

## Privacy-Preserving Abnormal Gait Detection Using Computer Vision and Machine Learning

A. Naz<sup>1</sup>, P. Prasad<sup>2</sup>, S. Mccall<sup>3</sup>, C. Leung<sup>3</sup>, I. Ochi<sup>3</sup>, L. Gong<sup>4</sup> and M. Yu<sup>3,\*</sup>

<sup>1</sup>School of Computer Science, University of Lincoln, Lincoln, UK

<sup>2</sup>International Institute of Information Technology, Bangalore, India

<sup>3</sup>School of Engineering and Physical Sciences, University of Lincoln, Lincoln, UK

<sup>4</sup>School of Agri-food Technology and Manufacturing, University of Lincoln, Lincoln, UK

### Abstract

Gait analysis plays a pivotal role in diagnosing a spectrum of neurological and musculoskeletal disorders. Variations in gait patterns often serve as early indicators of underlying health conditions, underscoring the importance of precise and timely analysis for effective intervention and treatment. In recent years, computer vision techniques have emerged as robust tools for automated gait analysis, offering non-invasive, cost-effective, and scalable solutions. However, existing approaches often overlook the critical aspect of privacy preservation. In this study, we propose the world's pioneering computer vision-based abnormal gait detection system with a privacy-preserving mechanism. Specifically, we extract 2D skeletons from encrypted images using a deep neural network model, which is facilitated by an optical system incorporating a custom-made refractive optical element. These extracted features are then fed into machine learning models for the detection of normal versus abnormal gait patterns. Evaluations across various models including random forest, decision tree, K-nearest neighbor, support vector machine, neural network, and convolutional neural network reveal that the random forest model attains the highest classification performance based on 2D skeletons extracted from encrypted images.

**Keywords:** Computer Vision, Gait Analysis, Abnormal Gait Detection

Received on 28 August 2024, accepted on 01 November 2024, published on 16 April 2025

Copyright © 2025 A. Naz *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetpht.11.9094

\*Corresponding author. Email: MYu@lincoln.ac.uk

### 1. Introduction

Gait embodies the rhythmic sequence of limb movements observed in animals, encompassing humans, as they traverse a stable surface. Human gait is intricately influenced by the coordinated function of the nervous, musculoskeletal, and cardiorespiratory systems, each playing a vital role in orchestrating fluid locomotion. Abnormal gaits serve as notable indicators of underlying neurological, musculoskeletal, or biomechanical irregularities [1]. Thus, timely identification of such aberrant gait patterns holds profound significance, offering a gateway to early diagnosis and intervention across a spectrum of medical conditions impacting human mobility.

Historically, identifying gait abnormalities relied heavily on subjective visual observation by clinicians, a method known for its inherent subjectivity and limited accuracy. However, the landscape has evolved with the emergence of modern sensor technologies and artificial intelligence (AI) techniques. Recent advancements have seen the integration of various sensor modalities/systems, such as pressure sensors [2], wearable sensors [3], and vision sensors [4], aimed at capturing intricate data pertaining to human movement during walking cycles. These data are subsequently processed and analyzed using state-of-the-art machine learning algorithms, facilitating the detection of abnormal gait patterns. Among the diverse sensor modalities, vision sensors have garnered significant attention for abnormal gait detection, owing to several advantages. Notably, vision sensors offer a cost-effective

solution, are non-invasive, and possess the capability to capture subtle nuances in gait. This adaptability and precision make vision sensors a preferred choice in the pursuit of more accurate and accessible abnormal gait detection methodologies.

However, a notable drawback of vision-sensor-based gait analysis is the issue of privacy. In such approaches, images of individuals are captured to analyze their gait patterns, potentially compromising their identities. Presently, to the best of our knowledge, there are no vision sensor-based solutions that adequately address the imperative of preserving privacy in gait analysis tasks. This underscores the critical need for developing innovative techniques that reconcile the objectives of gait analysis with safeguarding individuals' privacy rights.

In this study, we introduce a pioneering approach to privacy-preserving abnormal gait detection using vision sensors. Our developed algorithm effectively distinguishes between normal and abnormal gait patterns by analyzing 'blurred' images that safeguard individuals' identities. Specifically, our methodology entails the integration of an additional optical element, termed an 'optical encoder,' placed in front of the vision sensor lens. This encoder encrypts the original video frames, ensuring privacy protection. Subsequently, these encrypted frames are processed by a deep neural network to extract 2D skeletons, which are then inputted into machine learning models for gait classification. The parameters of the optical encoder and the deep neural network are trained collaboratively, to facilitate the successful extraction of 2D skeletons for gait analysis while simultaneously maximizing the differentiation between original and encrypted video frames, thereby ensuring robust privacy preservation. This innovative methodology represents a significant advancement in the field, offering a promising solution for vision privacy-conscious abnormal gait detection in real-world scenarios.

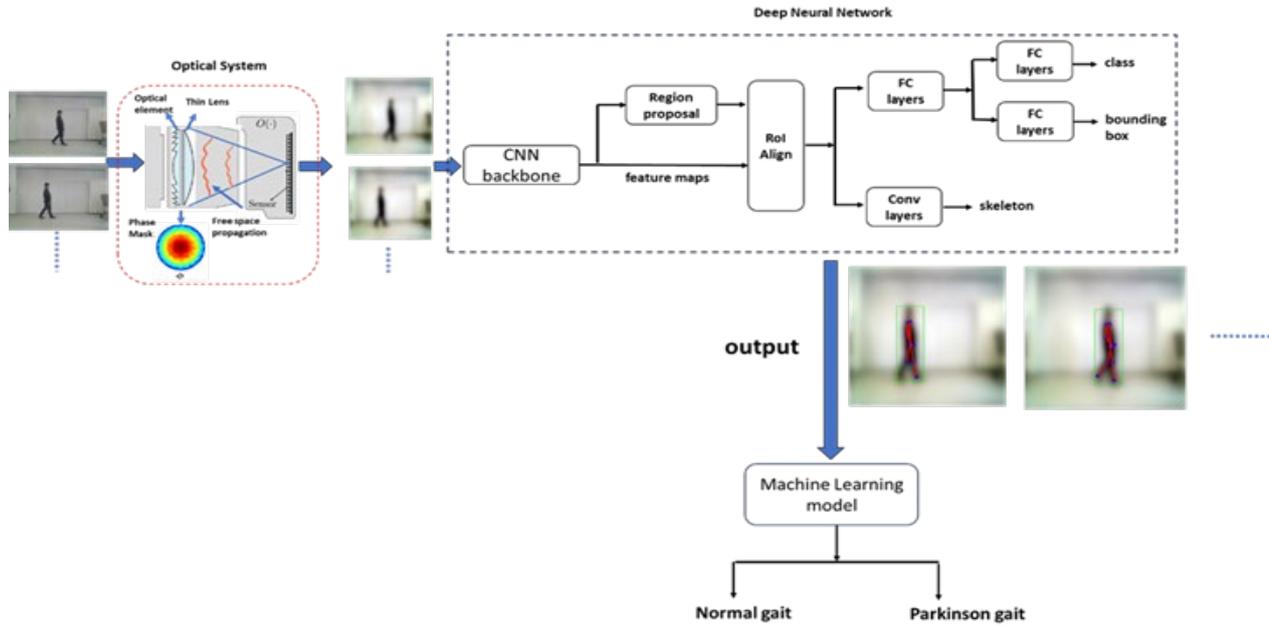
## 2. Related works

### 2.1. Abnormal gait detection

Privacy-preserving based motion analysis for healthcare in recent years, there has been significant research devoted to leveraging computer vision and machine learning methodologies for abnormal gait detection, as evidenced by various studies documented in the literature [4,5]. Specifically, in [5], researchers employed support vector machine (SVM) techniques to distinguish between normal and abnormal gait patterns based on 2D lattice features derived from silhouettes extracted from video sequences. Their experimentation on a relatively modest gait video dataset, comprising 5 individuals, yielded an approximate classification accuracy of 80%. In [6], a clinically interpretable approach rooted in computer vision was proposed. This method involved extracting 2D poses from recorded video sequences, from which features characterizing walking speed, arm swing, and postural

control were derived. These features were then inputted into a random forest classifier to estimate the Unified Parkinson's Disease Rating Scale (UPDRS) rating. Their experiments, conducted on a substantial dataset comprising over 600 subjects, demonstrated a notable correlation between clinician assessments and model predictions. In [7], the researchers employed an Extreme Learning Machine (ELM) for gait classification, particularly focusing on children, utilizing Salient Gait Features. Their experimental findings suggested that ELM offered improved classification accuracies along with reduced training time and implementation complexity compared to SVM. In [8], an algorithm was introduced for classifying gait disorders stemming from neurodegenerative diseases such as Parkinson's and Hemiplegia. This method involved capturing the motion trajectories of body joints using a Kinect sensor, followed by preprocessing to ensure position and scale invariance. Subsequently, trajectory descriptors representing joints' relative positional and speed characteristics were extracted from segmented motion sequences. Abnormal gait detection was then performed using a K-nearest neighbor (KNN) approach, which compared trajectory descriptors between testing sequences and a training dataset. In [9], researchers proposed utilizing full gait energy images (F-GEI) as features extracted from gait images, along with a COP-K-means semi-supervised clustering method for normal/abnormal gait classification. The experimental results demonstrated the effectiveness and efficiency of the proposed approach in terms of accuracy, robustness, and computational efficiency when compared with other state-of-the-art abnormality detection techniques. [10] introduced a novel approach that only leverages normal gait samples to train a one-class support vector machine (OCSVM) model for classifying normal versus Parkinsonian gait patterns utilizing Gait Energy Images (GEIs) extracted from the mask RCNN neural network model. Experimental results obtained from a gait dataset encompassing 30 participants, as well as online evaluations, showcased the potential for achieving high accuracy (exceeding 97%) using OCSVM trained solely on normal gait samples.

Researchers have not only focused on employing specific machine learning models but have also conducted evaluations and comparisons of various models for abnormal and normal gait classification. In [11], the Kinect Motion system was utilized to capture spatiotemporal gait data during walking. Subsequently, different machine learning models, including convolutional neural networks (CNN), support vector machines (SVM), K-nearest neighbors (KNN), and long short-term memory (LSTM) neural networks, were applied to classify three distinct walking patterns: normal gait, pelvic-obliquity-gait, and knee hyperextension-gait, for seven subjects. The findings revealed that SVM achieved the highest accuracy in classifying gait patterns, reaching approximately 95%. In [12], an approximation of skeleton joints was derived from



**Figure 1.** The flowchart of the proposed methodology

silhouettes extracted from RGB side-view gait video sequences. Various parameters, such as heel strike, toe-off, stride length, and time, were then extracted and inputted into different classifiers, including KNN, Support Vector Machine (SVM), and Bayesian classifier algorithms, to discern normal from abnormal gaits. The experiments demonstrated that the most accurate results were attained with KNN, particularly on leg-angle variables, achieving a remarkable 100% accuracy rate on a relatively small dataset comprising only 30 samples.

Recently there has been a surge in the application of deep learning methodologies for the classification of normal and abnormal gait, yielding promising outcomes. In [13], a range of machine learning and deep learning techniques, such as support vector machines (SVM), multilayer perceptron (MLP), Vanilla long short-term memory (LSTM), and Bidirectional LSTM, were employed to diagnose Cerebral Palsy (CP) gait based on the linear velocity of seven body joints. The findings indicated that LSTM achieves the best performance in detecting CP gait. In [14], various deep neural network models, including Temporal Convolutional Networks, Gated Recurrent Units, and Long Short-Term Memory Autoencoders, were applied to monitor activities of daily living (ADLs) for the identification of abnormal gaits, based on gait features extracted from 3D skeletons obtained from video sequences. Experimental results showcased high classification accuracies of more than 96% across multiple datasets for different deep neural network models.

Vision data can be effectively integrated with other modalities for enhanced abnormal gait classification. In [15], a hybrid neural network model, comprising both recurrent neural network (RNN)-based encoding layers and convolutional neural network (CNN)-based encoding layers, was developed by incorporating both 3D skeleton

data and pressure data aiming to classify normal and abnormal gait patterns. Experimental findings demonstrated that the proposed hybrid model yielded superior performance, achieving an accuracy of 95.66%. This outperformed models relying solely on a single sensor modality, highlighting the efficacy of leveraging multiple data sources for gait classification.

## 2.2. Privacy-preserving based motion analysis for healthcare

Currently, there has been some research focusing on the application of computer vision and deep learning techniques to analyze images and videos encrypted via software/hardware-based methods for privacy preservation. The related research spans various vision tasks such as privacy preserving-based pose estimation [16], action recognition [17], and image caption generation [18]. However, only a limited number of studies have delved into computer vision-based human motion analysis for healthcare with privacy preservation. [19] utilizes compressed sensing to detect falls while safeguarding privacy through the utilization of low-quality compressed data, whereas [20] proposes a hardware-based privacy-preserving solution, which involves the application of a custom-designed optical element whose parameters are jointly trained with a deep neural network for fall detection from encrypted images. To the best of our knowledge, there has been no prior work addressing the privacy preservation challenge in vision-based gait analysis.

### 3. Methodology

The flowchart of the proposed methodology is shown in Fig. 1, from which we can see that the proposed methodology contains three main components, an optical system for image encryption, a deep neural network for skeleton features extraction, and a machine learning model for gait classification. Different components are introduced in detail in the next few sections.

#### 3.1. Optical System

The optical system, which comprises an optical element used for image encryption to preserve privacy, a lens, and an image sensor as shown in Fig. 1, captures the encrypted images from the original scene. In specific, assuming that the scene is at optical infinity, then the light wave field arriving before the optical element can be represented as a planar wave field, denoted as  $U$ . The refractive optical element introduces phase delay of this incident wavefront for blurring/encrypting original scene, by an amount proportional to the surface profile  $\phi$  of the optical element at each point  $(x,y)$ , which is represented by:

$$t_\phi(x,y) = \exp(ik(n(\lambda) - 1)\phi(x,y)) \quad (1)$$

where  $n(\lambda)$  is the wavelength-dependent refractive index of the optical element material,  $k = \frac{2\pi}{\lambda}$  is the wavenumber, and  $\phi(x,y)$  is represented as:

$$\phi = \sum_{j=1}^N \alpha_j Z_j \quad (2)$$

where  $Z_j$  is the  $j$ -th Zernike polynomial in Noll notation as in [17], and  $\alpha_j$  is the coefficient that is trainable. After the optical element, the light wave continues to propagate to the camera lens, which induces the following phase transformation:

$$t_l(x,y) = \exp\left(-i\frac{k}{2f}(x^2 + y^2)\right) \quad (3)$$

Considering that a lens has a finite aperture size, we use a binary circular mask  $A(x,y)$  with diameter  $D$  to model the aperture and block light in regions outside the open aperture. The light wave field after the lens can be represented as:

$$\tilde{U}(x,y) = A(x,y)t_\phi(x,y)t_l(x,y)U \quad (4)$$

Finally, the light wave field propagates a distance  $d_2$  to the sensor with the exact transfer function as in [17]

$$T(f_{x'}, f_{y'}) = \exp\left(ikd_2\sqrt{1 - (\lambda f_{x'})^2 - (\lambda f_{y'})^2}\right) \quad (5)$$

where  $(f_{x'}, f_{y'})$  represents spatial frequency. Based on this transfer function, the light wave field arriving at the image sensor could be represented as:

$$U(x', y') = \mathcal{F}^{-1} \left\{ \mathcal{F} \left( \tilde{U}(x, y) \right) T_\lambda(f_{x'}, f_{y'}) \right\} \quad (6)$$

where  $\mathcal{F}\{\}$  and  $\mathcal{F}^{-1}\{\}$  represent the Fourier and inverse Fourier transforms. The magnitude of  $U(x',y')$ , which is calculated as:

$$H(x',y') = |U(x',y')|^2 \quad (7)$$

is defined as the point spread function (PSF). The encrypted image acquired by the optical system can then be modeled by the PSF as:

$$y = H \otimes x + \eta \quad (8)$$

where  $y$  represents the encrypted output image generated by the optical system,  $x$  represents the original scene which is represented as a discrete *color image*,  $\eta$  represents the optical system noises and  $\otimes$  represents the convolutional operation.

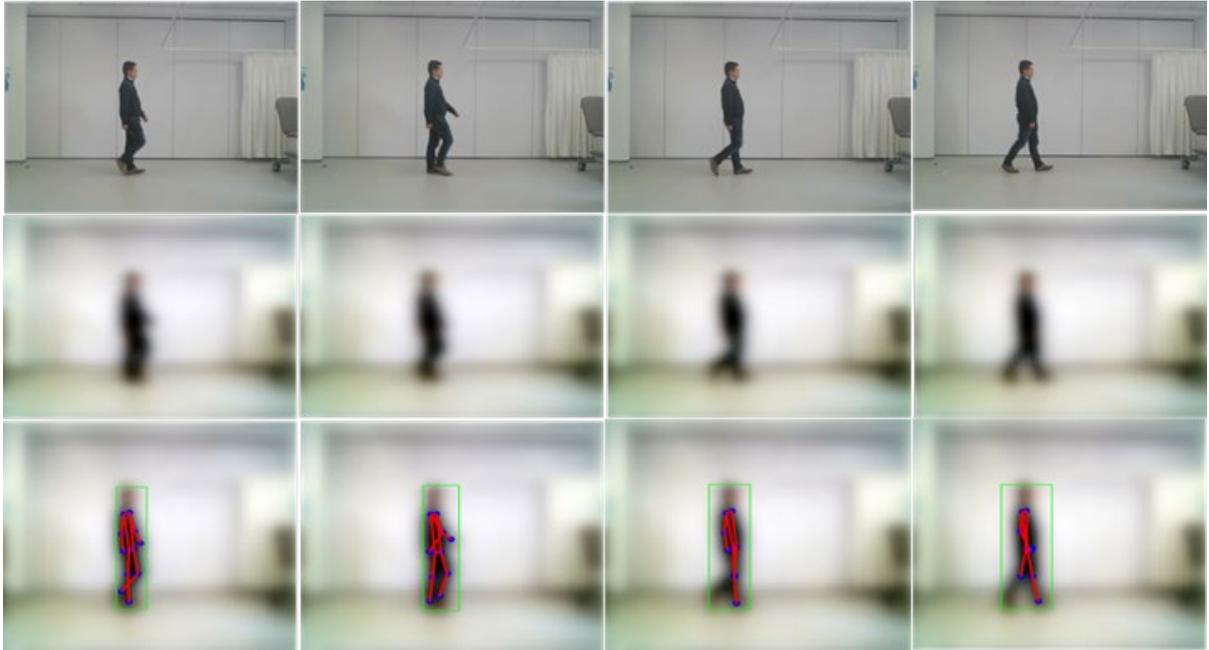
#### 3.2. Deep neural network for skeleton extraction

The encrypted image outputs generated by the optical system serve as the input for a sophisticated deep neural network model designed for 2D pose estimation. Our chosen architecture for this task is the mask RCNN (Region-based Convolutional Neural Network), renowned for its capability in object detection and instance segmentation. This architecture unfolds in two distinct stages, each contributing to the accurate extraction of human pose information.

In the initial stage, a Region Proposal Network (RPN) takes charge, leveraging convolutional neural network (CNN) backbones, such as Resnet-50 or Resnet101, to extract essential feature maps. The RPN then proceeds to propose regions of interest (RoIs) by enclosing candidate objects within bounding boxes. This step serves as a crucial precursor for identifying relevant areas that warrant further analysis in the subsequent stage. The second stage involves refining the identified RoIs using an ROIAlign layer, a pivotal component that ensures the preservation of spatial information without introducing misalignments. The ROIAlign process entails dividing each RoI into a fixed grid and subsequently sampling the feature map at regularly spaced points within each grid cell. The interpolation is achieved through bilinear methods, resulting in re-sized RoIs that possess a standardized and consistent format. The processed RoIs are then directed to two distinct heads within the model. The first head is equipped with fully connected layers responsible for classifying human objects and regressing the human region. This component plays a pivotal role in determining the presence of human subjects within the image and accurately delineating their spatial boundaries. Simultaneously, the second head employs a series of

convolutional layers to extract detailed information regarding the human 2D skeletons. This intricate process enables the model to capture the poses and skeletal structures, contributing to the overall precision of the 2D pose estimation.

In this work, we define the following loss function to jointly train the optical element coefficients in (2) and mask RCNN network:



**Figure 2.** The original video frames, encrypted video frames and the 2D skeleton extraction results.

$$L = L_{cls} + L_{reg} + L_{skeleton} - \lambda \cdot L_{difference} \quad (9)$$

where  $L_{cls}$ ,  $L_{reg}$ , and  $L_{skeleton}$  represent the classification, bounding box regression, and skeleton extraction losses associated with the mask RCNN,  $L_{difference}$  represents the loss of the difference between the original scene images and encrypted images obtained from the optical system. The balance factor  $\lambda$  is used to adjust the relative importance of these components. Minimizing the above loss function enables the mask RCNN detector to effectively extract 2D skeletons while simultaneously encrypting the images through the optical system for privacy-preserving purposes.

### 3.3. Machine learning models for gait classification

The 2D skeletons, derived from a video sequence through our deep neural network model, play a pivotal role in the classification of gaits as normal or indicative of Parkinson's disease, leveraging a specialized

classification model. Prior to delving into the classification process, we initiate a series of pre-processing steps. In the initial phase of pre-processing, we establish the midpoint of the shoulders as the reference point (origin). This serves as a fundamental anchor, and we subsequently adjust the position of each key point accordingly to obtain their relative positions. Following the establishment of relative key point positions, we further standardize these positions using

the length of the trunk as a normalization factor. This normalization step is pivotal in accounting for variations in individual body sizes in the image planes. With the

completion of these pre-processing steps, we proceed to transform individual key points, denoted as  $p_i$  via the following formula:

$$\hat{p}_i = \frac{p_i - p_1}{L} \quad (10)$$

where  $p_i$  represents the position of the shoulder and  $L$  represents the trunk length. By implementing the aforementioned pre-processing techniques, we ensure the stability and consistency of the extracted 2D poses, making them invariant with respect to both scale and position variations. This robustness enables our developed technique to effectively classify normal/abnormal gaits across diverse positions and distances from the camera when processing videos. For each frame, denoted as the  $i$ -th frame, we construct a concatenated vector, represented as  $c_i = [\hat{p}_{i,1}, \dots, \hat{p}_{i,N}]$ , by amalgamating the pre-processed positions of all keypoints. Here,  $\hat{p}_{i,j}$  represents the pre-processed position of the  $j$ -th keypoint at the  $i$ -th frame.

Subsequently, we compile these concatenated vectors for frames within a video sequence, forming an input array  $[c_1, \dots, c_N]$ , where  $N$  denotes the sequence length. Concatenated vectors obtained from video sequences in training/testing datasets are applied to train/test machine learning models for normal/abnormal gait classification. In this study, we evaluate various machine learning models, including k-NN, decision tree, random forest, support vector machine, neural network, and convolutional neural network. This comprehensive analysis ensures the selection of an optimal model that can effectively discern gait abnormalities.

#### 4. Experimental studies

The devised technique underwent evaluation using a dataset containing videos illustrating both normal and Parkinson's side-walking gaits of sixteen subjects within the laboratory at the University of Lincoln. This video dataset was meticulously segmented, resulting in a total of 320 samples, of which 192 were designated for training and 128 for testing. Fig. 2, Line 1 provides a glimpse of samples extracted from the recorded video frames. Initially, the stochastic gradient descent (SGD) algorithm was employed to jointly train the Zernike polynomial coefficients within a simulated optical system and the Mask RCNN using a substantial COCO dataset [23]. The base learning rate was configured at 0.00003, accompanied by a weight decay of 0.00001 and a momentum of 0.9. To facilitate stable training, a warm-up iteration of 100 was implemented, with the overall training epoch set to 50. Leveraging the trained optical system and Mask RCNN, image encryption and 2D pose extraction were conducted. Notably, exemplary results of encryption and pose extraction are presented in lines 2 and 3 of Fig. 2 for reference.

Skeleton features are extracted from both the training and testing video sequences and subjected to preprocessing steps outlined in Section 3.3. Subsequently, the preprocessed features derived from the training dataset are utilized to train various machine learning models, which are subsequently evaluated on the test datasets. This study employs a diverse array of classifiers, encompassing traditional methods such as random forest (RF), K-Nearest Neighbor (KNN), support vector machine (SVM), and decision tree (DT), alongside advanced deep learning approaches including convolutional neural network (CNN) and artificial neural network (ANN)

*Machine learning classifiers evaluation:* We trained machine learning models, including those based on the scikit-learn Python library [24]. For each model, we utilized the GridSearchCV function within scikit-learn to search for optimal hyperparameters, ensuring optimal model performance. The confusion matrices obtained for various models under the optimal hyperparameters are depicted in Fig. 3. Notably, the analysis reveals that the

random forest approach yields the most favorable outcome, exhibiting the fewest misclassified samples.

For a more comprehensive quantitative evaluations, we calculate different evaluation metrics including accuracy, precision, recall and f1-score, whose definitions are shown as below:

$$precision = \frac{TP}{TP + FP} \quad (11)$$

$$recall = \frac{TP}{TP + FN} \quad (12)$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (13)$$

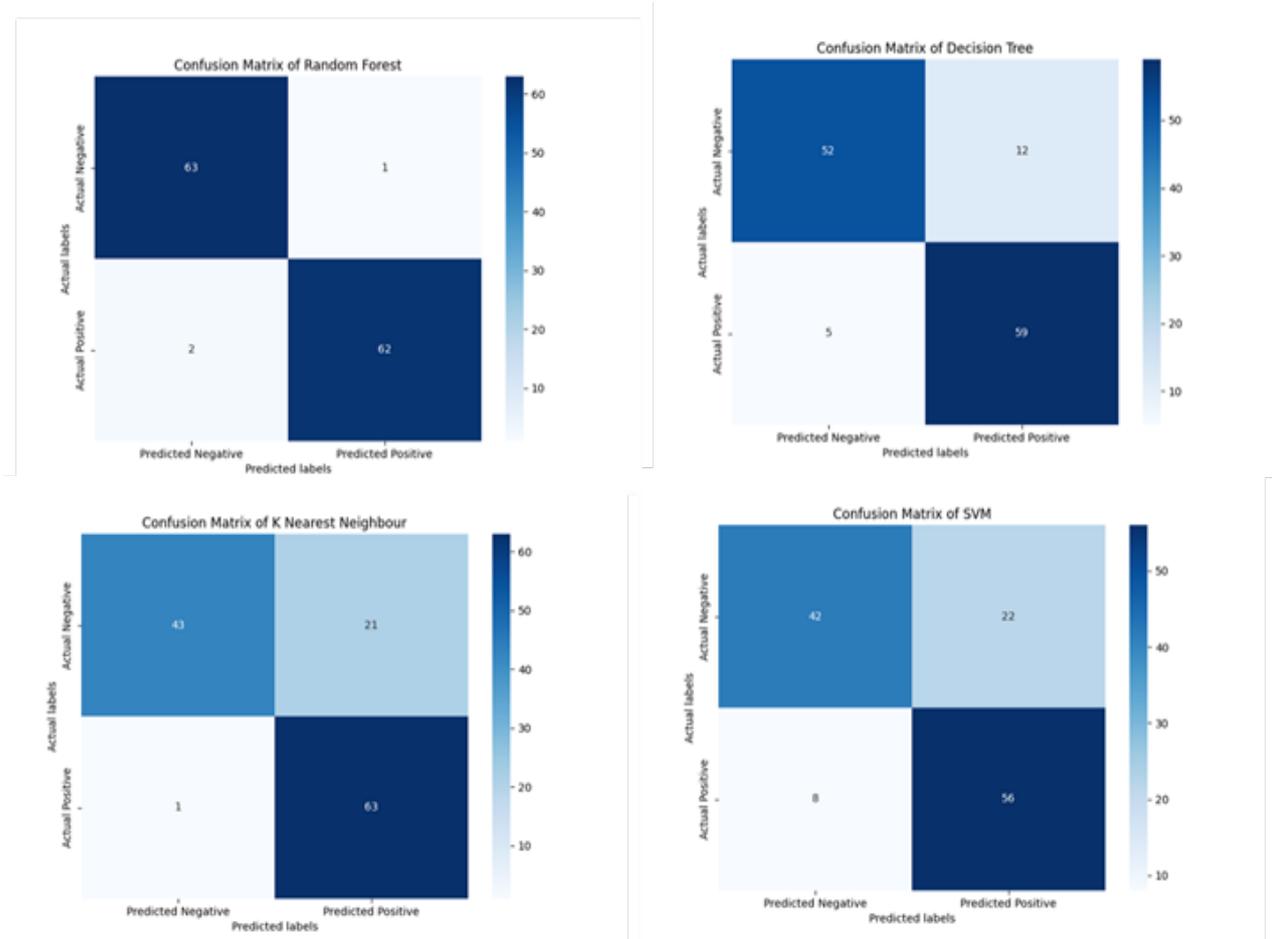
$$f1_{score} = \frac{2 \cdot precision \cdot recall}{precision + recall} \quad (14)$$

where TP, TN, FP, and FN represent true positive (correctly classified normal gait sample), true negative (correctly classified abnormal gait sample), false positive (incorrectly classified normal gait sample), and false negative (incorrectly classified abnormal gait sample) respectively. The closer these values to 1, the better a model is. The average results are shown in Table 1, from which we can see that the random forest based approach achieves the best performance with respect to the majority of evaluation metrics, especially with a much higher average accuracy than other models.

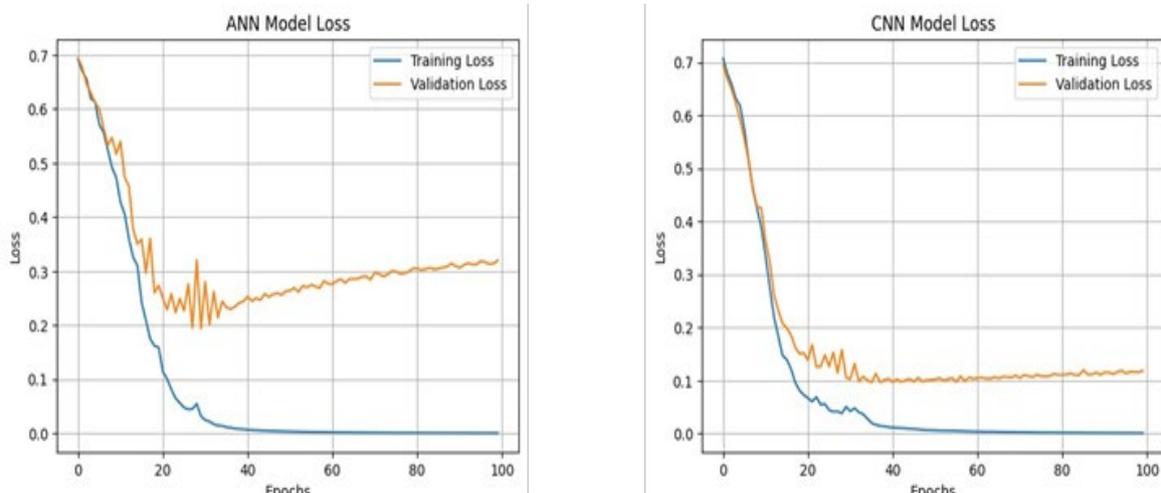
Table 1: Performance comparisons of different machine learning models

Model	Accuracy	Precision	Recall	F1 Score
RF	<b>0.9766</b>	<b>0.9841</b>	0.9688	<b>0.9764</b>
SVM	0.7656	0.7179	0.8750	0.7887
KNN	0.8281	0.7500	<b>0.9844</b>	0.8514
DT	0.8672	0.8310	0.9219	0.8741

*Deep learning classifiers evaluation:* We also experimented with two variations of deep neural network models for evaluation purposes. The first model is a standard artificial neural network, featuring four hidden layers with neuron counts of 128, 64, 16, and 32



**Figure 3.** Confusion matrices for different machine learning models.



**Figure 4.** Training/validating loss curves for the ANN & CNN model

respectively. Relu activation functions are applied to the hidden layers, while the output layer comprises a single neuron utilizing the sigmoid activation function, which outputs the probability of a normal gait. The CNN model comprises three convolutional layers with 128, 256, and 256 filters respectively, each with a kernel size of 3, followed by fully connected layers. During the training

of these network models, the binary cross-entropy loss function is employed and defined as follows:

$$L(y,p) = -(y \cdot \log(p) + (1 - y) \cdot \log(1 - p)) \quad (15)$$

where  $y$  is the true class label and  $p$  is the predicted probability. The Adam algorithm is used for the network

training. The loss curves on both the training/testing datasets during the network training are presented in Fig. 4. The diverse metrics obtained for both network models are presented in Table 2. It is evident from the results that the performance of the ANN model surpasses that of the CNN, exhibiting higher values across all metrics.

Upon examining the experimental results, it becomes evident that the Random Forest model outperformed all other approaches across various evaluation metrics. Both ANN and CNN exhibited inferior performance compared to traditional algorithms, as indicated by lower precision, recall, and F1 score values relative to the Random Forest model. This discrepancy can be partly attributed to the size of the dataset. With only 192 datapoints in the training set and 128 in the testing set, neural networks might struggle to generalize effectively due to their high capacity.

Table 2: Performance metrics of ANN and CNN models

Model	Precision	Recall	F1 Score
ANN	<b>0.9254</b>	<b>0.9688</b>	<b>0.9466</b>
CNN	0.8806	0.9219	0.9008

Random Forest, being an ensemble method, demonstrates resilience in such scenarios with limited data, relying on decision trees that are less prone to overfitting. Additionally, the dataset's features might be inherently more suitable for Random Forest without extensive feature processing. Moreover, Random Forest is better equipped to handle outliers and noise within the data. The decision trees within Random Forest are less influenced by outliers, as the final prediction is aggregated from multiple trees, diminishing the impact of individual outliers. Consequently, Random Forest exhibits superior performance in scenarios where outliers and noise are prevalent.

## 5. Conclusion

This study introduces a pioneering approach to computer vision-based normal/abnormal gait classification, prioritizing privacy preservation. Our methodology integrates the optimization of optical system parameters and deep neural network models, facilitating the precise extraction of human 2D skeletons from encrypted images. Subsequently, these 2D skeletons undergo thorough preprocessing before being input into machine learning models for the classification of normal and abnormal gaits. Our experimental findings underscore the efficacy of our approach, with the random forest model emerging as the top performer across various evaluation metrics, validating its suitability for our dataset

However, one notable limitation of our current study lies in the size of the dataset utilized for training and evaluation. We acknowledge that the dataset employed in this research is relatively small. To address this, we plan to expand our dataset significantly in future studies, thereby enhancing the robustness and generalizability of our models. Additionally, we aim to explore the development of more advanced and complex machine learning models tailored specifically for gait classification. Through these efforts, we aspire to further advance the state-of-the-art in gait classification while ensuring the preservation of privacy in our methodologies.

## Acknowledgements

This research has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 778602 ULTRACEPT.

## References

- [1] Pirker, W., Katzenschlager, R.: Gait disorders in adults and the elderly: A clinical guide. *Wien Klin Wochenschr* **129**(3), 81–95 (2017)
- [2] Dai, Y., Gao, J., Zhang, W., Wu, X., Zhu, X., Gu, W.: Smart Insoles for Gait Analysis Based on Meshless Conductive Rubber Sensors and Neural Networks. *Journal of Physics: Conference Series* **2500**(2023), 1–8 (2007)
- [3] Potluri, S., Ravuri, S., Diedrich, C., Schega, L.: Deep Learning based Gait Abnormality Detection using Wearable Sensor System. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 2019
- [4] Hundal, J., Badu, B.: Computer Vision and Abnormal Patient Gait: A Comparison of Methods. *The Open Artificial Intelligence Journal* **6**, 29–34 (2020)
- [5] Bauckhage, C., Tsotsos, J., Bunn, F.: Detecting Abnormal Gait. In: The 2nd Canadian Conference on Computer and Robot Vision (CRV'05), Victoria, BC, Canada, 2005
- [6] Ruppel, S., et al.: A Clinically Interpretable Computer-Vision Based Method for Quantifying Gait in Parkinson's Disease. *Sensors* **21**(16), 1–21 (2021)
- [7] Rani, M., Arumugam, G.: Children Abnormal GAIT Classification Using Extreme Learning Machine. *Global Journal of Computer Science and Technology* **22**(20), 11–22 (2022)
- [8] Li, Q., Wang, Y., Sharf, A., Cao, Y., Tu, C., Chen, B., Yu, S.: Classification of gait Anomalies from Kinect. *The Visual Computer* **34**, 229–241 (2018)
- [9] Yang, Z.: An Efficient Automatic Gait Anomaly Detection Method Based on Semi-supervised Clustering. *Computational Intelligence and Neuroscience* **2021**
- [10] Gong, L., Li, J., Yu, M., Zhu, M., Clifford, R.: A novel computer vision-based gait analysis technique for normal and Parkinson's gaits classification. In: 2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology | Congress

- (DASC/PiCom/CBDCCom/CyberSciTech), Calgary, AB, Canada, 2020
- [11] Chen, B., Chen, C., Hu, J., Sayeed, Z., et al.: Computer Vision and Machine Learning-Based Gait Pattern Recognition for Flat Fall Prediction. *Sensors* **22**(20), 11–22 (2022)
- [12] Nieto-Hidalgo, M., Ferrandez-Pastor, F., Valdivieso-Sarabia, R., Mora-Pascual, J., Garcia-Chamizo, J.: A vision-based proposal for classification of normal and abnormal gait using RGB camera. *Journal of Biomedical Informatics* **63**, 82–89 (2016)
- [13] Chakraborty, S., Sambhavi, S., Nandy, A.: Deep Learning in Gait Abnormality Detection: Principles and Illustrations. In: Suresh, A., et al. (eds.) *Bioinformatics and Medical Applications: Big Data Using Deep Learning Algorithms*, pp. 63–72. Scrivener Publishing LLC, 2022
- [14] Diraco, G., Manni, A., Leone, A.: Integrating Abnormal Gait Detection with Activities of Daily Living Monitoring in Ambient Assisted Living: A 3D Vision Approach. *Sensors* **24**(1), 1–22 (2024)
- [15] Jun, K., Lee, S., Lee, D., Kim, M.: Deep Learning-Based Multimodal AbnormalGait Classification Using a 3D Skeleton and Plantar Foot Pressure. *IEEE Access* **9**, 161576–161589 (2021)
- [16] Hinojosa, C., Niebles, J., Arguello, H.: Learning privacy-preserving optics for human pose estimation. In: *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, 2021
- [17] Hinojosa, C., Marquez, M., Arguello, H., Adeli, E., Fei-Fei, L., Niebles, J.: PrivHAR: Recognizing Human Actions from Privacy-Preserving Lens. In: *European Conference on Computer Vision*, Tel Aviv, Israel, 2022
- [18] Arguello, P., Lopez, J., Hinojosa, C., Arguello, H.: Optics Lens Design for Privacy Preserving Scene Captioning. In: *IEEE International Conference on Image Processing (ICIP)*, Bordeaux, France, 2022
- [19] Liu, J., Xia, Y., Tang, Z.: Privacy-preserving video fall detection using visual shielding information. *Vis. Comput.* **37**, 359–370 (2021)
- [20] Gong, L., Mccall, S., Yu, M.: Enhancing privacy with optical element design for fall detection. *Electronic Letters* **59**(20), (2023)
- [21] Goodman, J.: *Introduction to Fourier optics*. 4th edn. Macmillan Learning, 2017
- [22] He, K., Gkioxari, G., Dollr, P., Girshick, R.: Mask R-CNN. In: *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017
- [23] COCO: Common Objects in Context, <https://cocodataset.org/home>. Last accessed 28 Mar 2024
- [24] scikit-learn Machine Learning in Python, <https://scikit-learn.org/stable/>. Last accessed 28 Mar 2024