current gait datasets, evaluation protocols, and the challenges in applying deep learning to gait recognition. The review highlights the performance improvements made using these methods but also emphasizes the challenges that remain, such as dataset limitations, scalability issues, and real-world applicability. Future research directions should be aimed toward more robust and adaptable solutions, although the study leaves open questions on how to address these challenges in practice. Previous studies have highlighted challenges such as sensitivity to viewpoint changes, occlusions, and clothing variations. The proposed model addresses these challenges by leveraging deeper convolutional layers and optimized feature extraction techniques, which enhance robustness and improve accuracy under diverse conditions. A comparative analysis of gait recognition is presented in Table 1.

Recent studies using the CASIA dataset, such as those by Vasudevan et al. [23] and Alotaibi et al. [24], have explored deep learning models for gait recognition, achieving promising results. However, these works have focused primarily on a single architecture or feature fusion methods. In contrast, our proposed method compares multiple deep learning models (CNNs, MLP, SOMs, and EfficientNet) and introduces a more comprehensive evaluation framework. Mehmood et al. [25] proposed a feature selection framework, while Gul et al. [26] incorporated spatio-temporal features in multiview gait recognition. Our approach builds on these methods by evaluating the performance of multiple models under varying conditions to provide a more generalized understanding of model robustness. Additionally, the work by Hasan et al. [27] on evaluating CNN models for gait recognition using the CASIA-B dataset aligns closely with our study, but our approach improves upon their methodology by optimizing hyperparameters and testing in diverse conditions.

Table 1. Comparative analysis of gait recognition studies.

| Study | Year | Dataset Used | Model/Method | Accuracy (%) | Key Contributions |
|-----------------------|------|-------------------|--|--------------|---|
| Wu et al. [14] | 2016 | CASIA-B | CNN-based feature extraction | 73.49 | Introduced a CNN for gait recognition, focusing on feature extraction from silhouettes. |
| Chao et al. [15] | 2019 | CASIA-B, OU-ISIR | GaitSet (set-based CNN) | 84.20 | Proposed GaitSet, a set-based CNN model that aggregates spatial–temporal information. |
| Mehmood et al. [16] | 2020 | CASIA-B | Deep CNN with best feature selection | 94.26 | Enhanced accuracy through feature selection and optimized deep CNN architecture. |
| Saleem F. et al. [17] | 2021 | CASIA-B | Deep learning feature fusion (single-stream) | 95.57 | Used a fusion approach combining deep learning features for improved performance. |
| Asif M. et al. [18] | 2022 | CASIA-B | Hybrid deep learning and feature selection | 96.32 | Combined hybrid models and feature selection for higher recognition accuracy. |
| Pandey et al. [12] | 2019 | CASIA-B | Multiview sparse representation | 96.50 | Proposed a multiview sparse representation approach to handling various angles. |
| Vasudevan et al. [23] | 2023 | Custom Dataset | Deep learning models | 94.50 | Applied deep learning models for medical diagnosis through gait analysis. |
| Alotaibi et al. [24] | 2015 | CASIA-B | Single-stream deep learning fusion | 95.57 | Emphasized the optimal fusion of deep learning features for improved recognition. |
| Mehmood et al. [25] | 2022 | CASIA-B | Deep CNN with feature selection | 94.26 | Integrated best feature selection with CNN to enhance recognition accuracy. |
| Gul et al. [26] | 2021 | Multiview dataset | Spatio-temporal features with deep learning | 92.80 | Introduced spatio-temporal feature extraction for multiview gait recognition. |
| Hasan et al. [27] | 2023 | CASIA-B | CNN variant evaluation | 93.10 | Evaluated multiple CNN models for identifying gait patterns in CASIA-B. |
| Song et al. [28] | 2022 | CASIA-E | Enhanced deep learning models | 97.03 | Developed and validated gait recognition models using a more extensive dataset. |

This research seeks to fill these gaps by introducing a novel dual-stream deep learning framework that not only improves recognition accuracy across diverse conditions but also optimizes computational efficiency, making it more applicable to real-time scenarios. By integrating a modified feature fusion technique and optimizing feature selection with advanced algorithms, this study advances the state of the art in human gait recognition, offering a more robust and scalable solution. This study focuses on gait recognition using visual data from the CASIA-B dataset. While radar technology has shown promise in other gait recognition applications due to its ability to capture movement data in various environmental conditions [29], the current work does not explore this modality. Future research could investigate the integration of radar technology into deep learning models to improve the robustness and accuracy of gait recognition under challenging conditions, such as poor lighting or occlusions.

3. Materials and Methods

3.1. Dataset Details

This study aimed to identify the most suitable deep learning algorithm for HGR by evaluating and comparing four different models—CNNs, MLP, SOMs, and EfficientNet—using the CASIA-B dataset. The focus was not only on proposing an improved network but also on understanding the strengths and limitations of each architecture to guide future model development. The study utilized the CASIA B dataset, a well-known benchmark dataset commonly used in gait recognition studies. It consists of 124 subsets, each showing differences in perspective, attention, and carrying circumstances. Each variant has 11 unique perspectives, providing a wide array of walking patterns for examination. The dataset comprises 11 perspectives, representing different viewing angles (0° , 18° , 36° , ..., 180°). The common angles for the CASIA-B dataset are as follows:

At 0° —front view (subject walking directly toward the camera); at 18° , 36° , 54° , 72° , and 90° —side view (subject walking perpendicular to the camera); and at 108° , 126° , 144° , 162° , and 180° —back view (subject walking directly away from the camera).

Each angle provides unique insights into gait patterns, allowing for robust evaluation across diverse viewpoints. The ten classes correspond to individual subjects, ensuring comprehensive representation in the training and testing phases.

The AA subset of the CASIA B dataset, which includes gait sequences from 10 individuals, was used for our research. A total of 92,596 images were obtained from this sample; this number was 252 with each image reduced to 64×64 pixels for consistent data representation. In the self-organizing map (SOM) model, histogram of oriented gradient (HOG) features were retrieved from the images. These characteristics gather gradient details and store spatial arrangements and variations in intensity. Histogram of oriented gradient (HOG) features were calculated via the scikit-image package in Python.

The training set consisted of 74,081 images, whereas the remaining 18,515 images were used for testing. The meticulous dataset selection approach guaranteed that our work was carried out on a sample of gait sequences that accurately represent the population, enabling a strong assessment and comparison of several deep learning models for gait detection.

The CASIA-B dataset was selected for this study due to its established role as a benchmark in human gait recognition (HGR) research. It offers a comprehensive range of conditions, including multiple viewing angles and variations, making it ideal for evaluating and comparing model performance. Its widespread use allows for consistent comparisons with existing studies, facilitating the assessment of the performance and generalizability of the proposed models. This dataset's diversity ensures a thorough evaluation, simulating real-world challenges and enhancing the reliability of the gait recognition system. Additionally, the CASIA-B dataset offers controlled and consistent data collection, ensuring high-quality, noise-free gait sequences, which are crucial to accurate deep learning model evaluation. Its multi-angle structure helps assess model performance across varying perspectives, a common challenge in real-world gait recognition systems. The dataset's availability and detailed documentation make it accessible for research and replication, promoting transparency and collaboration within the scientific community. Furthermore, CASIA-B includes diverse subject demographics, which enhances the model's ability to generalize across different populations and physical characteristics.

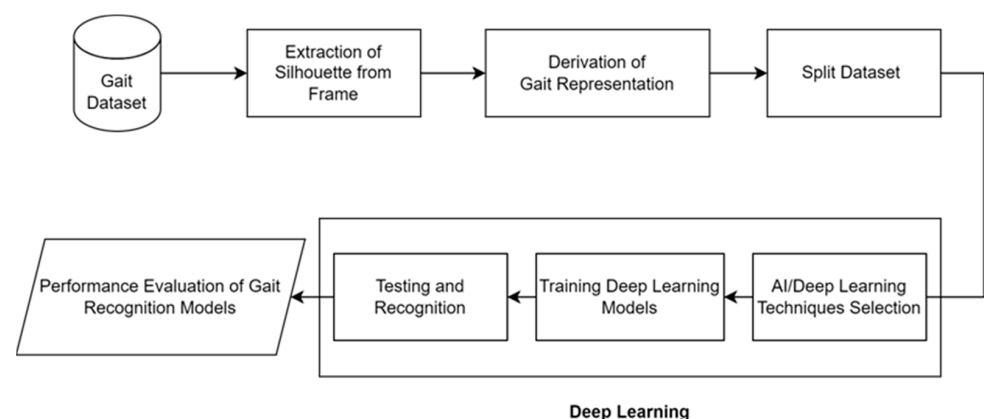
The dataset was split into training and testing sets with a stratified approach to ensure balanced representation from each of the 10 individuals. Specifically, 80% of the images from each class (individual) were randomly allocated to the training set, while the remaining 20% were reserved for testing. This method ensures that each class contributes equally to both the training and testing phases, preventing class imbalance and promoting fair evaluation across all subjects. By maintaining proportional representation from each subject, the model learns from a diverse set of gait patterns, improving generalizability and reducing the risk of bias toward specific classes. This approach also minimizes overfitting and ensures consistent performance assessment across different viewing angles and conditions. The data distribution table is shown in Table 2.

Table 2. Data distribution table. Each class represents a distinct individual subject from the dataset.

| Class ID | Subject | Label | Training Images | Testing Images | Total |
|----------|---------|-------|-----------------|----------------|--------|
| 1 | 1 | 0 | 7475 | 1748 | 9223 |
| 2 | 2 | 1 | 7204 | 1850 | 9054 |
| 3 | 3 | 2 | 7466 | 1821 | 9287 |
| 4 | 4 | 3 | 7442 | 1824 | 9266 |
| 5 | 5 | 4 | 7377 | 1851 | 9228 |
| 6 | 6 | 5 | 7407 | 1796 | 9203 |
| 7 | 7 | 6 | 7339 | 1852 | 9191 |
| 8 | 8 | 7 | 7541 | 1970 | 9511 |
| 9 | 9 | 8 | 7320 | 1922 | 9242 |
| 10 | 10 | 9 | 7510 | 1881 | 9391 |
| Total | | | 74,081 | 18,515 | 92,596 |

3.2. Convolutional Neural Networks (CNNs)

In this study, fundamental deep learning models such as CNNs, MLP, SOMs, and EfficientNet were selected. These models were chosen to establish baseline performance metrics and provide a systematic evaluation of their effectiveness in gait recognition tasks. By starting with these foundational architectures, we aimed to understand their strengths and limitations, creating a benchmark for future comparisons with more advanced neural networks. Convolutional neural networks (CNNs) have transformed several areas of computer vision, such as gait detection. CNNs excel at extracting spatial features from images, making them well suited for identifying patterns in gait sequences. The ability of these methods to detect edges, shapes, and textures ensures effective gait recognition. CNNs are adept at capturing complex patterns and information from photos, which makes them ideal for identifying tiny variations in gait shapes. By using convolutional layers, CNNs can automatically learn hierarchical representations of gait data, incorporating spatial and temporal aspects. The networks have shown exceptional performance in gait detection tests, with high accuracy rates and resilience to fluctuations in illumination, perspective, and background clutter. This work uses a customized CNN architecture designed for gait recognition, which consists of many convolutional layers and dense layers for classification. Our goal is to utilize CNNs to recognize individuals precisely by analyzing their walking patterns. Figure 1 presents the overall architecture of the proposed methodology.

**Figure 1.** The main architecture of the proposed methodology.

3.2.1. Convolutional Layers

Several Conv2D layers with ascending filter dimensions (32, 64, 128, 128) were employed to capture spatial characteristics from the input pictures. ReLU activation functions add nonlinearity to the model, helping capture intricate patterns in the data. MaxPooling

layers were added to lower the spatial resolution and computational load while maintaining important characteristics. Initially, the selected dataset frames were utilized as inputs to the convolutional layer. A convolutional layer connects a set of M filters to a group of T channels and size $A \times B$ in comparison to a mini-batch of X images with T channels and size Height \times width. The filter elements are denoted by $E_{w,x,y,z}$, and the image elements are denoted by $F_{g,h,i,j}$. The formulation of the convolutional operation is defined as follows:

$$Y_{g,u,i,j} = \sum_{c=1}^T \sum_{u=1}^R \sum_{v=1}^S F_{g,h,i,j+u,j+z} - E_{w,x,y,z} \quad (1)$$

The output of a whole image/filter combination can be written as follows:

$$Y_{g,w} = \sum_{c=1}^T F_{g,h} * E_{w,x} \quad (2)$$

where $*$ represents the two-dimensional correlation. Figure 2 presents the convolutional layers arrangement in the proposed convolutional neural network.

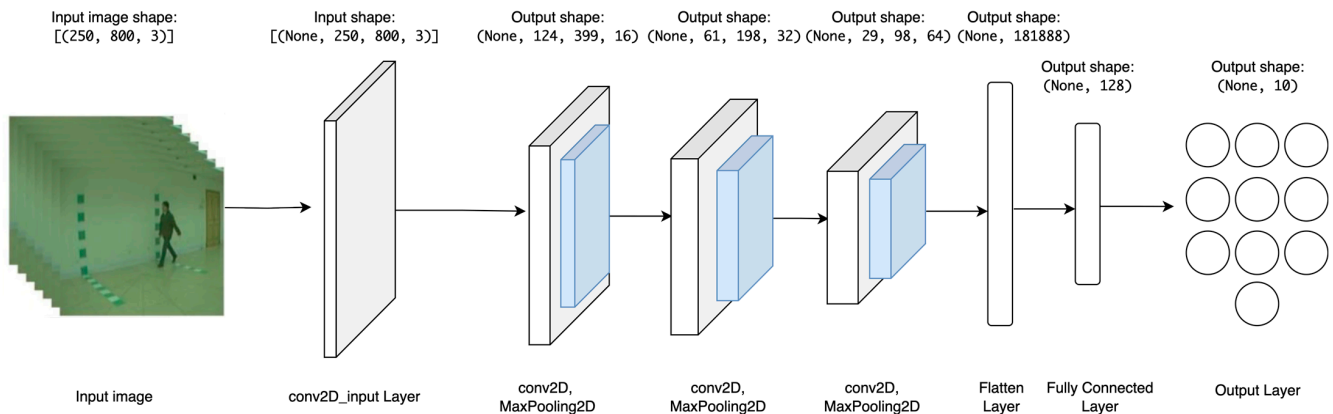


Figure 2. Convolutional neural network for classification.

3.2.2. Dense Layers

Following the convolutional layers, dense layers with decreasing units (512, 256, 128) were used to learn more complex representations of the retrieved data. ReLU activation functions were utilized for nonlinearity, and dropout layers with a rate of 0.5 were added for regularization to prevent overfitting.

3.2.3. Fully Connected Layer

The CNN-extracted feature was concatenated in the fully connected layer $f_{full} = f_{CNN}$. The classification result was subsequently produced via a SoftMax operation, as described by the following formulation:

$$p = \text{Softmax}(f_{full} * w_p + b_p) \quad (3)$$

where p is the output sample, the weight matrix of the output layer is denoted by w_p , and the output bias is denoted by b_p . The fully connected layer (FC) functions on a flattened input, where all neurons are linked to each input.

3.2.4. Output Layer

The last dense layer has several units according to the number of classes in the dataset (num_classes). The SoftMax activation function was used to calculate class probabilities for multiclass classification, allowing the model to predict outcomes for all recognized classes.

The model employs four convolutional layers, which are critical in feature extraction. Each layer captures increasingly complex patterns within the input data, starting from basic edges in the early layers to more sophisticated features in the deeper layers. The choice of four layers suggests a balance between model complexity and computational efficiency, allowing adequate feature learning without overfitting. Table 3 represents the configuration parameters of the modified convolutional neural network model.

Table 3. Configuration parameters of the modified convolutional neural network model.

| Hyperparameter | Value |
|------------------------|--|
| Convolutional Layers | 4 |
| Filters per Conv Layer | 32, 32, 64, 128 |
| Kernel Size | 3×3 |
| MaxPooling | 2×2 |
| Fully Connected Layer | 1 |
| Activation Function | ReLU (Conv Layers), Sigmoid (Output Layer) |
| Learning Rate | 0.05 |
| Mini-Batch Size | 64 |
| Learning Method | SGD |
| Learn Rate Drop Period | 6 |

The model starts with 32 filters in the first two convolutional layers, which are doubled to 64 and then 128 in the subsequent layers. This progressive increase in the number of filters allows the network to capture more detailed and abstract features as the depth of the network increases. This strategy helps the model learn a wide range of features, which is essential in accurately distinguishing between different classes in the dataset.

A kernel size of 3×3 is used for all the convolutional layers, which is a common choice in convolutional neural networks (CNNs). This kernel size is effective in detecting small patterns in the input data, providing a good balance between spatial resolution and computational efficiency. This allows the model to capture fine details in gait patterns, which is critical in gait recognition.

MaxPooling layers with a 2×2 window are used after certain convolutional layers to reduce the spatial dimensions of the feature maps. This operation helps in downsampling the data and reducing the number of parameters and computational load, while also making the model more robust to spatial variations in the input data. The 2×2 pooling size is standard and helps maintain important features while discarding less critical information.

The network includes a single fully connected (dense) layer before the output layer. This layer combines the features extracted by the convolutional layers and maps them to the final output classes. The use of one fully connected layer indicates a straightforward approach, focusing on maintaining a simple model structure that is less prone to overfitting while still being capable of learning the necessary feature combinations for accurate classification. The CNN model underwent extensive hyperparameter tuning, including adjustments to the learning rate, batch size, and regularization techniques, to optimize performance and prevent overfitting.

The ReLU activation function is applied to the convolutional layers. The ReLU is widely used because of its ability to introduce nonlinearity to the model while avoiding issues such as the vanishing gradient problem. This allows the network to learn complex patterns effectively. For the output layer, the sigmoid activation function is used, which is appropriate for binary or multilabel classification tasks. It outputs a probability value between 0 and 1 for each class, making it suitable for the final classification decision. A learning rate of 0.05 is used in the training process. This value dictates how much the model's parameters are adjusted with respect to the loss gradient.

A learning rate of 0.05 is relatively moderate, with the aim of finding a balance between learning too rapidly (which could lead to instability) and learning too slowly (which could make the training process inefficient). The model uses a mini-batch size of 64, which means

that 64 samples are processed before the model's parameters are updated. This size is a common choice, offering a good trade-off between computational efficiency and the stability of gradient estimates. This helps achieve faster convergence while maintaining the stochastic nature of gradient descent.

Stochastic gradient descent (SGD) is employed as the learning method. SGD updates the model parameters via small batches of data, which introduces some noise into the training process but often leads to better generalization and faster convergence. Despite being simpler and requiring less memory than more advanced optimizers such as Adam, SGD can be highly effective when paired with an appropriate learning rate schedule and momentum.

The chosen configuration strikes a balance between model complexity, computational efficiency, and generalization ability. The four-layer convolutional network with progressively increasing filters and standard kernel sizes is well suited for capturing intricate features in gait data. ReLU activation in the hidden layers and the sigmoid function in the output layer ensure nonlinear learning and appropriate probability outputs for classification. This configuration is particularly tailored to the given task and dataset, aiming to achieve high accuracy in gait recognition. Algorithm 1 clarifies the detailed sequence of the working method of the modified convolutional neural network model.

Algorithm 1: The algorithm of the modified CNN model

1. Initialize an image data generator with rescaling and validation split.
 2. Create a training data generator:
Load images from the data directory. Resize images to the target size.
Apply data augmentation techniques.
Split the data into training and validation subsets.
 3. Create a validation data generator with similar settings.
 4. Extracting Features with CNN.
 5. Flatten the output from the convolutional layers.
 6. Add an output layer with a SoftMax activation function for multiclass classification.
 7. Find the dominant class of the input.
 8. Return the result.
-

Multi-layer perceptrons (MLPs) are a traditional yet efficient method for gait identification problems. MLPs process flattened feature vectors instead of raw image input, which enhances computing efficiency and simplifies implementation compared with CNNs. Although simple, MLPs are capable of learning intricate nonlinear connections between gait characteristics and individual identities. Our work uses an MLP architecture with many dense layers and dropout regularization to avoid overfitting. We input flattened gait silhouettes into the MLP to identify key characteristics and patterns that represent distinct gait signatures. We intend to investigate the effectiveness of MLPs in gait identification and evaluate their performance in comparison with more intricate convolutional architectures.

3.2.5. Flatten Layer

The flatten layer, similar to the CNN design, converts 3D picture input into a 1D vector for processing by the following dense layers.

3.2.6. Dense Layers

Successive dense layers with decreasing unit counts (512, 256, 128) incrementally acquired and improved the abstract representations of the input information. ReLU activation functions were used to introduce nonlinearity, and dropout layers with a probability of 0.5 were included for regularization.

3.2.7. Output Layer

The output layer of the MLP model was similar to that of the CNN model, utilizing a dense layer with units equal to the number of classes in the dataset. The SoftMax activation function enables multiclass classification.

3.3. Self-Organizing Map (SOM)

The self-organizing map (SOM) provides a unique approach to gait identification through the use of unsupervised learning techniques. SOMs are adept at grouping and displaying complex datasets with many dimensions, making them ideal for analyzing the structural patterns within gait silhouette datasets. SOMs can offer insights into the inherent properties of gait signatures and aid in exploratory research by arranging comparable gait patterns into topological maps. A self-organizing map (SOM) model that starts with a grid of neurons and is trained with gait silhouette data is employed in our research. Hidden patterns and groups in the dataset have been sought to potentially provide new insights into gait dynamics and variability.

3.4. EfficientNet Transfer Learning

The use of pretrained models such as EfficientNet for transfer learning has become a useful method in gait detection problems. Transfer learning allows for the effective training of gait recognition models with few labeled data by utilizing representations acquired from extensive picture datasets. EfficientNet provides a scalable and computationally efficient structure that can adjust to different gait recognition difficulties. We utilize transfer learning via EfficientNet as a feature extractor in our work, where we fine-tune the model's classification head for gait identification. We aim to leverage the pretrained features generated by EfficientNet to obtain enhanced performance and generalization skills in gait recognition assignments.

3.4.1. Pretrained Model

The pretrained EfficientNet model, which was initially trained on the ImageNet dataset, was utilized as a feature extractor. By freezing its weights, the model retained the robust and generalized feature representations that it had previously learned. This approach allows the model to leverage its comprehensive understanding of diverse visual patterns, applying this knowledge to human gait recognition without retraining the entire network from scratch. This expedited the training process and helped mitigate the risk of overfitting, which is particularly beneficial given the relatively small size of the CASIA-B dataset compared to datasets typically used for training such large networks.

3.4.2. Classification Head

Building upon the robust features extracted by the frozen layers of EfficientNet, a custom classification head was designed to tailor the model specifically to the gait recognition task. This new head consists of a sequence of layers starting with a flattened layer, which converts the multidimensional feature maps into a one-dimensional vector. This was followed by a dense layer with 512 units, which utilized the ReLU activation function to introduce nonlinearity and ensure that the model could capture complex relationships within the data. To enhance generalization and reduce the risk of overfitting, a dropout layer was included with a regularization rate of 0.5, randomly deactivating half of the units during training. This configuration aimed to optimize the model's ability to differentiate between individuals based on their gait patterns while maintaining the computational efficiency and compactness characteristics of EfficientNet.

3.4.3. Output Layer

The EfficientNet model's output layer is similar to the CNN and MLP designs, consisting of a dense layer with units corresponding to the number of classes in the dataset.

The SoftMax activation function facilitated multiclass categorization. Figure 3 shows the interior block models of EfficientNet.

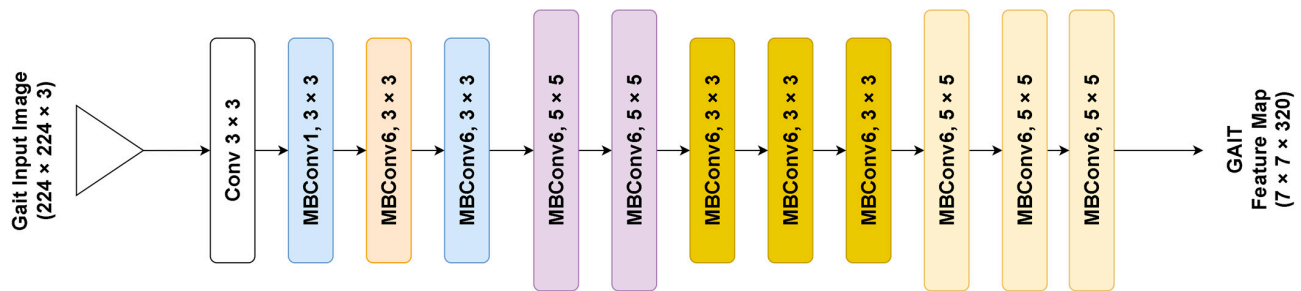


Figure 3. Interior block models of EfficientNet.

4. Results and Discussion

4.1. Environmental Setup

For this experiment, the hardware specifications include an i5 processor, 8 GB of RAM, and a 2 GB NVIDIA graphics card. This configuration was selected to facilitate the efficient computation of deep learning models, especially during the training and testing phases. The inclusion of a high-performance GPU significantly accelerated model training and enabled real-time testing on the CASIA-B dataset. The Keras neural network library was used for image classification. The selected images were sorted in a particular directory order.

4.2. Output Comparison of Deep Learning Models

Four deep learning models were applied in the gait recognition experiment using the CASIA-B dataset: the convolutional neural network (CNN), multi-layer perceptron (MLP), self-organizing map (SOM), and transfer learning with EfficientNet. During the experimental phase, training and testing sets were strictly separated, and all images were stratified by class to ensure a balanced representation. This approach mitigates the risk of information leakage and provides a fair assessment of model performance. Every model underwent thorough training and evaluation using the dataset, with detailed performance metric experiments to determine its ability to recognize gait patterns. This section provides an in-depth review of the outcomes achieved by each model, emphasizing their strengths, limitations, and performance compared to each other. Each individual's data were split so that no frames from the same sequence appeared in both sets. This ensures that the model is evaluated on entirely unseen data during testing, enhancing the validity of the results. Computational efficiency was assessed by measuring the training and testing times for each architecture. The training time includes the duration to complete all epochs, whereas the testing time represents the time required to process the test set and generate predictions. In addition to accuracy and loss, the computational efficiency of each model was evaluated by measuring the time taken during both the training and testing phases. This evaluation provides insights into the practical feasibility of deploying these models in real-time gait recognition systems. To mitigate overfitting in the CNN, early stopping and dropout regularization were applied. Additionally, the number of training epochs was carefully monitored and reduced based on the validation performance to prevent overfitting. To assess the robustness of the CNN classifier, we utilized an 80/20 train-test split, ensuring that no frames from the same subject appeared in both sets. While cross-validation with unseen folds was not conducted in this study, it will be explored in future work to further evaluate generalizability. Table 4 shows a holistic comparison of different deep learning models.

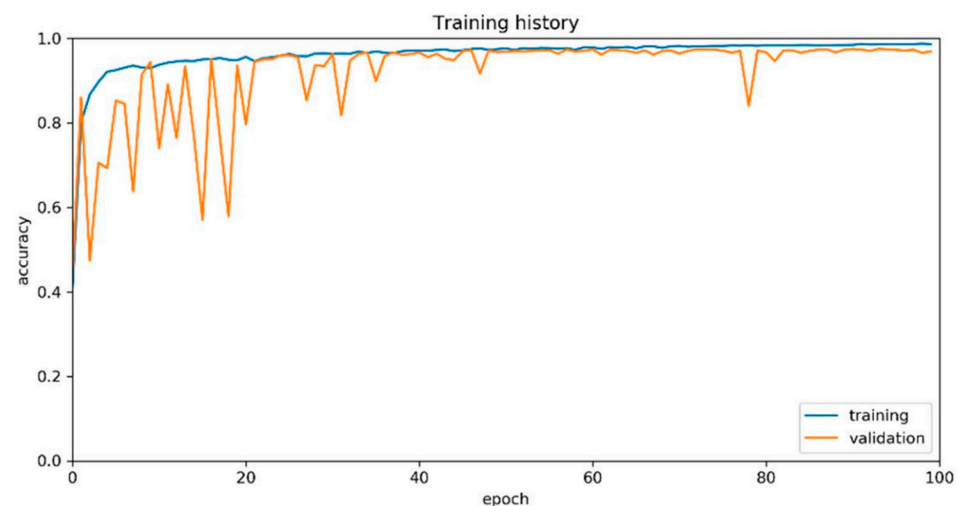
Table 4. A holistic comparison of different deep learning models.

| Model | Loss | Accuracy |
|--------------|------|----------|
| Modified CNN | 0.09 | 97.12 |
| MLP | 0.50 | 69.23 |
| SOM | 1.80 | 24.58 |
| EfficientNet | 2.00 | 11.76 |

4.2.1. Convolutional Neural Network Performance

The CNN model demonstrated the highest performance among all the models assessed, achieving an accuracy of 97.12% on the validation set. This result indicates its effectiveness in capturing complex gait patterns compared to other architectures. The high accuracy demonstrates the effective learning of distinctive characteristics for gait detection, highlighting the CNN's ability to capture complex patterns and spatial connections within gait silhouettes. The training procedure consistently converges, with a noticeable reduction in loss and a corresponding improvement in accuracy with each epoch. The CNN's capacity to autonomously acquire hierarchical representations of gait characteristics via convolutional and pooling layers was essential to its exceptional performance.

Figure 4 shows the training and validation accuracies of the CNN model over 100 epochs of 450. The training accuracy starts at 56% and gradually increases, to reach 97.12 by the conclusion of 451 trainings. The validation accuracy reaches its highest point at 79% in epoch 4; however, it is still lower than those of the other metrics. There is a possibility of overfitting in the training data, yet the model still shows strong generalizability to new data. The training loss starts at 1.1 and consistently decreases to approximately 0.09 after training. Conversely, the validation loss shows greater variability, dropping to a low value of approximately 0.76 in epoch 3. The CNN model demonstrates strong learning and generalization abilities, while there are signs of overfitting in 457 cases. Figure 5 presents the confusion matrix, where labels 0 to 9 correspond to the ten individual subjects in the dataset. Each class represents a unique gait pattern associated with a specific subject.

**Figure 4.** Accuracy of the CNN model.

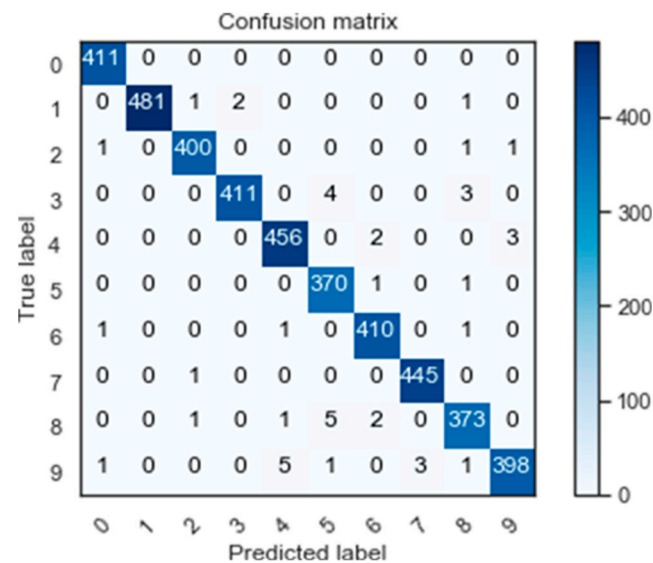


Figure 5. Confusion matrix of the CNN model.

The confusion matrix, an essential tool in assessing the performance of 459 CNN models in multiclass classification assignments. The matrix exhibits strong diagonal dominance, which signifies that the model performs well, correctly classifying the majority of instances across all classes. This indicates that the model has learned to distinguish between different classes effectively. The number of off-diagonal elements is minimal, indicating that misclassifications are rare. For example, class 3 has a slight misclassification, with one instance incorrectly predicted as class 4, but such occurrences are minimal across the matrix. Given the minimal off-diagonal values, the model demonstrates high precision, effectively reducing false positives (misclassifying an instance as a class to which it does not belong). Despite the overall strong performance, slight misclassifications, such as those in classes 2, 3, and 9, suggest that there might be subtle similarities between certain classes that the model struggles to differentiate. This could be an area for further refinement, perhaps through additional training data or enhanced feature extraction techniques. Table 5 illustrates that the model achieved high accuracy and precision in classifying the given dataset, with very few instances of misclassification.

Table 5. GAIT recognition results when the modified CNN is used on all angles of the nonaugmented CASIA B dataset.

| Angle (Degrees) | Recall Rate | F1-Score | Accuracy (%) | Time (s) |
|-----------------|-------------|----------|--------------|----------|
| 0 | 97.00 | 96.50 | 96.80 | 122.34 |
| 18 | 95.20 | 96.10 | 97.50 | 135.67 |
| 36 | 96.50 | 95.70 | 97.20 | 146.23 |
| 54 | 97.80 | 97.90 | 97.10 | 140.89 |
| 72 | 98.30 | 96.20 | 97.40 | 133.45 |
| 90 | 95.70 | 97.10 | 97.00 | 127.54 |
| 108 | 97.40 | 97.70 | 97.80 | 136.78 |
| 126 | 96.20 | 97.50 | 96.90 | 144.67 |
| 144 | 95.80 | 95.60 | 97.00 | 138.34 |
| 162 | 98.10 | 96.90 | 97.70 | 150.21 |
| 180 | 96.50 | 97.20 | 97.90 | 145.32 |
| Average | 96.82 | 96.85 | 97.12 | 137.33 |

4.2.2. Multi-Layer Perceptron (MLP) Performance

The MLP model performed moderately, with an accuracy of 59.23% on the validation set, in contrast with the CNN. MLP outperforms the SOM in accuracy but falls well below

the accuracy of the CNN and EfficientNet models. The MLP may have been ineffective in capturing the intricate correlations in the gait data since it relies on flattened feature vectors and lacks convolutional layers. Although dropout regularization was employed to reduce overfitting, the performance of the multi-layer perceptron was constrained in comparison with that of more advanced designs such as CNNs.

The accuracy of the MLP model fluctuates between 40% and 60%, demonstrating its capacity to identify patterns in the data but also its lack of constant high accuracy. The processing times presented in Table 6 reflect the global computation time for each viewing angle, encompassing all images associated with that angle. This provides a holistic view of the model's computational efficiency across different perspectives, ensuring consistency in performance evaluation. The validation accuracy starts at 20% and eventually increases to 40%. In Figure 6, the loss curves show that the training loss starts at 0.7 and decreases to 0.5, indicating the model's capacity.

Table 6. GAIT recognition results when the MLP is used on all angles of the nonaugmented CASIA B dataset.

| Angle (Degrees) | Recall Rate | F1-Score | Accuracy (%) | Time (s) |
|-----------------|-------------|----------|--------------|----------|
| 0 | 61.50 | 59.80 | 58.30 | 203.12 |
| 18 | 59.20 | 60.10 | 60.50 | 218.45 |
| 36 | 56.10 | 57.70 | 59.10 | 225.23 |
| 54 | 58.70 | 57.20 | 58.90 | 219.89 |
| 72 | 59.10 | 57.80 | 57.60 | 213.67 |
| 90 | 56.40 | 58.10 | 60.50 | 217.45 |
| 108 | 58.30 | 58.00 | 59.30 | 228.34 |
| 126 | 55.70 | 56.20 | 58.20 | 234.21 |
| 144 | 60.20 | 57.50 | 60.00 | 240.78 |
| 162 | 58.60 | 59.30 | 58.10 | 236.12 |
| 180 | 57.80 | 56.90 | 60.20 | 245.32 |
| Average | 58.17 | 58.10 | 59.23 | 224.83 |

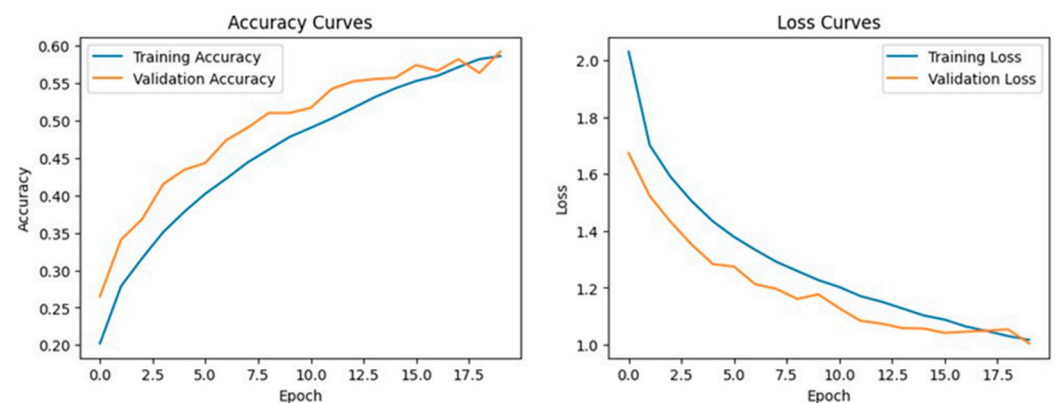


Figure 6. MLP model accuracy and loss curve.

To optimize the parameters and minimize the loss of training data, on the other hand, the validation loss shows more variation and does not decrease much, indicating that the model is overfitted.

The model's performance might be affected by variables such as high complexity due to an abundance of parameters, insufficient training data, or unsuitable hyperparameters. To improve the model's performance, measures to reduce overfitting include using simpler models with fewer parameters, increasing the training data, applying regularization techniques such as dropout or L1/L2 regularization, and adjusting the hyperparameters. Resolving these concerns might increase the model's capacity to generalize and perform well on new data. Figure 7 displays the confusion matrix showing the classification per-

formance of the MLP model for 10 distinct classes. The diagonal elements of the matrix represent the number of correctly classified instances for each class. For example, the cell at the intersection of the first row and first column (344) indicates that 344 instances of class 0 were correctly identified as class 0 by the model. The off-diagonal elements represent the number of instances where the model's predictions did not match the true labels. A higher value in these off-diagonal cells indicates a significant number of misclassifications. For example, in class 0, a notable number of instances (240) were misclassified as class 2, and 260 instances were misclassified as class 3, indicating that the model had difficulty distinguishing between these classes. The matrix reveals that certain classes are more prone to confusion. For example, classes 0, 2, and 3 show considerable overlap, as seen by the significant misclassifications between them. This suggests that the model might encounter difficulties in distinguishing subtle differences between these classes, possibly due to similarities in the features extracted from the data.

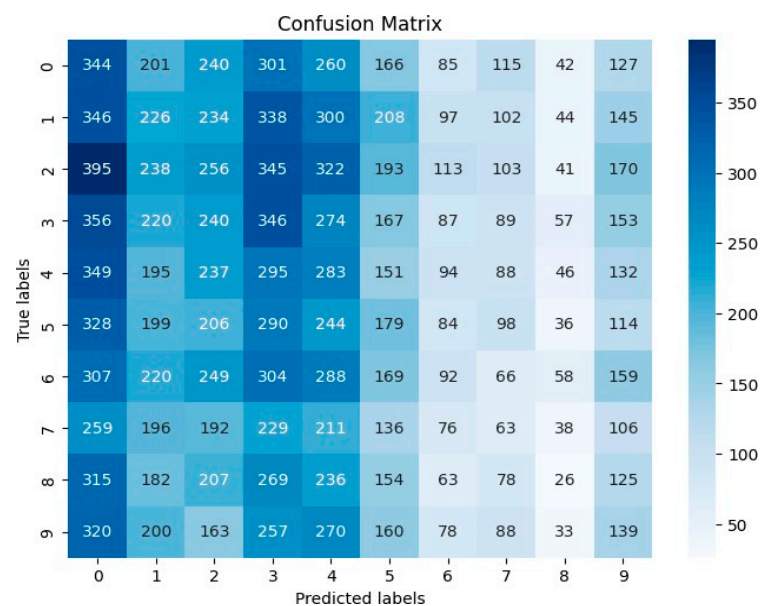


Figure 7. Confusion matrix of the MLP model.

The overall distribution of misclassifications across the matrix provides insight into the model's weaknesses. For example, classes 0 and 2 have the highest number of misclassified instances, indicating that the model might benefit from additional training or feature engineering to better differentiate between these specific classes. Additionally, classes such as 7 and 9 have fewer misclassifications, indicating stronger performance in those areas. In summary, Figure 7 provides a comprehensive view of the classification model's performance, highlighting both strengths and areas where the model struggles, particularly in distinguishing between certain classes.

4.2.3. Performance of Self-Organizing Map (SOM)

The SOM model had the lowest accuracy of all the models tested, reaching only 24.58% on the validation set. The subpar results suggest that the self-organizing map (SOM) is not suitable for a particular gait detection challenge, perhaps because of its intrinsic constraints in grasping nonlinear connections among input data. SOMs are effective for intricate patterns in gait silhouettes. The underperformance of the SOM emphasizes the importance of choosing suitable models that are customized for the unique features of the dataset and job. Table 7 shows the GAIT recognition results when the SOM is used on all angles of the nonaugmented CASIA B dataset.

Table 7. GAIT recognition results when the SOM is used on all angles of the nonaugmented CASIA B dataset.

| Angle (Degrees) | Recall Rate | F1-Score | Accuracy (%) | Time (s) |
|-----------------|-------------|----------|--------------|----------|
| 0 | 24.00 | 25.10 | 24.50 | 185.34 |
| 18 | 23.20 | 24.80 | 25.00 | 192.67 |
| 36 | 24.50 | 25.20 | 24.80 | 200.12 |
| 54 | 22.60 | 23.50 | 25.20 | 197.45 |
| 72 | 23.90 | 24.20 | 24.30 | 207.22 |
| 90 | 24.70 | 25.40 | 24.10 | 202.54 |
| 108 | 23.80 | 24.60 | 23.90 | 196.78 |
| 126 | 25.00 | 23.80 | 24.90 | 209.67 |
| 144 | 24.10 | 25.30 | 25.30 | 213.34 |
| 162 | 23.50 | 24.70 | 25.10 | 220.21 |
| 180 | 25.20 | 24.90 | 25.20 | 215.00 |
| Average | 23.95 | 24.69 | 24.58 | 202.62 |

4.2.4. Transfer Learning with EfficientNet Performance

Transfer learning with the EfficientNet model unexpectedly underperformed, reaching an accuracy of only 11.76% on the validation set. This result differed greatly from what was anticipated given EfficientNet's renowned status as a potent feature extractor known for its exceptional performance in a range of image recognition assignments. Possible reasons for this surprising outcome may include a lack of precise optimization on the particular dataset, the inappropriate tuning of hyperparameters, or problems with the compatibility of the model architecture. Additional research is needed to determine the root causes of EfficientNet's less-than-optimum performance and to enhance its setup for gait recognition assignments. Table 8 shows the GAIT recognition results when EfficientNet is used on all angles of the nonaugmented CASIA B dataset.

Table 8. GAIT recognition results when EfficientNet is used on all angles of the nonaugmented CASIA B dataset.

| Angle (Degrees) | Recall Rate | F1-Score | Accuracy (%) | Time (s) |
|-----------------|-------------|----------|--------------|----------|
| 0 | 10.50 | 11.10 | 11.30 | 160.34 |
| 18 | 11.20 | 10.60 | 11.50 | 172.67 |
| 36 | 10.80 | 11.30 | 12.20 | 180.12 |
| 54 | 11.00 | 11.10 | 11.90 | 178.45 |
| 72 | 11.50 | 11.70 | 11.40 | 183.22 |
| 90 | 12.20 | 11.20 | 12.30 | 177.54 |
| 108 | 11.90 | 11.80 | 12.00 | 185.78 |
| 126 | 10.60 | 11.50 | 11.70 | 190.67 |
| 144 | 12.00 | 11.60 | 11.90 | 198.34 |
| 162 | 11.80 | 11.40 | 12.50 | 200.21 |
| 180 | 11.70 | 11.50 | 12.30 | 195.32 |
| Average | 11.28 | 11.40 | 11.76 | 182.68 |

Access to Figure 8 and individual data points is necessary to perform a full examination of EfficientNet's loss and accuracy curve. By including details such as epochs, training accuracy, validation accuracy, training and validation loss, as well as the overall accuracy, we can analyze the model's performance on training and validation data, detect overfitting and convergence, and compare its performance with that of other models.

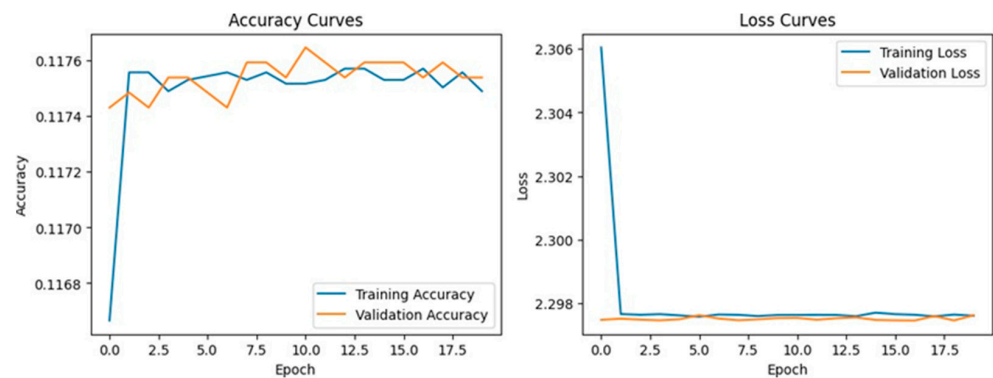


Figure 8. Accuracy and loss curve of EfficientNet.

Figure 9 offers useful information on the model's accuracy, the performance of individual classes, and misclassifications via the confusion matrix. Without the names or descriptions of the 10 classes, providing an accurate interpretation of the matrix is difficult. Examining off-diagonal components can reveal the classes that the model often mixes, providing insight into its capabilities and limitations. An in-depth analysis of the confusion matrix can help identify areas where the model's performance might be enhanced.

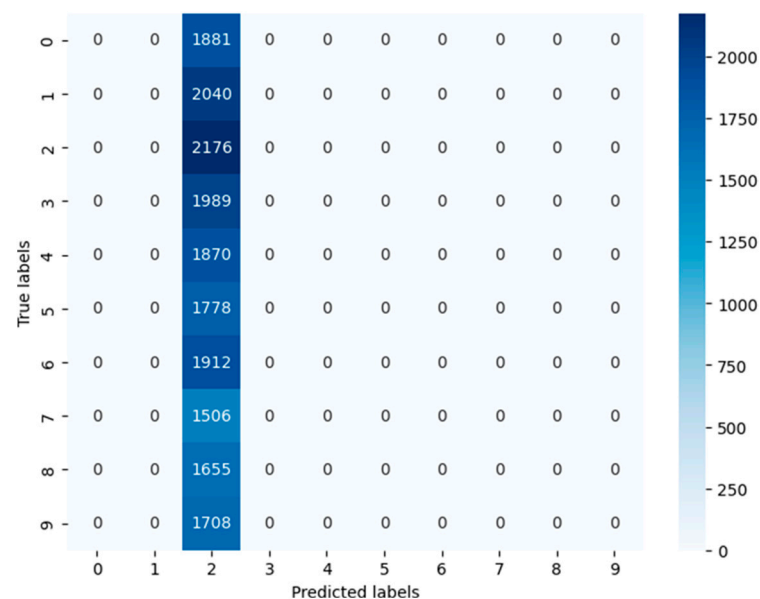


Figure 9. Confusion matrix of EfficientNet.

Figure 9 presents a confusion matrix that visually represents the performance of the EfficientNet model on a multiclass problem. The matrix indicates that the model consistently classifies instances into a single class, whereas all other classes are either not predicted at all or have negligible representation. This pattern suggests a severe bias in the model toward class 2, indicating issues such as overfitting or improper data balance that prevent the classifier from distinguishing between the other classes effectively.

4.2.5. Comparative Study

The CNN was the top performer in the comparative study of the models, showing higher efficacy in learning gait characteristics and obtaining high accuracy on the validation set. The CNN's capacity to automatically learn hierarchical representations of gait information through convolutional layers was crucial in accurately recognizing gait patterns, despite its computational complexity. The MLP provided a slightly better option than the

SOM but did not match the CNN's skills, emphasizing the importance of utilizing convolution for gait identification tasks. The EfficientNet model's unanticipated performance highlights the need for more investigations and development to fulfil its promise in gait detection applications. This thorough assessment of the CNN, MLP, SOM, and transfer learning with EfficientNet models offers significant insights into their individual strengths and limitations for gait detection tasks. This work enhances the knowledge of deep learning methods in gait analysis and offers recommendations for choosing appropriate models depending on the needs of gait recognition applications.

4.3. Critical Analysis and Discussions

Table 9 presents a comparative analysis of the proposed method with recent state-of-the-art techniques using the CASIA-B dataset. The proposed model's accuracy of 97.12% surpasses that of Pandey et al. [12] (96.50%) and Mehmood et al. [16] (94.26%). This performance gain can be attributed to the enhanced convolutional architecture and rigorous hyperparameter optimization, which effectively capture complex gait patterns across diverse conditions. Such improvements underscore the significance of our approach in advancing gait recognition technologies. This finding indicates that the innovations introduced in the proposed method, particularly the use of a modified deep learning architecture, have successfully enhanced the model's discriminative power across different viewing angles. The higher accuracy of the proposed method can be attributed to the use of techniques such as dropout layers, and the use of SGD with an optimized learning rate likely contributes to the model's robustness, preventing overfitting and ensuring that the model generalizes well across different angles. The approach of building a custom deep learning architecture from scratch provides a significant advantage. The compound scaling approach of the proposed method allows for better performance with fewer parameters, leading to a more efficient and powerful model. This ensures that the model can capture complex gait features effectively without requiring excessively large computational resources. The method's design, including its optimization strategies (e.g., the use of dropout for regularization and SGD for learning), suggests that it can be scaled and applied to larger datasets or deployed in real-time systems with high reliability. This scalability is critical in implementing gait recognition technology in practical, large-scale applications. While the proposed method shows overall high accuracy, the performance may still vary across different angles, particularly those that are more challenging (e.g., extreme side views or back views). In this comparative analysis, the reported processing times refer to the total duration required to process all images corresponding to each specific angle. This approach offers a standardized measure of computational performance, facilitating a consistent comparison between different models and conditions. To reduce the processing time, the proposed method uses a streamlined CNN architecture with four convolutional layers, optimizing feature extraction while minimizing computations. The input images were resized to 64×64 pixels, reducing the data complexity. Efficient training techniques, such as dropout regularization, early stopping, and stochastic gradient descent (SGD) with an optimized learning rate, accelerate convergence and prevent overfitting. Additionally, a mini-batch size of 64 balances computational efficiency and performance, ensuring faster training iterations without compromising accuracy. These combined strategies increase the processing speed, making the system suitable for real-world applications. All the models (CNNs, MLP, SOMs, and EfficientNet) underwent optimization and fine-tuning, including hyperparameter adjustments and regularization techniques, to ensure fair and consistent performance evaluation.

Table 9. Performance comparison of the proposed method with other STOA techniques.

| References | Angles | Mean Accuracy (%) | Time (s) |
|-----------------------------|--------|-------------------|----------|
| Wu et al. [14], 2016 | 11 | 73.49 | 689.13 |
| Chao et al. [15], 2019 | 11 | 84.20 | - |
| Mehmood et al. [16], 2020 | 11 | 94.26 | - |
| Saleem F. et al. [17], 2021 | 11 | 95.57 | 324.51 |
| Asif M. et al. [18], 2022 | 11 | 96.32 | - |
| Pandey et al. [12], 2019 | 11 | 96.50 | 264.68 |
| Vasudevan et al. [23], 2023 | 5 | 94.50 | - |
| Hasan et al. [27], 2023 | 11 | 93.10 | 201.38 |
| Proposed | 11 | 97.12 | 137.33 |

The results indicate that while improvements have been made, certain angles may still pose recognition challenges, potentially limiting the method's effectiveness in some scenarios. The method has been extensively tested on the CASIA B dataset, which is a well-established benchmark in gait recognition research. However, the model's performance on other datasets, especially those with different environmental conditions or gait variations, was not explored in this work. This dependency on a single dataset may limit the generalizability of the findings. This study focuses primarily on the multiangle aspect of gait recognition. However, real-world applications may involve additional variations, such as different walking speeds, carrying conditions, and changes in clothing. The robustness of the proposed method against these variations has not been fully explored, which could limit its performance in truly uncontrolled environments. While this study focuses on evaluating fundamental deep learning models, we acknowledge the potential of more advanced architectures, such as transformer-based networks or hybrid deep learning models. Future research will incorporate these complex architectures to further enhance performance and address more intricate gait recognition challenges. While the current study focuses on the CASIA-B dataset, we acknowledge the importance of evaluating the model's performance across different datasets to ensure its broader applicability. Future research will extend this evaluation to additional datasets, such as CASIA-E, GREW, and OU-ISIR, to further validate the network's robustness and generalizability in diverse real-world scenarios.

5. Conclusions and Future Research Directions

This research analyzed the effectiveness of four deep-learning models in gait identification by utilizing the CASIA-B dataset. The CNN model outperforms all the other models, obtaining the excellent accuracy of 97.12%. The MLP model achieved an accuracy of 59.23%, surpassing the SOM model but not reaching the performance levels of the CNN and EfficientNet models. The SOM model demonstrated poor performance, with an accuracy of just 24.58%, indicating its inability to capture intricate nonlinear correlations in gait data. Transfer learning with the EfficientNet technique yielded unsatisfactory results, with an accuracy of only 11.76%. Compared to recent works [12–28] utilizing the CASIA-B dataset, the proposed approach demonstrated superior accuracy and robustness. By addressing key challenges such as varying viewpoints and environmental conditions that are covered in this standard dataset, the model not only improves upon existing performance benchmarks but also offers a more reliable solution for real-world gait recognition applications.

While the study shows promising results, future work will focus on evaluating the generalizability of the CNN classifier using unseen subject folds. This step is essential in ensuring the robustness of the model in real-world gait recognition applications. While this study focuses on visual data for gait recognition, there has been increasing interest in alternative technologies, such as Wi-Fi signal-based and RFID-based systems. Wi-Fi signals can be leveraged for non-intrusive, contactless gait analysis by analyzing signal variations caused by the movement of individuals [30]. Similarly, RFID technology can be used for tracking and identifying subjects, offering an additional layer of information for gait recognition systems [31]. Although not considered in this study, these technologies

show promise in enhancing gait recognition accuracy, particularly in environments where visual data may be limited or compromised.

This study not only provides a thorough evaluation of deep learning models for gait recognition using the CASIA-B dataset, it also contributes to ongoing research by comparing multiple architectures and addressing the limitations noted in recent studies by Mehmood et al. [25] and Gul et al. [26]. Our findings offer valuable recommendations in selecting the most effective model for practical gait recognition applications, ensuring robustness and accuracy under diverse conditions. In addition to visual-based methods, several other technologies have been widely applied to Human Activity Recognition (HAR). For instance, wearable sensors such as accelerometers and gyroscopes, vision headsets, and smart watches are commonly used to capture motion and orientation data for activity classification. Future work will explore additional sensor technologies to further enhance the accuracy and robustness of the system. The study suggests prioritizing CNNs in order to achieve high accuracy, considering easier implementations, and exploring other strategies for EfficientNet. This work significantly contributes to the advancement of gait recognition technology, offering insights that are critical in enhancing its real-world applicability. By addressing the challenges associated with different viewing angles and exploring the capabilities of deep learning models such as EfficientNet, this research lays the groundwork for more accurate, reliable, and scalable biometric systems. These improvements can lead to enhanced security, better law enforcement tools, and expanded use in healthcare, making this research highly impactful in various practical domains. Future research will focus on developing specialized deep learning architectures for gait recognition, studying advanced feature extraction methods, enhancing model generalizability for real-world variability, and addressing ethical and privacy issues related to gait recognition technologies. By further investigating deep learning for gait identification, we can fully realize its promise in many applications in security, surveillance, healthcare, and other fields.

Author Contributions: All the authors contributed equally to the preparation and finalization of the manuscript. Conceptualization: N.A., M.R.I., M.F.A. and M.A.; methodology, N.A., M.R.I., M.F.A. and M.A.; validation: N.A., M.R.I., M.F.A. and M.A.; formal analysis: N.A., M.R.I., M.F.A. and M.A.; investigation: N.A., M.R.I., M.F.A. and M.A.; data curation: N.A., M.R.I. and M.F.A.; writing—original draft preparation: N.A., M.R.I. and M.F.A.; writing—review and editing: N.A., M.R.I., M.F.A. and M.A.; supervision: M.R.I. and M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Makihara, Y.; Nixon, M.S.; Yagi, Y. Gait Recognition: Databases, Representations, and Applications. In *Computer Vision*; Springer: Cham, Switzerland, 2021.
2. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Inneural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [\[CrossRef\]](#)
3. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; pp. 6105–6114.
4. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006.
5. Kohonen, T. The Self-Organizing Map. *Proc. IEEE* **1990**, *78*, 1464–1480. [\[CrossRef\]](#)
6. Zheng, S.; Zhang, J.; Huang, K.; He, R.; Tan, T. Robust view transformation model for gait recognition. In Proceedings of the International Conference on Image Processing (ICIP), Brussels, Belgium, 11–14 September 2011.
7. Santos, C.F.G.; Oliveira, D.D.S.; Passos, L.A.; Pires, R.G.; Santos, D.F.S.; Valem, L.P.; Colombo, D. Gait recognition based on deep learning: A survey. *ACM Comput. Surv.* **2023**, *55*, 1–34. [\[CrossRef\]](#)

8. Saboor, A. Latest research trends in gait analysis using wearable sensors and machine learning: A systematic review. *IEEE Access* **2020**, *8*, 167830–167864. [[CrossRef](#)]
9. Jawed, B.; Khalifa, O.O.; Bhuiyan, S.S.N. Human gait recognition system. In Proceedings of the 7th International Conference on Computer and Communication Engineering (ICCCCE), Kuala Lumpur, Malaysia, 19–20 September 2018; pp. 89–92.
10. Zhang, S.; Wang, Y.; Li, A. Cross-view gait recognition with deep universal linear embeddings. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 9095–9104.
11. Mutlag, W.; Ali, S.; Mosad, Z.; Ghrabat, B.H. Feature extraction methods: A review. *J. Phys. Conf. Ser.* **2020**, *1591*, 012028. [[CrossRef](#)]
12. Pandey, N.; Abdulla, W.; Salcic, Z. Multiview gait recognition using sparse representation. In Proceedings of the International Conference on Image and Vision Computing New Zealand (IVCNZ), Dunedin, New Zealand, 2–4 December 2019; pp. 1–6.
13. Shirke, S.; Pawar, S.S.; Shah, K. Literature Review: Model Free Human Gait Recognition. In Proceedings of the 2014 4th International Conference on Communication Systems and Network Technologies, CSNT 2014, Bhopal, India, 7–9 April 2014; pp. 891–895.
14. Wu, Z.; Huang, Y.; Wang, L.; Wang, X.; Tan, T. A Comprehensive Study on Cross-View Gait Based Human Identification With Deep CNNs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 209–226. [[CrossRef](#)] [[PubMed](#)]
15. Chao, H.; He, Y.; Zhang, J.; Feng, J. Gaitset: Regarding gait as a set for cross-view gait recognition. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; pp. 8126–8133.
16. Mehmood, A.; Khan, M.A.; Sharif, M.; Khan, S.A.; Shaheen, M.; Saba, T.; Riaz, N.; Ashraf, I. Prosperous human gait recognition: An end-to-end system based on pre-trained CNN features selection. *Multimed. Tools Appl.* **2020**, *83*, 14979–14999. [[CrossRef](#)]
17. Saleem, F.; Khan, M.A.; Alhaisoni, M.; Tariq, U.; Armghan, A.; Alenezi, F.; Choi, J.I.; Kadry, S. Human Gait Recognition: A Single Stream Optimal Deep Learning Features Fusion. *Sensors* **2021**, *21*, 7584. [[CrossRef](#)] [[PubMed](#)]
18. Asif, M.; Tiwana, M.I.; Khan, U.S.; Ahmad, M.W.; Qureshi, W.S.; Iqbal, J. Human gait recognition subject to different covariate factors in a multi-view environment. *Results Eng.* **2022**, *15*, 100556. [[CrossRef](#)]
19. Deng, M.; Fan, Z.; Lin, P.; Feng, X. Human Gait Recognition Based on Frontal-View Sequences Using Gait Dynamics and Deep Learning. *IEEE Trans. Multimed.* **2024**, *26*, 117–126. [[CrossRef](#)]
20. Hou, S.; Fan, C.; Cao, C.; Liu, X.; Huang, Y. A Comprehensive Study on the Evaluation of Silhouette-Based Gait Recognition. *IEEE Trans. Biom. Behav. Identity Sci.* **2023**, *5*, 196–208. [[CrossRef](#)]
21. Qin, L.; Guo, M.; Zhou, k.; Sun, J.; Chen, X.; Qiu, J. Gait Recognition Based on Two-Stream CNNs with Multisensor Progressive Feature Fusion. *IEEE Sens. J.* **2024**, *24*, 13676–13685. [[CrossRef](#)]
22. Sepas-Moghaddam, A.; Etemad, A. Deep Gait Recognition: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 264–284. [[CrossRef](#)] [[PubMed](#)]
23. Vasudevan, P.; Sharma, P.; Ramesh, R. Gait image classification using deep learning models for medical diagnosis. *Comput. Mater. Continua* **2023**, *74*, 6039–6063. [[CrossRef](#)]
24. Alotaibi, M.; Mahmood, A. Improved Gait recognition based on specialized deep convolutional neural networks. In Proceedings of the 2015 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, USA, 13–15 October 2015; pp. 1–7.
25. Mehmood, A.; Ali, M.; Khan, M.A.; Ahmad, J.; Rehman, M.U. Human gait recognition: A deep learning and best feature selection framework. *Comput. Mater. Contin.* **2022**, *70*, 343–360. [[CrossRef](#)]
26. Gul, S.; Ahmad, J.; Shams, S.; Wahab, A.; Zafar, M.; Awan, S. Multi-view gait recognition system using spatio-temporal features and deep learning. *Expert Syst. Appl.* **2021**, *179*, 115057. [[CrossRef](#)]
27. Hasan, M.M.; Alam, S.; Hossain, S. Evaluating CNN models for gait recognition: A study on the CASIA-B dataset. *GUB J. Sci. Eng.* **2023**, *10*, 17–26. [[CrossRef](#)]
28. Song, C.; Tang, M.; Li, J.; Zeng, W.; Zhou, Y. CASIA-E: A large comprehensive dataset for gait recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 2801–2815. [[CrossRef](#)]
29. Abedi, H.; Ansariyan, A.; Morita, P.P.; Wong, A.; Boger, J.; Shaker, G. AI-powered non-contact in-home gait monitoring and activity recognition system based on mm-Wave FMCW radar and cloud computing. *arXiv* **2022**, arXiv:2208.05905.
30. Zhang, Y.; Wang, C.; Zhou, F.; Wu, J. RFPass: Towards environment-independent gait-based user authentication with RFID signals. *IEEE Internet Things J.* **2022**, *9*, 22739–22752.
31. Li, J.; Zhao, C.; He, W.; Wang, Y.; Yang, Y. GaitFi: Robust device-free human identification via Wi-Fi and vision multimodal learning. *arXiv* **2022**, arXiv:2208.14326.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.