Preprint of paper from ICONS 2024, Arlington, VA, July-August 2024.

Realtime Person Identification via Gait Analysis using IMU Sensors on Edge Devices

Shanmuga Venkatachalam, Harideep Nair, Prabhu Vellaisamy, Yongqi Zhou, Ziad Youssfi, and John Paul Shen
Department of Electrical & Computer Engineering, Carnegie Mellon University
{*shanmugv, hpnair, pvellais, yongqiz2, zyoussfi, jpshen*}*@andrew.cmu.edu*

***Abstract*—All individuals have a unique gait signature, or walking style, that can serve as their biometric identifier. While recent research has shown that deep neural networks can perform effective gait recognition, these studies often prioritize classification accuracy with little regard to model complexity and hardware efficiency. In this paper, an efficient four-layered convolutional neural network (CNN) model is proposed for real-time gait-based person identification targeting edge devices such as wearables, microcontrollers, and neuromorphic chips. This model is trained on a public gait dataset augmented with custom collected data, totaling to 24 classes. Despite its relatively small size (236 KB), the model achieves a high accuracy of 96.7%. It delivers fast inference time of 70 ms requiring only 5 KB of RAM and consumes just 125 mW power during continuous inference on an Arduino Nano 33 BLE Sense. Model efficacy is further demonstrated by converting proposed CNN to a spiking neural network (SNN) and running realtime inference on BrainChip's neuromorphic platform, Akida SoC, within 72 mW power.**

***Index Terms*—Human Gait, Biometric Identification, Inertial Sensors, Arduino, Neuromorphic Akida, Edge AI**

I. Introduction and Background

Mobile phones and wearable edge devices have become an integral part of our everyday lives. Sensitive private information and data are increasingly being stored in these devices. Therefore, it is essential to safeguard these data in a trusted manner to prevent breaches from unauthorized and malicious threats. While biometric features are utilized for identification, including passcodes, facial recognition, and fingerprints, they can be compromised [1] and are only presented as an initial authentication. Unauthorized users can gain access to these devices if the authentication is not constantly and dynamically performing identification, which, with traditional authentication, can either be expensive or impractical.

Gait analysis is seen as a promising alternative to traditional authentication methods. Gait is the distinct walking signature pertaining to an individual, carrying inherent characteristics that can be leveraged for non-intrusive person identification. Compared to traditional and static biometrics, gait analysis taps into the dynamic and behavioral aspects of an individual's movement. Each person has a distinct gait, influenced by factors like anatomy, musculoskeletal structure, and personal habits. Previous studies have demonstrated that the gait signature of an individual cannot be replicated or copied [2]. With all smartphones and wearable devices containing inertial measurement unit (IMU) sensors, it is possible to dynamically profile a person's gait for more secure and personal authentication. Traditionally, machine learning algorithms have been used for gait identification; deep learning techniques have been employed to great effect [3] more recently.

Existing gait identification works predominantly focus on model accuracy rather than model efficiency [4]–[8]. Hardware efficiency becomes a metric of paramount importance for edge-based gait recognition on devices such as wearables, microcontrollers, smartphones, and neuromorphic chips. Akida System-on-Chip (SoC) [9] is a neuromorphic processor developed by BrainChip targeting efficient edge inference with online learning. MetaTF [10] is their Machine Learning framework facilitating the training, testing, and deployment of neural networks on Akida. It contains four main components - (i) a model zoo containing TensorFlow/Keras-defined models that can be quantized and are compatible with Akida, (ii) a framework for quantizing deep learning (DL) models, (iii) a conversion tool to convert DL models to neuromorphic spiking neural network (SNN) models compatible with Akida, and (iv) an interface to the Akida processor.

In this paper, a streamlined four-layer convolutional neural network (CNN) model is proposed that is both lightweight and highly efficient, fit for edge and mobile deployment for dynamic real-time gait analysis. The model is trained on a version of the public dataset whuGAIT [11], supplemented with additional user-based gait data belonging to the authors to demonstrate real-time inference demonstration. The Arduino Nano 33 BLE Sense board is selected as the vehicle platform for evaluation and inference. Additionally, the four-layer CNN model is converted to an analogous event-based spiking neural network using the Brainchip MetaTF framework to obtain power and latency measurements on the Brainchip Akida board. The highlights and contributions of this work include:

- A light-weight four-layered CNN model for gait analysis is presented that performs close to prior state-of-the-art CNN models, while supporting four additional classes.
- Augmentation of a public dataset with user-based gait data is performed to report live inference classification, demonstrating quick re-training for the model.
- Inference results through deployment of the four-layered CNN model on an Arduino Nano 33 BLE Sense board are analyzed and discussed. The model only consumes 5 KB RAM and 236 KB flash storage.
- Further, an event-based SNN analogous to the proposed CNN model is generated and deployed on the Brainchip Akida neuromorphic processor, and the resultant latency and power metrics are compared to see the efficacy of

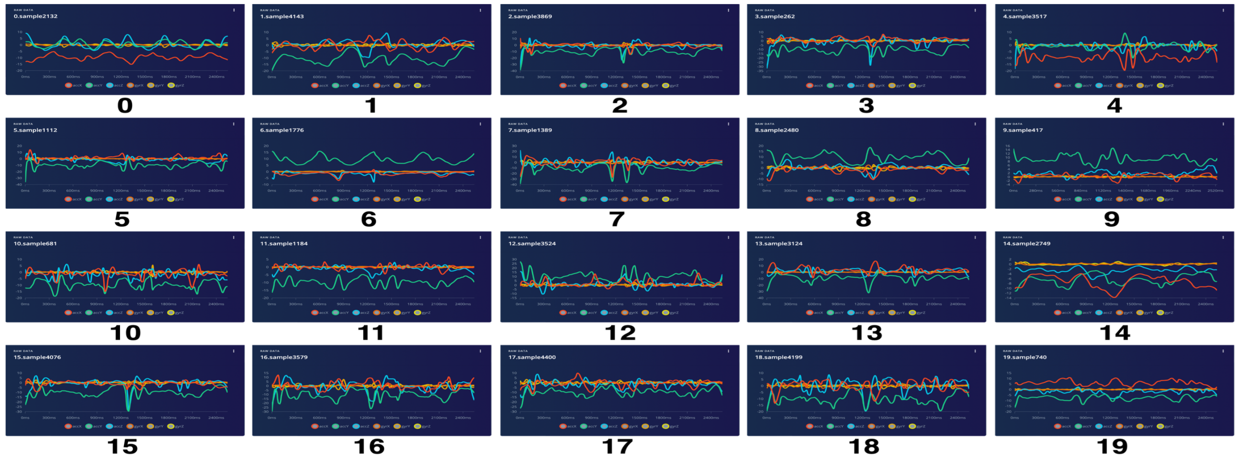

Fig. 1. Raw data samples from each of the 20 classes of whuGAIT. Each sample has 128 length with 3-axis accelerometer and 3-axis gyroscope data.

event-based SNN models for gait analysis.

The paper is structured as follows. Section II discusses the methods used in this work, including dataset preparation, model architecture, training, and inference pipeline. In Section III, model deployment on various platforms and the resulting inference are reported and analysed, comparing against previous works. Finally, Section IV summarizes the results and findings of the study and discusses future research directions.

II. METHODS

This section details the methods and experimental setup used in this work. First, the publicly available dataset used is discussed, followed by our custom data collection procedure for data augmentation. We then describe our proposed model architecture, training methodology and the edge inferencing pipeline on two hardware devices, namely Arduino Nano 33 BLE Sense and neuromorphic Akida SoC from BrainChip.

A. whuGAIT Dataset

The primary dataset used is whuGAIT [11]. The overall data collection process involves a total of 118 subjects, out of which 20 subjects gathered thousands of samples each over a span of two days. The remaining 98 subjects undertook a more concise data collection spanning one day resulting in hundreds of samples each. Each data sample comprises both 3-axis accelerometer and 3-axis gyroscope data, all recorded at a uniform sampling rate of 50 Hz.

Among the various sub-datasets of whuGAIT, dataset#2 is chosen for this work since it utilizes only the first set of 20 subjects as opposed to all 118, keeping the total number of classes to 20. It consists of 49,275 samples, of which 44,339 samples are used for training and the rest 4,936 for testing. Each raw data sample has a length of 128. One example of raw data from each class is shown in Figure 1.

B. Custom Data Collection

Further, we augment the whuGAIT dataset with custom collected data for four new classes. The custom data not only

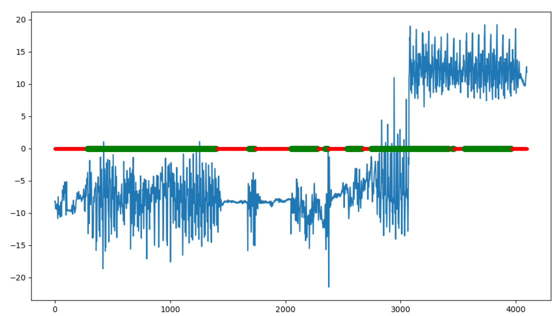

Fig. 2. Automatic segmentation of raw gait data using DCNN to extract walking periods.

illustrates the model's capability to learn additional classes but also enables live demonstration as will be discussed later in Section III. The data was collected for walking at multiple paces back and forth, on carpeted and non-carpeted floors for better generalization. Our final model uses spectral feature extraction on this raw data for ease of live deployment and demonstration. We have also created a custom data preprocessing pipeline as explained next.

Our initial implementation involves manually splitting the collected data into 3-second gait segments using the visualization tool provided by Edge Impulse (EI) [12]. However, this manual methodology is inconvenient for collecting large amounts of data. To facilitate automated preprocessing, we use a one-dimensional DCNN [11] model to extract walking period based on its semantic difference with non-walking (e.g., standing, running) period. The result is shown in Figure 2. The blue zigzag line represents the collected activity data, which includes walking, stopping, and random movement in different directions and slopes. The green part of the straight line indicates a walking data segment, while the red part indicates a noise period that should be discarded. The extracted walking data is then segmented into fixed-length segments to

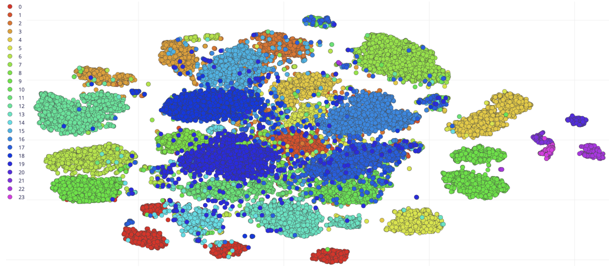

Fig. 3. UMap Data Visualization

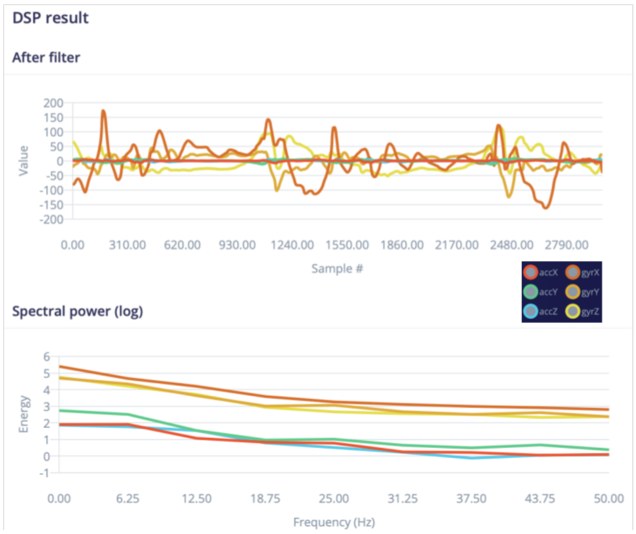

Fig. 4. FFT Features

align them with other data in the training set.

C. Training Methodology

1) Feature Extraction: We use spectral analysis preprocessing block in Edge Impulse to extract spectral features from raw data samples. The corresponding UMap separation diagram visualizing the clusters for all 24 classes are shown in Figure 3, with visible separation between most clusters. FFT analysis with FFT length of 16 on a window size of 3 seconds is used. Figure 4 illustrates the result after filtering as well as logarithmic spectral power for an example sample. An interesting observation from Figure 4 is that the gyroscope data carries more energy than the accelerometer data. This implies that collected gait data has richer features in angular momentum relative to acceleration. This observation is consistent for all data collected with Arduino. Future investigations can explore equalizing the energy between accelerometer and gyroscope for more uniform integration of signals.

2) Model Architecture: Proposed model (Figure 5) consists of four layers, where the first layer is a 2D Convolution layer with 32 output channels and 3x3 filters. The input to Convolution layer consists of a vector of 78 FFT-applied features which is reshaped to 13x6. The Convolution layer is followed by flatten, and subsequently 3 dense layers with 256, 128 and 32 neurons respectively. The final layer is a softmax layer with 24 classes. This network was chosen after extensive hyperparameter tuning resulting in the best tradeoff between accuracy performance and model complexity.

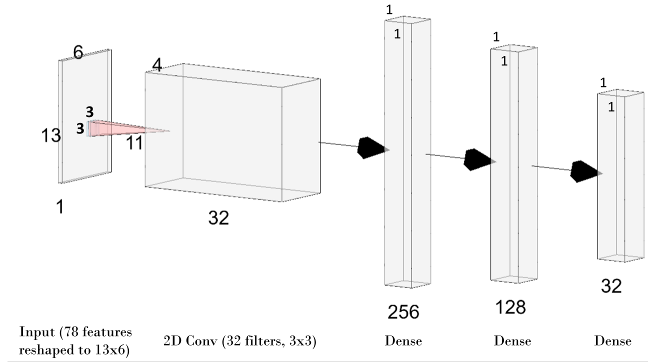

Fig. 5. Proposed four-layer CNN architecture with one convolutional layer and three fully-connected (dense) layers. Input vector with 78 features is reshaped to 13x6 before first layer and output layer has 24 classes.

TABLE I
INFERENCE ACCURACY COMPARISON ON WHUGAIT DATASET#2

Model	Inference Accuracy
CNN+LSTM Baseline [11]	97.33%
CNN+Attention Baseline [13]	97.67%
Proposed CNN	**96.23%**

3) Model Hyperparameters: The model is trained for 20 epochs with 0.0005 learning rate. During training, 20% of the training set is used for validation. Batch size is set to 32.

D. Inferencing Pipeline

1) Arduino Nano 33 BLE Sense: Edge Impulse is used to convert the trained model from FP32 to INT8 precision via quantization-aware training. Quantized model is then downloaded and deployed on to Arduino via runtime scripts that orchestrate timed loops for inference and latency measurement.

2) BrainChip Akida: BrainChip's MetaTF [10] software development platform is used to convert the pretrained CNN to an equivalent spiking neural network (SNN), with minimal impact on model accuracy. Two BrainChip Akida v1 boards [9] installed on the server (connected to computers through PCIe) are utilized for deployment of the converted SNN model.

III. Deployment Results

The trained model is deployed on two edge devices, Arduino Nano 33 BLE Sense, and BrainChip Akida. As an extension to a third edge inferencing platform, model deployment is also performed on a mobile smartphone. This section details the corresponding model accuracy and hardware metrics.

A. Inference Accuracy

The trained model achieves 96.3% validation accuracy and 96.23% testing accuracy with almost perfect confusion matrix. As shown in Table I, the proposed small CNN model is able to achieve performance very close to more complex baseline

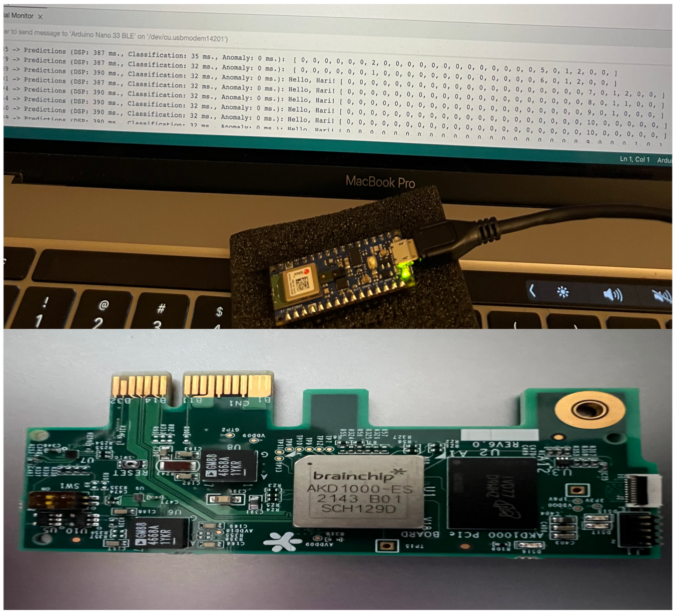

Fig. 6. Top: Arduino inference setup with laptop screen showing the output recognizing one team member via their gait (Arduino was placed inside pant pocket while walking). Bottom: BrainChip Akida PCIe board used for neuromorphic inference.

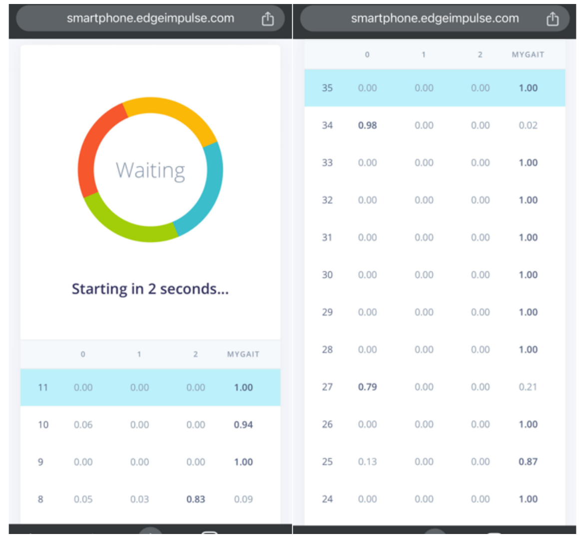

Fig. 7. Mobile Phone Deployment

models - CNN+LSTM [11] and CNN+Attention [13]. It is to be noted that proposed model accounts for classification of four additional classes belonging to the authors along with the original 20 classes from whuGAIT.

B. Deployment on Arduino Nano 33 BLE Sense

During live deployment on Arduino (Figure 6 top), we successfully demonstrated accurate prediction of each of the four team members from live gait. We observed a 2 second delay from the onset of walking to the generation of appropriate predictions from the model. This overhead could be related to the prediction smoothing function within the Arduino deployment code and is a topic for future investigation and improvement. However, once it starts generating predictions, the inference latency is only about 70 ms with 5 KB active RAM usage. Further, its power consumption is measured using a power jive to be 125 mW (25 mA current at 5 V). The results underscore the feasibility of a lightweight model capable enough to classify 24 classes that can easily fit into a very small form factor.

C. Deployment on BrainChip Akida

The converted SNN is first mapped to BrainChip Akida processor (Figure 6 bottom) through MetaTF's "map" function defined in the Akida python package. The Akida board comes with power measurement utilities on device that can be probed through python functions defined in the Akida package. Through these utilities, the SNN model is measured to consume about 72 mW with an average frame rate of 24.82 fps during inference, with model accuracy of 96.12%. The inference energy consumed is 2.9 mJ/frame. This demonstrates the potential for huge power and energy efficiency of proposed CNN model and its amenability to neuromorphic platforms while delivering near state-of-the-art inference accuracy.

D. Deployment on Mobile Phone

To simulate real-world application scenarios, further experimentation is conducted on a mobile device, an iPhone 13, using Edge Impulse platform. The EI platform has accessible mobile motion sensor only for a three-axis accelerometer, but is unable to retrieve the gyroscopic data. Thus the input features of the model had to be adjusted from six to three. Due to differences in coordinate systems, precision, and amplification between sensors on mobile devices and those on Arduino, the model deployed on phone requires training with sensor data collected from the phone (following the same procedures in Section II). As shown in Figure 7, label *MyGait* denotes the gait of the phone owner; label *0* represents a stationary pattern; labels *1* and *2* are others' gaits from the dataset. Once the mobile device owner starts walking steadily, the model is capable of identifying the identity based on gait.

IV. Conclusion

Our work serves as a feasibility proof for deployment of very efficient yet highly effective lightweight models on to small form factor edge devices. Our model trained on a standard dataset augmented with the team members' gait data is able to achieve 96% accuracy on 24 classes, while consuming only 70 ms inferencing time with 5 KB RAM, and 125 mW power on Arduino. Further, our work serves as a first step towards deploying a gait recognition model on a neuromorphic device such as the BrainChip Akida, incurring just 72 mW power and 2.9 mJ/frame. Future investigation will focus heavily on optimizing the model, adding data for physically challenged and other users for diversification and bias reduction, and building and optimizing the neuromorphic SNN model from scratch without conversion.

REFERENCES

[1] H. Zhang, J. Liu, K. Li, H. Tan, and G. Wang, "Gait learning based authentication for intelligent things," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4450–4459, 2020.

[2] A. Peinado-Contreras and M. Munoz-Organero, "Gait-based identification using deep recurrent neural networks and acceleration patterns," *Sensors*, vol. 20, no. 23, p. 6900, 2020.

[3] A. Sepas-Moghaddam and A. Etemad, "Deep gait recognition: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 45, no. 1, pp. 264–284, 2022.

[4] V. Narayan, S. Awasthi, N. Fatima, M. Faiz, and S. Srivastava, "Deep learning approaches for human gait recognition: A review," in *2023 International Conference on Artificial Intelligence and Smart Communication (AISC)*. IEEE, 2023, pp. 763–768.

[5] C. Filipi Gonçalves dos Santos, D. d. S. Oliveira, L. A. Passos, R. Gonçalves Pires, D. Felipe Silva Santos, L. Pascotti Valem, T. P. Moreira, M. Cleison S. Santana, M. Roder, J. Paulo Papa *et al.*, "Gait recognition based on deep learning: a survey," *ACM Computing Surveys (CSUR)*, vol. 55, no. 2, pp. 1–34, 2022.

[6] C. F. G. d. Santos, D. D. S. Oliveira, L. A. Passos, R. G. Pires, D. F. S. Santos, L. P. Valem, T. P. Moreira, M. C. S. Santana, M. Roder, J. P. Papa *et al.*, "Gait recognition based on deep learning: a survey," *arXiv preprint arXiv:2201.03323*, 2022.

[7] V. Kukreja, D. Kumar, and A. Kaur, "Deep learning in human gait recognition: An overview," in *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*. IEEE, 2021, pp. 9–13.

[8] M. Alotaibi and A. Mahmood, "Improved gait recognition based on specialized deep convolutional neural network," *Computer Vision and Image Understanding*, vol. 164, pp. 103–110, 2017.

[9] B. H. Ltd. Akida neuromorphic system-on-chip. [Online]. Available: https://brainchip.com/akida-neural-processor-soc/

[10] ——. Metatf neuromorphic ml framework. [Online]. Available: https://doc.brainchipinc.com

[11] Q. Zou, Y. Wang, Q. Wang, Y. Zhao, and Q. Li, "Deep learning-based gait recognition using smartphones in the wild," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3197–3212, 2020.

[12] V. Janapa Reddi, A. Elium, S. Hymel, D. Tischler, D. Situnayake, C. Ward, L. Moreau, J. Plunkett, M. Kelcey, M. Baaijens *et al.*, "Edge impulse: An mlops platform for tiny machine learning," *Proceedings of Machine Learning and Systems*, vol. 5, 2023.

[13] H. Huang, P. Zhou, Y. Li, and F. Sun, "A lightweight attention-based cnn model for efficient gait recognition with wearable imu sensors," *Sensors*, vol. 21, no. 8, p. 2866, 2021.