Gait Pattern Recognition via Fusing Static and Dynamic Features in Single Stream Network

1st Isha Jain
Electrical Engineering
Indian Institute of Technology
Kanpur, INDIA
ishajain21@iitk.ac.in

2nd Navya
Electrical Engineering
Indian Institute of Technology
Kanpur, INDIA
navya21@iitk.ac.in

3rd Dharini Reddy
Electrical Engineering
Indian Institute of Technology
Kanpur, INDIA
dharinir21@iitk.ac.in

4th Vijay Kumar Pandey
Electrical Engineering
Indian Institute of Technology
Kanpur, INDIA
vijaykp23@iitk.ac.in

5th Tushar Sandhan
Electrical Engineering
Indian Institute of Technology
Kanpur, INDIA
sandhan@iitk.ac.in

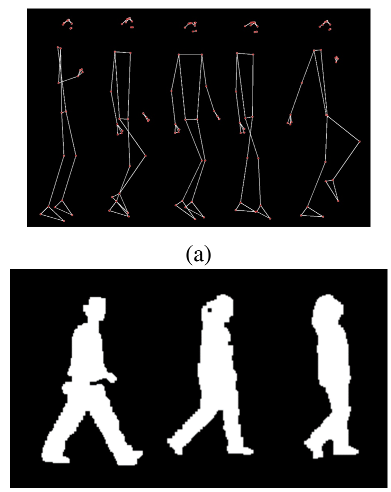
***Abstract*—Gait refers to the walking and motion characteristics of an individual. In this paper, we suggest a unique method for analyzing human gait patterns using static as well as dynamic features passing through the same convolutional neural network. Generally, the gaits are processed using contour detection and segmentation to perform time series analysis with extracted features. Furthermore, machine learning models such as decision trees, support vector machine are also utilized for the gait classification. We initially trained our model on both cropped average images and optical flow of each gait independently. Subsequently, we fused the average image of each gait (building static information) with its corresponding optical flow (dynamic information). This novel fusion approach facilitated the training of a single stream convolutional neural network model. Our findings reveal a notable accuracy of 98.75% in successfully recognizing individuals based on their gait patterns. Simplicity of the model and succinct fusion of static and dynamic features produce improved results with very less computational time.**

***Index Terms*—Human gait recognition, Time-series analysis, Convolutional neural network, Silhouette, Image feature integration**

I. Introduction

The intricate biomechanical pattern of human gait has long been a subject of fascination and inquiry within the scientific community. A basic method for examining how people move, gait analysis has numerous uses in the fields of healthcare, rehabilitation and security [1]–[3]. Traditionally, gait analysis has been carried out in a room-scale setting with costly sensors and sophisticated cameras [4], [5]. Gait analysis, sometimes referred to as walking or motion analysis, is an assessment of an individuals standing and walking posture. The unique characteristic of a human gait is influenced by various elements, including weight, limb length, posture, footwear, and characteristic motion [6].

Advancements in gait analysis have been crucial for clinical applications, as highlighted by Baker and M in "Gait Analysis: Clinical Facts" [7]. Their work emphasizes the clinical significance of gait analysis, offering essential insights into the diagnosis and rehabilitation of various musculoskeletal conditions [7].



(a)

(b)

Fig. 1: Gait is the distinctive signature of human movement. Skeleton and silhouette representation of images: (a) skeleton of the image obtained using MediaPipe library, (b) Silhouette images from the CASIA dataset [8]

Furthermore, gait assessment methodologies have facilitated enhanced diagnostic precision and efficiency for individuals afflicted with motor skill impairments due to injury. These advancements also enable comparative analyses of key indicators between rehabilitative and forensic gait assessments [9].

Researchers have noticed in their study that just like fingerprint and iris, almost every individual has different walking style [10]. If we compare with fingerprint and iris based identification, human gait identification does not require the user's interaction and can be done at a distance, as long as the gait is visible. Gait analysis helps in variety of applications such as biometric identification, security and access control and discrete identification [11], [12]. The features of human gait patterns are very difficult to conceal and can be utilized in security systems as well as for person re-identification [13]. The most important feature of gait as a biometric characteristic is its unnoticeable charecteristics, allowing it to be gathered without the subject's agreement. Gait analysis offers an advantage in situations where traditional biometrics like fingerprints or facial recognition might not be suitable, such as in low-light conditions or when subjects are at a distance from cameras [14]. In order to proceed with gait analysis techniques, gaining a better understanding of the machine and the procedure of utilizing those machines is equally important. With the support of improved utilization of techniques, data collection has become much simpler. Few general approaches of representing the walking patterns like skeleton images using pose estimation and silhouette images contributes to the dataset of interest [15], [16].

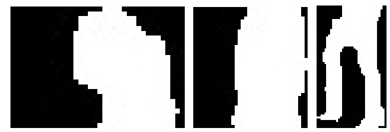
The intersection of human activity recognition and machine learning comes to the forefront in the work of Ronao and Cho [17], about human activity recognition using smart phones, underscores the potential of advanced technologies, paving the way for innovative applications in gait analysis. The most popular approaches for automated gait analysis have been found to be those based on SVM and convolutional neural network (CNN) [18]. Another approch of support vector machines (SVM) has few advantages over CNN such as high computing efficiency and the capacity to work with relatively samaller datasets [19], [20].

The proposed work focuses particularly on machine learning approaches mainly CNN, to transform our understanding of human mobility. Our approach involves a multi-modal analysis, incorporating both visual and coordinate-based information to comprehensively capture and interpret the nuances of human gait. We have fused the static and dynamic feature information from the silhouette images and used it as input to train our CNN model which enhances the performance of our model. Through this multi-modal approach, combining visual cues with coordinate-based information, our research aims to overcome the limitations of traditional gait analysis methods.

II. Related work

Contact-based methods involve the placement of sensors or markers on specific body parts to capture biomechanical data during gait analysis. These methods are often used in clinical settings and biomechanics research [7], [21]. The sensors measure parameters such as force, pressure, and joint angles to analyze gait patterns [22].

1) Wearable Sensors and Machine Learning:: While wearable sensors have many benefits, they also have drawbacks. Firstly, they only provide localized data, which leaves a lack of comprehensive full-body kinematics and secondly, data processing complexity arises from challenges in integrating and synchronizing data from diverse sensor types or brands [18]. Wearable sensors use wireless connectivity for real-time data transmission to a computer or mobile device. The collected data is then processed and analysis, employing algorithms to identify key gait events and assess biomechanical parameters such as step length and gait symmetry. This is followed by an interpretation of the data to detect abnormalities and key features for identification [18].



(a) (b) (c)

Fig. 2: Segmentation of individual cropped image into (a)head, (b)torso, and (c)legs for detailed analysis and time series formation

A. Non-contact based Methods:

1) Computer vision and machine learning:: Vision-based methods utilize computer vision techniques and cameras to capture and analyze gait patterns. These methods offer non-intrusive monitoring and have applications in surveillance, healthcare, and biometrics. Computer vision systems track and identify key features of a person's body to analyze gait patterns [23], [24].

Learning-based approaches leverage machine learning and deep learning algorithms to analyze gait patterns from sensor data or images. These methods can automatically extract features and classify gait patterns, contributing to automated gait analysis systems. Machine learning (ML) algorithms, such as neural networks, are trained to recognize unique gait features and patterns [25], [26].

Novel methods based on ML include the usage of deep artificial neural networks to provide comprehensive models for complex physical, biological, and security systems. By inputting model predictions back to input variables, it demonstrates which variables are most important for gait analysis. Other authors used various algorithms like RIPPER, Random Forest, Decision Tree (DT), IB1, Bootstrap aggregating, and BayesNet where random forest classifiers showcased exceptional accuracy.

Integration of ML algorithms in gait analysis softwares for identification purposes greatly simplifies evaluation and increases accuracy.

B. Hybrid Methods:

Hybrid methods integrate multiple techniques, such as sensor-based data collection, computer vision, and machine learning based approaches to enhance the accuracy and robustness of gait analysis systems [27]. These approaches combine the strengths of different methodologies for comprehensive gait assessment [28].

Algorithm 1 Gait Pattern recognition

Input: $I(x, y)$ Silhouette gait images of each individual
Output: Trained CNN models with accuracy of 98.75%
Initialization:

- Grayscale conversion:
$$I_{\text{gray}}(x, y) = 0.299 \cdot R(x, y) + 0.587 \cdot G(x, y) + 0.114 \cdot B(x, y)$$
- Gaussian smoothing:
$$I_{\text{smoothed}} = \text{GaussianBlur}(I_{\text{gray}})$$
- Gradient calculation:
$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}$$
- Non-maximum suppression:
$$G_{\text{suppressed}}(x, y) = \begin{cases} G(x, y), & \text{if } G(x, y) \geq G(x \pm 1, y \pm 1) \\ & \text{and } G(x, y) \geq G(x \pm 1, y \mp 1) \\ 0, & \text{otherwise} \end{cases}$$
- Double thresholding:
$$E(x, y) = \begin{cases} 1, & \text{if } G_{\text{suppressed}}(x, y) > \text{otsu threshold} \\ 0, & \text{if } G_{\text{suppressed}}(x, y) < \text{otsu threshold} \\ \text{undefined}, & \text{otherwise} \end{cases}$$
- Edge tracking:
$$C(x, y) = \begin{cases} 1, & \text{if } E(x, y) \text{ is connected to a strong edge pixel} \\ 0, & \text{otherwise} \end{cases}$$

for each person **do**
    Find cropped average image:
$$I_A(x, y) = \frac{1}{n} \sum_{i=1}^{n} I_i(x, y)$$
    for each frame **do**
        Track pixel movement between frames:
$$I_F = \arg\min_{u,v} |\nabla I_1(x, y) - \nabla I_2(x + u, y + v)|^2$$
        Obtain the merged image by replacing the s channel of optical flow with the average cropped image:
$$I_M(x, y, s) = \begin{cases} I_A(x, y), & \text{if } s = \text{s\_channel} \\ I_F(x, y, s), & \text{otherwise} \end{cases}$$
    end for
end for
Train CNN Model using merged Images:
$$\text{CNN Model} = \text{Train}(I_M(x, y, s))$$

III. Methodology

We trained CNN models to detect individual cropped average images, optical flow, and optical flow paired with cropped average images [29]. Each image is transformed in a series of steps, beginning with grayscale conversion and binary thresholding and ending with contour identification and extraction of the largest contour. The moments of the greatest contour are computed, and centroid coordinates are obtained. All photos are normalized after being cropped from the contour bounding boxes. The bounding box is separated into three distinct areas: the head, the body, and the legs. The average image is calculated by taking the mean of the pixel values across all image arrays.

We investigated each person's optical flow using a set of photographs to analyze and quantify object motion across pictures or video frames. Tracking pixel movement between frames can reveal the direction of motion in a scene. Optical flow is used in gait analysis to examine movement and dynamic patterns within image sequences [29], [30].

We used derived features such as absolute energy, mean, sample entropy, standard deviation, and skewness for each individual's head, torso, and legs in our time series analysis as these features showed large variations among all individuals. The optical flow of the video was approximated using grayscale and HSV clipped pictures along with the bounding boxes. The tsfresh library was used to extract features, with a focus on time series coordinates for the head, torso, and legs. The retrieved features were used for the training of an artificial neural network (ANN) model.

A. Motion cues

The direction and intensity of body movement at the subsequent step are depicted by the optical flow of individual pixels. Calculating the optical flow matrix [30] assumes small movement and steady brightness.

$$I_x + I_y + I_t = 0 \tag{1}$$

where the picture gradients that can be generated with the Sobel filter are represented by the variables Ix in vertical and Iy in horizontal directions. Following that, an HSV image is created from the optical flow using the formula [30]

$$H_{ij} = tan^{-1}(u_{ij}/v_{ij}) \tag{2}$$

where Hij is the pixel's hue and uij and vij are the dynamic components in the vertical and horizontal directions of the pixel, for a given location (i, j).

$$S_{ij} = \sqrt{(u_{ij}^2 + v_{ij}^2)} \tag{3}$$

where S = Sij is map of the magnitude and Sij denotes pixel's magnitude.

$$V_{ij} = \frac{S_{ij}}{max(S) - min(S)} \tag{4}$$

where the pixel value is denoted by Vij. The cropped average image I2 is combined with the HSV image I1, as illustrated in Figure 1. After trnsforming the source image to grayscale image, the Canny operator is used to identify edges or the character image's border in order to produce I2. After cropping the picture, we discovered the average cropped picture. We define the gait optical flow image model I3 [30] as

$$I_3 = \sigma I_1 + (1 - \sigma) I_2 \tag{5}$$

where the intensity factor is denoted by σ.

IV. Experiment

A. Dataset

We have used CASIA dataset (Dataset A) [8], [31] that features data from 20 individuals. The original dataset contains the images from 3 different angles for each individual: 0, 90 and 180 degree. We have only analysed the 0 degree silhouette images. Each individual contributed to four sequences of images. The sequences vary in length (ranging from 37 to 127 frames) to account for the inherent differences in walking speeds among the subjects. The modified dataset comprises 6400 silhouette images, ensuring diversity for a comprehensive exploration of gait dynamics.

The dataset offers a detailed insight into human gait under various conditions. By capturing the walking patterns at different speeds and differences in clothes, the dataset allows for a detailed exploration of gait analysis [8].

B. Implementation details

We assessed the ML models such as CNN, ANN, SVM, and DT using a stratified 5-fold cross validation technique. Thus, three folds for the training data, one fold as a validation set for determining the ideal architecture and hyper-parameters, and the final fold as a test set. In ANN, widely used activation functions such as rectified linear units (ReLUs) or hyperbolic tangent functions, induce nonlinearity into the network layers. While these activation functions are useful for a variety of purposes, they may not have intrinsic self-normalizing qualities. To address the issue of vanishing or ballooning gradients, ANN requires extra strategies like as batch normalization or cautious weight initialization. Because of the good performance of typical ML methods like SVM and DT in classifying cerebral palsy related patterns of human gaits, we used them as baseline approaches. Further, We chose CNN because they produce decent results with images as compared to the other methods used previously [32].

We assessed the models using a stratified 5-fold cross validation technique. Thus, three folds facilitated training data, one fold as a validation set for determining the ideal architecture and hyper-parameters, and the final fold as a test set.

We constructed a CNN model with a convolutional base and fully connected layers. We used a pretrained VGG19 [33] model and loaded it without its top classification layer. We added additional layers to this base model, including flattening, a dense layer containing 256 units activated by Rectified Linear Units (ReLU), and an output layer with a softmax activation function. The neuron's count in the final output layer was chosen to match the dataset's classification requirements, which included 20 different classes. The training data

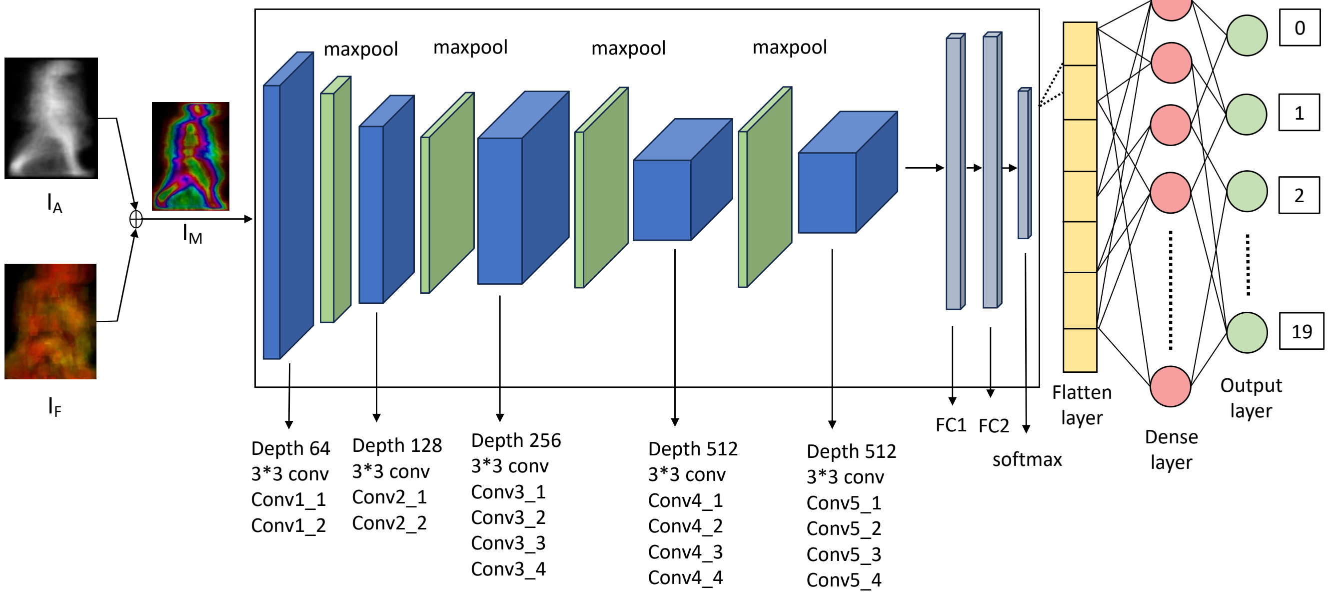
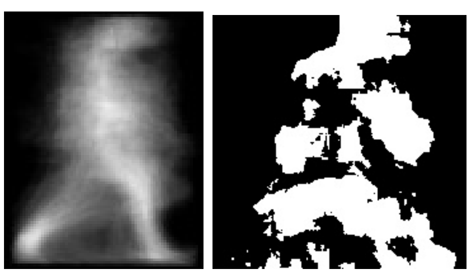


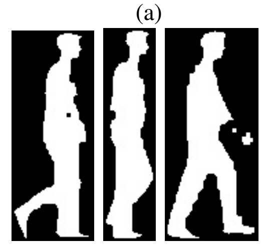
Fig. 3: Illustration of CNN model architecture with input as a fusion of averaged and optical flow images which outputs the class of the person the input gait belongs to.



(a) (b)

Fig. 4: a) cropped average image required for in-depth image analysis and used in CNN model, b) optical flow visualization

(a)



(b)

Fig. 5: Samples of silhouette images from CASIA dataset [31]

consisted of photos scaled to 224x224 pixels. We developed a batch generation technique with a size of 32 to allow for efficient model training by iteratively feeding labelled images into the network. The architecture consists of a vgg19 input layer of input shape same as that of the gait image and output shape (224,224,3). It then consists of a functional layer of input shape (224,224,3) and output shape (7,7,512). Then, it follows a flatten layer of output shape (25088) leading to a dense layer of output shape (256). This layer leads to the final result output layer of size (20).

Throughout the training phase, our model continually performed well, with an accuracy of 98.75 percent as shown in Fig. 8. The convolutional layers extract hierarchical characteristics from input images with precision, while the fully connected layers operate as classifiers, predicting class labels based on the learned features. The goal of this model architecture is to extract abstract features from incoming data and use them to make accurate predictions, resulting in image classification into one of 20 predetermined classes.

Furthermore, for experimental purposes, we trained the model three times using various data representations. Initially,

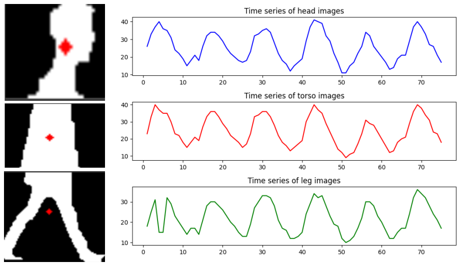


Fig. 6: Illustration of each gait by segmenting it into head, torso and legs and computing each of its centroid coordinates. The derived features of these segments in time series helped to classify the individuals with greater accuracy.

we trained it using optical flow data from the photos. Next, we trained the model on cropped average photos. Surprisingly, the highest accuracy was obtained when we trained the model on a dataset obtained after merging optical flow images and clipped average images. Through these iterations, we aimed to investigate performance variances and determine the most effective technique for our categorization problem.

Additionally, an ANN [34] with three dense layers, two of which were activated using ReLU and the third using the softmax function, was trained using the extracted features. The ANN was constructed using the Adam optimizer and categorical cross-entropy loss function. The classification yielded an accuracy of 94.10% on the test dataset.

To carry out a classification task utilizing a given set of characteristics and labels, we developed a DT classifier model. Here, different classes are used to express the labels. Because DT are nonparametric, nonlinear, and naturally interpretable, they are often used for gait classification [35]. On the testing subset, the model achieved an accuracy score of 44%. Since DT alone can be prone to smaller changes in the input, we further used SVM [26] as well. SVM offer a non-parametric approach for various classification tasks, and their appeal lies in their ability to handle non-linear relationships in the data while providing inherent interpretability. Upon evaluation, the SVM classifier achieved an accuracy of 22 percent on the test dataset which may not be notably high in this instance, but it serves as a valuable metric for assessing the effectiveness of the SVM model.

C. Results

The CNN model was initially trained separately on cropped average image for each gait giving an accuracy of 70 % and optical flow of each gait giving an accuracy of 82.5 %. We achieved the highest accuracy of 98.75% for the CNN model(Table 1), which used the fusion of optical flow and cropped average picture data.

Training metrics is depicted in Fig. 10 where, avg, opfl, merged denotes the training inputs and they are average crop images, optical flow images and fusion of averaged and optical flow of images respectively.

TABLE I: Performance comparison of relevant state of the art models with our proposed model

Model	Time taken	Performance accuracy(%)
kececi et al.(ANN) [34]	10 ms	94.15
Kim et al.(SVM) [26]	08 ms	22.22
Yan et al.(DT) [8]	13 ms	44.44
Karen et al. (Two Stream) [36]	113 s	69
CNN - cropped average image	53 s	70
CNN - optical flow	105 s	82.5
CNN - merged(**ours**)	52 s	98.75

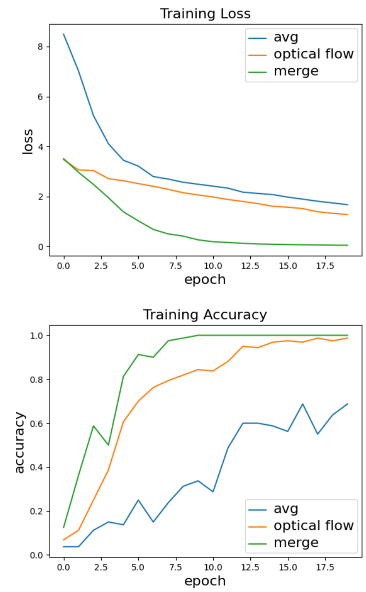


Fig. 7: Training metrics of the model using three different input images- average image, optical flow visualization, and merge of average image and optical flow image for training

1) Comparative analysis using two stream model: For video classification, Karen et al. presented two stream models with competitive performance [36]. We built a two-stream model that receives two distinct inputs in two different channels: a cropped average image and the optical flow visualization. It is trained independently on the previously trained VGG19 model before combining the two channels' flatten layers for final classification.

The results are then concatenated to train the hidden layer and get the final accuracy. We obtained an accuracy of 69% which is quite less than the one-stream model where we had used merged image as input to the model, merging the average cropped image in the s channel of the optical flow image, and then evaluated the model as shown in Fig. 8.

The time taken to run two stream model is 115 seconds, which is more than twice the time taken by one stream network. This is due to operating the same VGG19 model twice on two different input images and concatenating the results, rather than merging the 2 images and using it as a single input which is what has been processed in the one stream network.

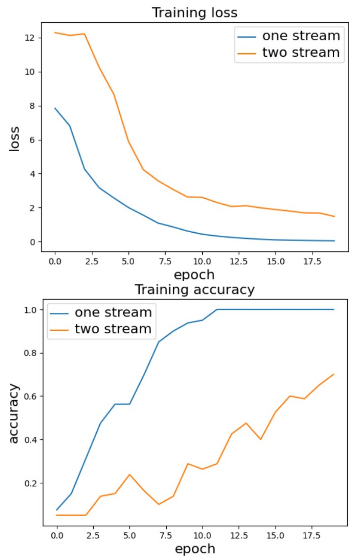


Fig. 8: Training metrics of a two stream model using two input images - average cropped image and optical flow visualization, and concatenation of outputs in the hidden layer to produce final result

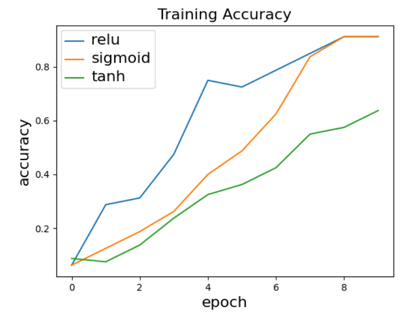
D. Ablation studies

We conducted a set of ablation experiments to analyze the impact of different features in our gait recognition model. The purpose of these tests were to determine how each one of them affected the model's overall performance.

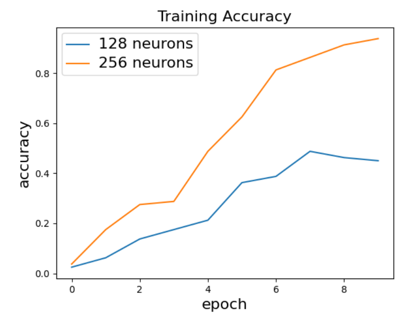
1) Assessment of activation functions:: ReLU Activation: Using the ReLU activation function, the model's accuracy was 98.75%. ReLU's non-saturating characteristics have demonstrated to be particularly beneficial in accelerating convergence during training.
Sigmoid Activation: The model attained a 94.99% accuracy rate by employing the sigmoid activation function. However, the model's performance was somewhat worse than ReLU because of its vanishing gradient issue.
Tanh Activation: Using the tanh activation function caused a considerable drop in the model's performance, to 73.75%.
Tanh results in same problems as sigmoid, which causes delayed convergence and gradient saturation, both of which have a negative impact on the accuracy of the model.

2) Ablation of fully connected layers:: We experimented using only 128 neurons instead of 256 in the hidden layers. The accuracy of the model dropped to 75% when the number of neurons were reduced from 256 to 128. A loss of representational power as a result of this capacity reduction led to a decline in performance.

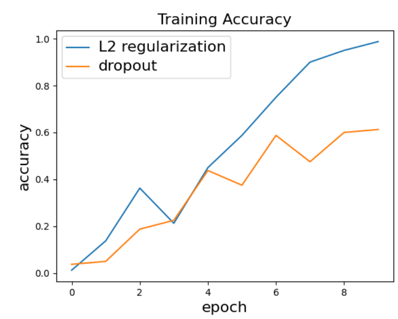
3) Assessment of regularization techniques:: We evaluated how regularization strategies worked to increase the model's capacity for generalization:
L2 Regularization: Using L2 regularization, the model attained an accuracy of 98.75% which was comparable to its performance in the absence of regularization. This suggests that L2 regularization has minimal impact on the model's functionality in this particular situation.
Dropout Layer: The accuracy dropped to 89.99% when a dropout layer was added. Dropout works to prevent over-fitting by randomly removing neurons during training, but as this result shows, high dropout rates can also cause under-fitting.
1



(a)



(b)



(c)

Fig. 9: Training metrics of the model when trained with (a) different activation functions, (b) different number of neurons, (c) L2 regularization or addition of a dropout layer

V. Conclusion and Future Work

Our methodology emphasises the need of using numerous categorization methods to achieve robust and precise gait detection, which improves the model's reliability. We have experimented on various techniques involving extracting optical flow, merging channels and usage of double stream to understand the gait pattern of each individual and classify them accurately. Our results show that our single stream model

TABLE II: Comparison of accuracy and time taken for various activation function, number of neuron in penultimate layer and regularization

Modifications	Model	Accuracy	Time taken
Activation function	Tanh function	73.75%	53 s
	Sigmoid function	94.99%	53 s
	ReLU function	98.75%	52 s
Penultimate layer	Number of neurons 128	75%	53 s
	Number of neurons 256	98.75%	52 s
Regularization	L2 regularization	98.75%	52 s
	Dropout layer	89.99%	53 s

outperforms the two stream model in both accuracy and time complexity, The future of gait analysis envisions a greater use of Machine Learning for improved accuracy and real-time applications in healthcare and human-computer interaction.

Presently, we are classifying the dataset based on images captured from 0 degree angle. Our future plans involve expanding this approach to include the dataset captured at 45 and 90 degrees for each individual to enhance accuracy. Furthermore, We aim to leverage recurrent neural network (RNN) to study the temporal sequence of each class giving scope for better reliability for real life based applications.

REFERENCES

[1] Michael W Whittle. *Gait analysis: an introduction*. Butterworth-Heinemann, 2014.

[2] Henry G Chambers and David H Sutherland. A practical guide to gait analysis. *JAAOS-Journal of the American Academy of Orthopaedic Surgeons*, 10(3):222–231, 2002.

[3] Christopher Kirtley. *Clinical gait analysis: theory and practice*. Elsevier Health Sciences, 2006.

[4] X. Gu, F. Deligianni, B. Lo, W. Chen, and G.Z. Yang. Markerless gait analysis based on a single rgb camera. In *2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages 42–45, 2018.

[5] Amandine Dubois and Francois Charpillet. A gait analysis method based on a depth camera for fall prevention. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4515–4518, 2014.

[6] Ellen Buckley, Claudia Mazzà, and Alisdair McNeill. A systematic review of the gait characteristics associated with cerebellar ataxia. *Gait Posture*, 60:154–163, 2018.

[7] Richard Baker, Alberto Esquenazi, Maria Grazia Benedetti, and Kaat Desloovere. Gait analysis: Clinical facts. *European journal of physical and rehabilitation medicine*, 52:560–574, 08 2016.

[8] Tieniu, Tan, and NLPR. Casia-a dataset for gait analysis. 2001. Accessed: 2023-09-02.

[9] Massimiliano Mangone, Enrico Marinelli, Gabriele Santilli, N Finanore, Francesco Agostini, V Santilli, Andrea Bernetti, Marco Paoloni, and Simona Zaami. Gait analysis advancements: rehabilitation value and new perspectives from forensic application. *European review for medical and pharmacological sciences*, 27:3–12, 01 2023.

[10] Anubha Parashar, Apoorva Parashar, Weiping Ding, Mohammad Shabaz, and Imad Rida. Data preprocessing and feature selection techniques in gait recognition: A comparative study of machine learning and deep learning approaches. *Pattern Recognition Letters*, 172:65–73, 2023.

[11] Saravanan .M, Atluri Raam Praneeth, and Santhosh Babu. Gait recognition for security and surveillance applications. 06 2020.

[12] Guglielmo Cola, Alessio Vecchio, and Marco Avvenuti. Continuous authentication through gait analysis on a wrist-worn device. *Pervasive and Mobile Computing*, 78:101483, 2021.

[13] Imen Chtourou, Emna Fendri, and Mohamed Hammami. Person re-identification based on gait via part view transformation model under variable covariate conditions. *Journal of Visual Communication and Image Representation*, 77:103093, 2021.

[14] Anubha Parashar, Rajveer Singh Shekhawat, Weiping Ding, and Imad Rida. Intra-class variations with deep learning-based gait analysis: A comprehensive survey of covariates and methods. *Neurocomputing*, 505:315–338, 2022.

[15] Ben Crabbe, Adeline Paiement, Sion Hannuna, and Majid Mirmehdi. Skeleton-free body pose estimation from depth images for movement analysis. In *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, pages 312–320, 2015.

[16] Chandra Prakash, Anshul Mittal, Rajesh Kumar, and Namita Mittal. Identification of gait parameters from silhouette images. In *2015 Eighth International Conference on Contemporary Computing (IC3)*, pages 190–195, 2015.

[17] Charissa Ann Ronao and Sung-Bae Cho. Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Systems with Applications*, 59:235–244, 2016.

[18] Abdul Saboor, Triin Kask, Alar Kuusik, Muhammad Mahtab Alam, Yannick Le Moullec, Imran Khan Niazi, Ahmed Zoha, and Rizwan Ahmad. Latest research trends in gait analysis using wearable sensors and machine learning: A systematic review. *Ieee Access*, 8:167830–167864, 2020.

[19] Chia-Yeh Hsieh, Wan-Ting Shi, Hsiang-Yun Huang, Kai-Chun Liu, Steen J Hsu, and Chia-Tai Chan. Machine learning-based fall characteristics monitoring system for strategic plan of falls prevention. In *2018 IEEE International Conference on Applied System Invention (ICASI)*, pages 818–821. IEEE, 2018.

[20] Reed D Gurchiek, Rebecca H Choquette, Bruce D Beynnon, James R Slauterbeck, Timothy W Tourville, Michael J Toth, and Ryan S McGinnis. Remote gait analysis using wearable sensors detects asymmetric gait patterns in patients recovering from acl reconstruction. In *2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages 1–4. IEEE, 2019.

[21] Maria Grazia Benedetti, Ettore Beghi, Antonio De Tanti, Aurelio Cappozzo, Nino Basaglia, Andrea Giovanni Cutti, Andrea Cereatti, Rita Stagni, Federica Verdini, Mario Manca, Silvia Fantozzi, Claudia Mazzà, Valentina Camomilla, Isabella Campanini, Anna Castagna, Lorenzo Cavazzuti, Martina Del Maestro, Ugo Della Croce, Marco Gasperi, Tommaso Leo, Pia Marchi, Maurizio Petrarca, Luigi Piccinini, Marco Rabuffetti, Andrea Ravaschio, Zimi Sawacha, Fabiola Spolaor, Luigi Tesio, Giuseppe Vannozzi, Isabella Visintin, and Maurizio Ferrarin. Siamoc position paper on gait analysis in clinical practice: General requirements, methods and appropriateness. results of an italian consensus conference. *Gait Posture*, 58:252–260, 2017.

[22] Weijun Tao, Tao Liu, Rencheng Zheng, and Hutian Feng. Gait analysis using wearable sensors. *Sensors (Basel, Switzerland)*, 12:2255–83, 12 2012.

[23] Odysseas Stavrakakis, Athanasios Mastrogeorgiou, Aikaterini Smyrli, and Evangelos Papadopoulos. Gait analysis with trinocular computer vision using deep learning. In *2023 IEEE International Conference on Image Processing Challenges and Workshops (ICIPCW)*, pages 3702–3706, 2023.

[24] Shaoxiong Zhang, Yunhong Wang, and Annan Li. Cross-view gait recognition with deep universal linear embeddings. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9091–9100, 2021.

[25] Huanghe Zhang, Yi Guo, and Damiano Zanotto. Accurate ambulatory gait analysis in walking and running using machine learning models. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(1):191–202, 2020.

[26] Wonjin Kim and Yanggon Kim. Abnormal gait recognition based on integrated gait features in machine learning. In *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, pages 1683–1688, 2021.

[27] Jie Kong, Yonghui Xu, and Han Yu. Deep transfer learning for abnormality detection. pages 233–237, 10 2019.

[28] Ziming Yin, Yi Jiang, Jianli Zheng, and Hongliu Yu. Stja-gcn: A multi-branch spatial–temporal joint attention graph convolutional network for abnormal gait recognition. *Applied Sciences*, 13:4205, 03 2023.

[29] Anurag Ranjan and Michael Black. Optical flow estimation using a spatial pyramid network. pages 2720–2729, 07 2017.

[30] Hongyi Ye, Tanfeng Sun, and Ke Xu. Gait recognition based on gait optical flow network with inherent feature pyramid. *Applied Sciences*, 13:10975, 10 2023.

[31] Shuai Zheng, Junge Zhang, Kaiqi Huang, Ran He, and Tieniu Tan.

Robust view transformation model for gait recognition. pages 2073–2076, 09 2011.
[32] Sai Chaganti, Ipseeta Nanda, Koteswara Pandi, Tavva Prudhvith, and Niraj Kumar. Image classification using svm and cnn. pages 1–5, 03 2020.
[33] Anuradha Khattar and Syed Quadri. “generalization of convolutional network to domain adaptation network for classification of disaster images on twitter”. *Multimedia Tools and Applications*, 81, 09 2022.
[34] Aybuke Kececi, Armağan Yildirak, Kaan Ozyazici, Gulsen Ayluctarhan, Onur Agbulut, and Ibrahim Zincir. Implementation of machine learning algorithms for gait recognition. *Engineering Science and Technology, an International Journal*, 23(4):931–937, 2020.
[35] Song-Hua Yan, Yan-Cheng Liu, Wei Li, and Kuan Zhang. Gait phase detection by using a portable system and artificial neural network. *Medicine in Novel Technology and Devices*, 12:100092, 2021.
[36] Karen Simonyan and Andrew Zisserman. Two-stream convolutional networks for action recognition in videos, 2014.