Revolutionary AI-Driven Skeletal Fingerprinting for Remote Individual Identification

Achraf BERRAJAA[1,2], Ayyoub El OUTMANI[3], Issam BERRAJAA[4], Nourddin SAIDOU[5]
Department of Computer Science, Faculty of Sciences, University Mohamed First, Oujda, Morocco[1]
School of Technology, Moulay Ismail University, Meknes, Morocco[2]
Research Center, High Studies of Engineering School, EHEI Oujda, Morocco[3]
Mohammed VI University Hospital, Oujda, Morocco[4]
Euromed Research Center, Euromed University of Fes, Fez, Morocco[5]

***Abstract*—This research aims to devise a distinct mathematical key for individual identification and recognition. This key, represented through signals, is constructed using Lagrange polynomials derived from the skeletal points. Consequently, we present this key as a novel fingerprint categorized within physiological fingerprints. It's crucial to highlight that the primary application of this fingerprint is for remote individual identification, specifically excluding any bodily masking. Subsequently, we implement an artificial intelligence model, specifically a Convolutional Neural Network (CNN), for the automated detection of individuals. The proposed CNN is trained on an extensive dataset comprising 10000 real-world cases and augmented data. Our skeletal fingerprint recognition system demonstrates superior performance compared to other physiological fingerprints, achieving a remarkable 98% accuracy in detecting individuals at a distance.**

Keywords*—*Artificial intelligence; recognition of individuals; new fingerprint; lagrange polynomials; CNN

I. Introduction

Biometrics is a method employed for quantifying human body attributes, encompassing the analysis of physical, biological, and behavioral characteristics of individuals. While traditionally utilized for individual identification, its significance has grown in the realm of countering terrorist and criminal threats, emphasizing authentication.

Biometrics can be broadly classified into two primary categories: physical and behavioral. Physical biometric modalities include features like fingerprints, hand and facial shapes, as well as characteristics such as vein patterns in the hand, the iris, and the ear. On the other hand, behavioral modalities encompass traits like signatures and gait patterns.

Biometric characteristics, fundamentally enduring, remain constant throughout an individual's lifetime, rendering them both unique and universal. This permanence offers a robust alternative to easily forgettable or susceptible means of personal identification, such as PINs or passwords, as well as vulnerable physical items like magnetic cards, which can be subject to theft, duplication, or loss. Automated processes within the realm of biometric systems address these distinctive characteristics, making them highly secure. Biometric systems are acknowledged as the epitome of security owing to their inherent resistance to fraudulent use. Continuous research endeavors are dedicated to exploring novel methods and improving existing ones. The diverse modalities within biometrics are generally classified into three analytical classes based on their nature, as highlighted in [19].

1) Physiological Biometrics: This category encompasses measurements derived from distinct physical attributes, including fingerprints, the unique patterns of the outer ear (oracular print), iris structure, and facial features.
2) Biological Biometrics: This type of measurement primarily relies on biological substances such as DNA, urine, saliva, and blood. Frequently utilized in criminal identification and anti-doping efforts, these measurements provide valuable insights into individual profiles.
3) Behavioral Biometrics: This technique involves measuring characteristics linked to behavior, such as voice recognition, keyboard typing speed, writing styles, and signatures. Unlike physiological biometrics, these traits are less stable, subject to variation with age, and influenced by psychological conditions like stress.

A biometric system relies on physiological, biological, or behavioral techniques to uniquely identify individuals. The landscape of biometric modalities is expansive and constantly evolving, with new methods continually emerging. In the subsequent discussion, we delve into notable research and highlight some prevalent modalities, including facial recognition, speech analysis, fingerprinting, hand structure, and iris scanning. Moreover, we introduce our innovative skeletal fingerprint recognition system, a distinctive addition to physiological fingerprints. It is noteworthy that the robustness of our system, powered by artificial intelligence, lies in its ability to identify individuals at a distance with an impressive accuracy rate of 98%, setting it apart from other physiological fingerprinting methods.

In terms of structure, this paper is organized as follows: Section 2 provides a comprehensive literature review, examining works pertaining to biometric systems, with a particular focus on notable research and highlighting prevalent modalities, such as facial recognition, speech analysis, fingerprinting, hand structure, and iris scanning. Section 3 delves into the intricacies of our data preparation pipeline and outlines the proposed model for our recognition system. Moving on, Section 4 is dedicated to digital experiments, where we present, discuss, and analyze the results obtained, demonstrating the effectiveness of our models. The concluding Section 5 summarizes the study and offers potential perspectives for further enhancing the current results.

II. Related Work

Biometrics involves the automated identification of individuals through their anatomical and/or behavioral characteristics, emphasizing personal identity based on inherent traits rather than possessions or memorized information.

In scholarly discussions, these characteristics are often referred to as identifiers, modalities, indicators, or biometric attributes. It's crucial to highlight that any behavioral or physiological trait can be deemed a biometric modality, provided it satisfies seven key properties: universality, uniqueness, permanence, measurability, performance, acceptability, and bypassing, as outlined by Jain et al. in their work [13].

Numerous distinctive morphological features characterize each individual, and each of these can be measured through various methodologies:

- Fingerprints (finger-scan): Since the 1970s, fingerprint recognition systems have been marketed and gradually attract the attention of researchers and security companies such as the FBI for example. Fingerprints are the traces left by the grooves of the finger pulps. The pattern formed is unique and differs from one individual to another. In practice, it is almost impossible to use all the information provided by this drawing, so we extract the main characteristics such as the bifurcations of ridges. A fingerprint contains on average a hundred characteristic points . Statistically, that it is impossible to find 12 identical points in 2 individuals.
  Among the methods used to process Fingerprints: In [7], the authors present a prototype ultrasonic sensor for detecting fingerprint patterns. Their working principle is based on amplitude measurements at selected points in the sound field of ultrasound waves diffracted by subsurface finger structures. Machine learning methods have also been proposed, such as in [11]. The authors propose an unsupervised approach based on object fingerprinting to detect activity without human labeling. Lately authors of [9] present Molecular Surface Interaction Fingerprinting, a conceptual framework based on geometric deep learning methods for detecting fingerprints important for specific biomolecular interactions.
- Hand geometry: The recognition of the shape of the hand is considered to be the ancestor of biometric technologies. In the sixties, Robert P. Miller filed a patent for a device to measure characteristics of the hand and record them for later comparison. This biometric modality consists in measuring several characteristics of the hand, such as the shape of the hand, the shape of the joints, the length and width of the fingers, etc. Several works have been carried out to extract these characteristics, we cite [20], the authors implemented a biometric system based on hand geometry recognition. Hand features are extracted from color photographs as users place their hands on a platform designed for this type of task. Various pattern recognition techniques have been tested for classification and/or verification. In recent work, an algorithm [3] is proposed, which is based on the geometry properties of the hand as keys for encryption/decryption audio files. In study [4] the authors adopted a solution that relies on the geometry of the hand to reduce mobile payment fraud.
- Iris: It is a reliable technology, and appears to be much more accurate than some other biometric means. This is because our iris has so many characteristics that can vary from one individual to another. The iris is made up of blood vessels and these are arranged differently from one individual to another. Each eye is unique. It is proven that the probability of finding two identical irises is lower than the inverse of the number of humans who have lived on earth. Once the image of the configuration of the blood vessels is obtained by the biometric system, the operation is almost identical to that of the system analyzing the fingerprint. This technique can be affected by several factors such as the distance between the eye and the camera, reflections, false eye detection etc. To reduce the risk of poor recognition, certain dynamic characteristics of the eye are called upon. This is summarized in the following study [2].
- Facial scan: The development of biometric systems based on recognition of the shape of the face is one of the most recent biometric techniques. This technique is based on a face image. The most popular face recognition methods are based on: 1) the shape and the location of facial attributes such as eyebrows, eyes, lips, nose, and chin and their spatial relationships, or 2) general (global) analysis of facial images, representing faces as a weighted combination of representing the number of canonical faces [12].
  In addition to physical characteristics, an individual also has behavioral characteristics that are unique to him :
- keystroke-scan: This is a technique for recognizing people based on their own typing rhythm. It is a "Software Only" biometric solution, because it only consists of a collection of data based on the typing dynamics of users. It is applied to the password which thus becomes much more difficult to imitate. Keystrokes are affected by several factors: time between keystrokes, frequency of errors, or the overall time it takes to type text. However, even if, this behavior is not unique to each individual, it offers sufficient discriminating information to identify an individual. A new study [16] is carried out in this field for deprived attacks, where the authors propose a new hybrid method based on typing dynamics for identification and biometric authentication integrated with artificial intelligence.
- Speech Recognition (voice-scan): The data used by speech recognition comes from both physiological and behavioral factors. Identification by the route is based on the size and the shape of the appendages (nasal cavities, lips and mouth) used in sound synthesis. The physiological characteristics of an individual's voice are invariable, but behavioral characteristics change over time and with age, depending on health conditions (sore throat) and emotional states, etc. which

decreases the accuracy of the identification rate. This identification technique is sensitive to a large number of factors such as noise. As an example of a security system based on voice identification as an access control key, we cite [18].

Other biometric techniques are currently being developed such as biometry by the geometry of the veins of the hand [21], biometry by the palm print [8], biometry by the geometry of the veins of the finger [25]. Despite their reliability, biometric identification systems do not guarantee recognition of the individual. The big concern is that most systems work with contact with the individual, which is not logical to avoid terrorist attacks or for remote identification of the individual. For this, the gait recognition is considered to be the most demanded method of biometrics in this situation.

Gait analysis is the systematic study of human movement using the eyes and brain of an observer, supplemented by instruments that measure body movement, body mechanics, and muscle activity. Gait recognition is a relatively new biometric technique [15], [22] which has attracted more interest in the Computer Vision community in recent years, due to its advantages over other biometrics [24]. Biometric methodologies are generally intrusive and require the collaboration of different methods of biometric in order to perform accurate data acquisition. The process, on the contrary, can be captured remotely and without collaboration of several biometric methodologies. This makes it a discreet method of recognizing people at a distance and without contacting the individual, this is more requested in order to avoid terrorist attacks. In study [10], the authors propose an algorithm to characterize gait using 3-dimensional skeletal information acquired by the Microsoft Kinect sensor. In study [17], the authors propose an self-similarity based gait recognition system for human identification using modified Independent Component Analysis (MICA). In study [23], a method has been proposed for in-depth gait recognition based on the characteristics of the local directional pattern (LDP) to extract information and a neural network for learning.

The primary objective of this research is to pioneer a novel fingerprint within the realm of physiological biometrics, specifically focusing on the parametrization of the human skeleton. Our methodology involves measuring the dimensions of the human skeleton in both static and dynamic states, particularly through gait analysis. These measurements are then mathematically modeled using Lagrange polynomials, which are subsequently translated into electronic signals. Furthermore, we have implemented a Convolutional Neural Network (CNN) to automate the detection of individuals. The proposed neural network is trained using the modeled Lagrange polynomials on a substantial dataset comprising 10 000 real-world cases and augmented data. The distinctive advantage of this approach lies in its capacity to mitigate morphological falsifications, such as fingerprints or facial masks. The combined performance of the proposed model, integrating Lagrange polynomials and CNN, proves highly promising. It enables remote identification, wherein the individual's skeleton can be captured through cameras, allowing for authentication decisions to be made remotely.

III. The Proposed Skeletal Fingerprint Recognition System

A biometric system functions as a pattern recognition system that acquires biometric data from an individual, extracts a set of features from this data, and compares these features with stored signatures in a database. Depending on the application environment, the biometric system can operate in either identification or verification mode. This same principle applies to the recognition of skeletal fingerprints, wherein a provided fingerprint is compared with one or more existing fingerprints stored in the system's biometric database. The system yields a positive result if the skeletal fingerprints match any of the templates, and a negative result otherwise.

Our skeletal fingerprint recognition system comprises five modules: acquisition (capturing the skeleton of the individual using a camera), pretreatments (including grayscale adjustment, normalization, binarization, and skeletonization), feature extraction (utilizing precise skeleton coordinates to form Lagrange polynomials), prediction and comparison (generating signals from each Lagrange polynomial combined with the CNN prediction), and finally, the decision.

The overall structure of our skeletal fingerprint recognition system is shown in Fig. 1.

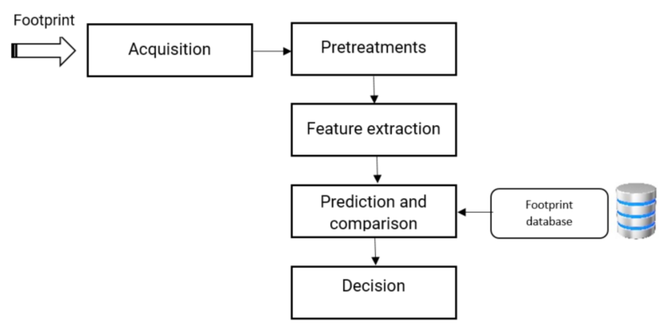

Fig. 1. General architecture of the proposed recognition system.

A. Skeletal Fingerprint Acquisition

The purpose of this phase is to procure digital images, specifically capturing images in both dynamic and static states using a specialized camera.

As previously elucidated, our methodology involves measuring the dimensions of the human skeleton in both static and dynamic states to create a unique key for individual identification and recognition. To achieve this, sixteen photos of the dynamic skeleton and a single photo of the skeleton in a static state are extracted for each sequence. The selection of the number of images for the dynamic state (sixteen photos) for key creation is determined through a reinforcement learning-based model.

We've implemented a reinforcement learning (RL) model to assess the efficacy of keys built from one to 50 images. In essence, the RL model compares 50 keys, where each key, denoted as k ($k \in 1, ..., 50$), is constructed from k dynamic images and one static image. Results derived from the RL model indicate that the key with $k = 16$ yields optimal outcomes in terms of accuracy and computation time cost.

We utilized a digital camera affixed to a tripod to capture gait sequences in an open-air setting. The duration of each sequence varies based on the time taken by each individual to traverse the field of view. However, our specific focus was on extracting sixteen photos of the dynamic skeleton and a single photo of the static skeleton for each sequence. This approach enabled the creation of an initial database comprising 2087 sequences.

To enhance the learning dataset, we incorporated publicly available images from the National Laboratory of Pattern Recognition (NLPR) gait database, as outlined in [17]. This supplementary database encompasses a total of 240 sequences.

For data augmentation, a technique employed in deep learning to generate new images from an original dataset through random transformations, we utilized the *ImageDataGenerator* class from the Keras library. Data augmentation aims to address the challenge of limited images during the training phase. This involves operations that modify the appearance of the image without altering its semantics, such as adjusting brightness or applying rotations (0°C, 45°C, 90°C with respect to the image plane). The implementation of data augmentation enabled the creation of a comprehensive database comprising 10 000 use cases.

B. Pretreatments

Prior to advancing to subsequent stages, the pretreatment step is imperative. The preprocessing of skeletal fingerprints entails three distinct stages: binarization, filtering, and skeletonization.

Binarization: Our initial step involved converting color images to grayscale. Given that the grayscale image encompasses 256 levels (ranging from 0 for black to 256 for white), the binarization process transforms the image into two levels (binary). The threshold is determined and set by the user (we opted for a threshold of 128). Each pixel's value is then compared to this threshold: if it exceeds the threshold, the pixel assumes a value of one (white); otherwise, it takes on a value of zero (black).

Filtering: Filtering is an operation designed to extract information or enhance the visual quality of an image, such as eliminating noise or refining the edges of a blurred image. In our context, a filter is employed to identify the person and extract the skeletal structure.

Skeletonization: In the binarized image (black and white) the lines can be seen clearly but they have different sizes. To be able to quickly detect minutiae (terminations, bifurcations), it is necessary to obtain a more schematic image of the skeleton, in which all the lines have the same thickness.

C. Feature Extraction

For the extraction of features there are several techniques such as moments, wavelet transforms, Markov models, neuron networks and fuzzy algorithms, in addition to the large family of extraction methods. But among the most important analyzes we cite [10] that the skeletal data were extracted in real time in a map of 20 body joints.

Likewise, we are basing ourselves on this study which has shown that it is possible to characterize people with reasonable precision, based on a set of 20 subjects which represents the important characteristics of walking (additionally, Doctor I.B., an author of this article, has confirmed the information). Hence Table I shows exactly the 20 points that we used to characterize a skeleton in static/dynamic state. These points will be modeled in the form of Lagrange polynomials which will be translated in the form of electronic signals. Also a convolutional neural network will be used for the automatic detection of individuals (this will be explained in the next paragraph). Fig. 2 shows an example of the preprocessing results.

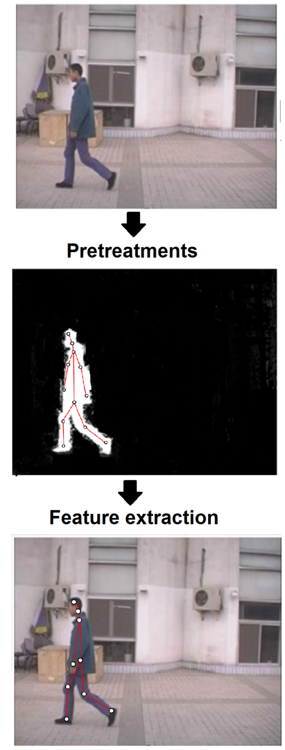

Fig. 2. Example of the preprocessing result.

TABLE I. SKELETON JOINT NOTATION

Joint Label	Joint Type	Joint Label	Joint Type
(x_0, y_0)	Hip Center	(x_{10}, y_{10})	Wrist Right
(x_1, y_1)	Spine	(x_{11}, y_{11})	Hand Right
(x_2, y_2)	Shoulder Center	(x_{12}, y_{12})	Hip Left
(x_3, y_3)	Head	(x_{13}, y_{13})	Knee Left
(x_4, y_4)	Shoulder Left	(x_{14}, y_{14})	Ankle Left
(x_5, y_5)	Elbow Left	(x_{15}, y_{15})	Foot Left
(x_6, y_6)	Wrist Left	(x_{16}, y_{16})	Hip Right
(x_7, y_7)	Hand Left	(x_{17}, y_{17})	Knee Right
(x_8, y_8)	Shoulder Right	(x_{18}, y_{18})	Ankle Right
(x_9, y_9)	Elbow Right	(x_{19}, y_{19})	Foot Right

D. Prediction and Comparison

In our system, we leverage the synergistic power of two robust techniques: Lagrange polynomials and Convolutional Neural Networks (CNN).

1) Lagrange polynomials: Lagrange interpolation polynomials are employed to transform each image set already generated (sequence), comprising the sixteen dynamic skeletons and the static skeleton, into a key represented by electronic signals. The key transformation process is outlined in the following algorithm:

For each use case we have j photos with $1 \leq j \leq 7$, identification of exposed points (x_i^j, y_i^j), with $0 \leq i \leq N_j$: N_j is the number of exposed points chosen (see Table 1) for each use case j. All the x-coordinates must to be different ($x_i^j \neq x_{i+1}^j$). The points are stored in the P_j.

Lagrange polynomial formulation for each use case j:

$$P_j(X) = \sum_{i=0}^{N_j} y_i^j L_i^j(X) = \sum_{i=0}^{N_j} a_i^j (X^i)$$

with $L_i^j(X) = \prod_{k=0,k\neq i}^{N_j} \frac{X - x_k^j}{x_i^j - x_k^j} = \frac{X - x_0^j}{x_i^j - x_0^j} \ldots \frac{X - x_{i-1}^j}{x_i^j - x_{i-1}^j} \frac{X - x_{i+1}^j}{x_i^j - x_{i+1}^j} \ldots \frac{X - x_n^j}{x_i^j - x_n^j}$

Fig. 3 summarizes the objective of Lagrange polynomials:

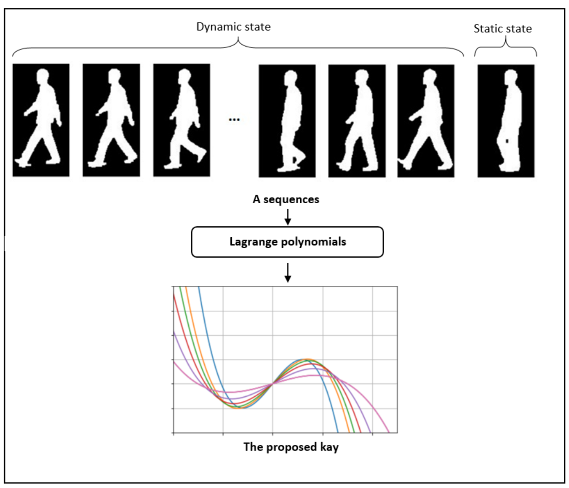

Fig. 3. Example of the proposed key for the identification of individuals.

2) Convolutional neural networks: To establish a neural network, various parameters need to be defined, including the type of neural network, training data, type of training, number of layers, number of neurons (connections), activation functions, and propagation rules, etc. As previously stated, we were provided with a set of image sequences. It is worth noting that when dealing with images, the utilization of Convolutional Neural Networks is essential.

Convolutional Neural Networks (CNNs) are a special type of neural networks specializing in processing grid-type topological data, such as images. The architecture of a CNN model generally contains three distinct types of layers: Convolution Layer, Pooling Layer and Fully connected Layer [1].

Convolution Layer: A convolutional layer is basically responsible for applying one or more filters to our input with the aim is to bring out certain features of the image. It is this layer that distinguishes CNN from other neural networks.

Each convolutional layer contains one or more filters. A filter is essentially a matrix of integers for a subset of the input image of the same size as the filter. Each pixel in the subset is multiplied by the corresponding value in the filter, and the results are summed to get a single value. Repeat this process until the filter "slides" over the entire image. This operation allows to extract the features of the image. An activation function (such as ReLU) is used to output the final features. ReLU essentially guarantees that there are no negative values in the feature output matrix, forcing negative values to zero.

Pooling Layer: Pooling is a technique used to decrease the dimensionality of input features, leading to a decrease in the total number of parameters and the model's complexity. Max pooling is among the most commonly used pooling methods, where only the highest value in the matrix is retained.

Fully Connected Layer: This layer contains traditional neurons that receive sets of weights from the previous layers. This last part which allows learning of the convolutional neural networks. It contains a number of intermediate layers and also a final layer. In the case of the classification problem, the number of neurons in the last final layer is exactly the same number of the classes of the problem treated.

Multiple methods exist for training a neural network to produce specific outputs for given inputs. The current training approach involves Forward/Backward Propagation, utilizing error propagation to adjust the network based on each neuron's contribution to the error. These weights are fine-tuned through gradient descent. An alternative technique for training neural networks is using genetic algorithms ([14]). By training the network on a dataset with known correct outputs, the network can generalize results for new data not part of the training set. In our case, we trained a CNN with 10,000 diverse use cases.

In the literature, there are variations in the architecture of convolutional neural networks, including differences in the number of layers and neurons, owing to various proposed architectures in the field. For our specific application, we propose the following architecture:

We initialize a sequential model then we start by configuring our first convolutional layer to process the inputs of form (352, 240, 1) which is the format of our images then we configure 32 kernels (filters) of form 3 x 3 pixels. The output of these filters will be passed to an activation function "relu" before being forwarded to the next layer. The second pooling layer reduces the representation of the inputs by taking the maximum value on the matrix defined by the "pool_size" parameter which has been configured at (2,2).

The third layer is very similar to the first layer except that this time we have 64 filters instead of 32 but we kept the same size of the filters and at the end of the process the output will go through the "relu" function so that we do not have negative values.

The fourth layer is exactly the same as our second layer.

The fifth layer we used 128 filters.

Once the convolutional and pooling layers have been executed, it becomes essential to incorporate a fully connected layer. This particular layer receives the output data from the convolutional networks, whereby the output of the convolutional network is flattened into a vector form before it is fed into the fully connected layer.

Following the hidden layers, a dropout layer is utilized in the network for regularization, in order to prevent overfitting of the model. The final output layer of the network consists of P neurons, where P denotes the number of sequences within the dataset. The activation function used for this layer is "softmax", which represents a probability distribution to predict the individual. It has been demonstrated that the neural network is trained by adjusting the weights through comparison of predicted results with the actual labels of the sequences in the dataset.

Now that the model has been established, the next step involves training it on the digital representation of the training data. The neural network is equipped with a cost function which needs to be minimized, for which the gradient descent algorithm is employed, specifically utilizing the Adam optimizer. The learning standard employed is a precision of 0,001.

Finally, the metric employed to evaluate the performance of our neural network is defined as accuracy. Accuarcy is defined as the ratio of correctly predicted observations to the total number of observations in the dataset.

$$Accuarcy = \frac{TN + TP}{TN + TP + FN + FP}$$

where, TN = True Negatives, TP = True Positives, FN = False Negatives and FP = False Positives.

3) Comparaison: An automatic fingerprint recognition system yields a positive or negative result by comparing the fingerprint under consideration with all other fingerprints stored in its database.

As previously outlined, our system relies on two techniques for making predictions:

The first technique involves Lagrange polynomials, where the comparison of two fingerprints is executed through the following algorithm: comparing the signals received with the stored signals in the database (denoted as $S_j(X)$ with coefficients s_i^j).

For all j with $1 \leq j \leq 7$ do :
For all i with $0 \leq i \leq N_j$ do :
if a_i^j = s_i^j
Then the imprint is confirmed
else
The imprint is invalid

The second prediction involves the use of Convolutional Neural Networks (CNN). The neural network was trained on an extensive database comprising 10 000 real-world use cases and additional cases generated through data augmentation. Our model enables remote identification and has the capability to accurately predict the target of a skeleton capture.

E. Decision of the Skeletal Fingerprint Recognition System

The ultimate decision is a combination of predictions from both Lagrange polynomials and CNN. In Fig. 4, we illustrate our skeletal fingerprint recognition system. To demonstrate the efficiency of this system, we will provide details in the "Results and Discussion" section.

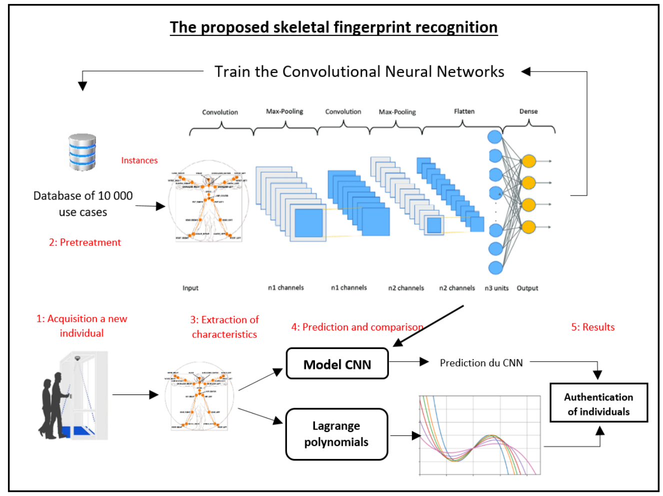

Fig. 4. Architecture of the system.

IV. Results and Discussion

This section is dedicated to evaluating the performance of our CNN model, with a focus on assessing the quality of the solution. It is important to highlight that all experiments were carried out on Google Colab utilizing a GPU.

A. The Learning and Test Rate

The learning rate serves as a critical indicator of model effectiveness. As illustrated in Fig. 5, the graph portrays the learning rate (98%) and validation rate (97%). The visualization of the validation accuracy indicates a successful performance of our CNN, and the minimal gap between the learning accuracy and validation accuracy suggests an absence of overfitting [6]. This observation is reinforced by the learning rate (98%) and validation rate (97%).

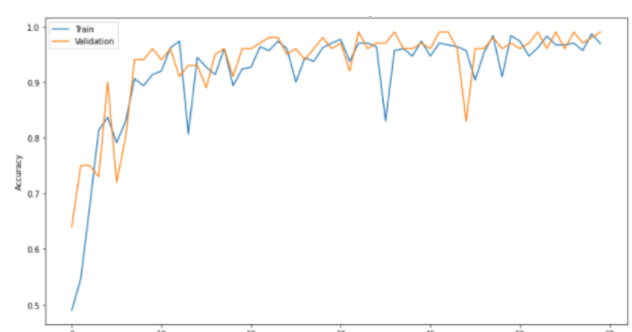

Fig. 5. Development of the training and test score per epoch.

To test the effectiveness of our automatic fingerprint recognition system, we chose 10 peoples and for each one we created 10 new different sequences (human skeleton in static state and in dynamic state), from where the totality of the test

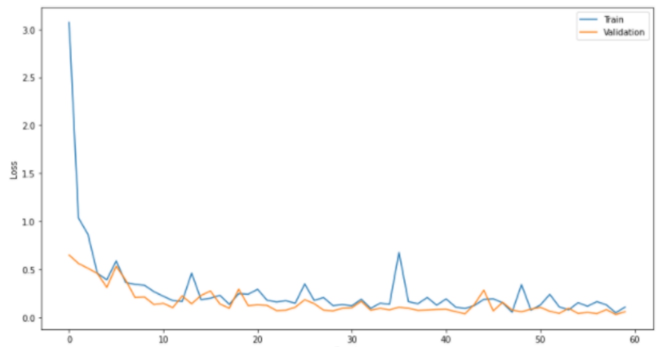

Fig. 6. Convergence of the cost function per epoch.

set equal to 100 sequences and we compared the predictions with the targets using accuracy metric [5]. The following confusion matrix shows (see Fig. 7) the results of the test where the accuracy metric is 95%. Fig. 6 shows convergence of the cost function per epoch.

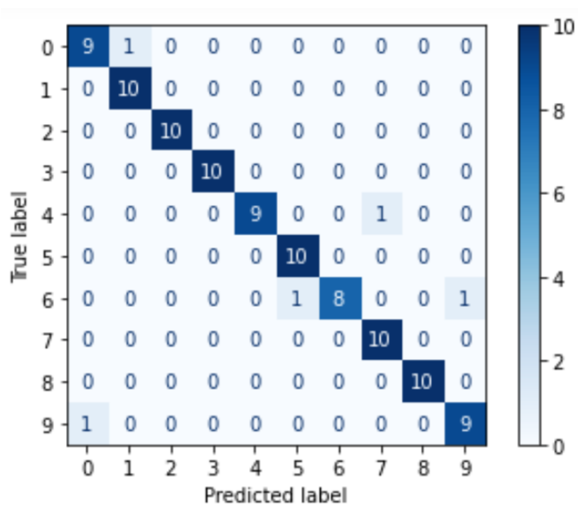

Fig. 7. The confusion matrix of the test set.

V. Conclusion

In this paper, we have introduced a novel skeletal fingerprint recognition system comprising five key modules: acquisition, pretreatments, feature extraction, prediction and comparison, and, ultimately, the decision-making process.

The proposed skeletal fingerprint recognition system is built upon a unique key designed for the identification and recognition of individuals. This key, represented by signals, is modeled using Lagrange polynomials derived from key points on the human skeleton. Consequently, this key can be regarded as a distinct fingerprint categorized within the set of physiological fingerprints. It is essential to emphasize that the primary utility of this fingerprint lies in the remote identification of individuals, free from any masking of the human body. To automate the detection of individuals, a Convolutional Neural Network (CNN) has been integrated. Training data was sourced from various companies within the same domain and augmented through data augmentation, resulting in a comprehensive dataset of 10 000 use cases. The proposed model exhibits a commendable success rate of 98%.

As part of our future endeavors, we aim to extend this approach to a GPU cluster platform, enabling the processing of more complex cases.

Declarations

- Availability of data and materials: All experiments were performed on the google colab under GPU.
- Competing interests: The authors declare that they have no competing interests.
- Funding: The research received no specific grant from any funding agency in the public, commercial, or notfor-profit sectors.
- Ethics approval and consent to participate: The authors declare that the manuscript is our original work. The ideas, views, innovations, and results presented in the above manuscript are totally mine.

Acknowledgment

The authors would like to thank the Editor-in-chief and anonymous reviewers for their comments and suggestions to improve the quality of the paper.

References

[1] A F M Saifuddin Saif, Trung Duong and Zachary Holden, "Computer Vision-based Efficient Segmentation Method for Left Ventricular Epicardium and Endocardium using Deep Learning" International Journal of Advanced Computer Science and Applications(IJACSA), 14(12), 2023

[2] Ahmed, D. M., Ameen, S. Y., Omar, N., Kak, S. F., Rashid, Z. N., Yasin, H. M., ... and Ahmed, A. M. (2021). A State of Art for Survey of Combined Iris and Fingerprint Recognition Systems. Asian Journal of Research in Computer Science, 18-33.

[3] Al-kateeb, Z. N. and Mohammed, S. J. (2020). A novel approach for audio file encryption using hand geometry. Multimedia Tools and Applications, 79(27), 19615-19628.

[4] Banga, L. and Pillai, S. (2021, July). Impact of Behavioural Biometrics on Mobile Banking System. In Journal of Physics: Conference Series (Vol. 1964, No. 6, p. 062109). IOP Publishing.

[5] Berrajaa, A. (2024). Solving the steel continuous casting problem using a recurrent neural network model. International Journal of Computing Science and Mathematics, 19(2), 180-192.

[6] Berrajaa, A. (2022). Natural language processing for the analysis sentiment using a LSTM model. International Journal of Advanced Computer Science and Applications, 13(5).

[7] Bicz, W., Gumienny, Z. and Pluta, M. (1995, June). Ultrasonic sensor for fingerprints recognition. In Optoelectronic and Electronic Sensors (Vol. 2634, pp. 104-111). International Society for Optics and Photonics.

[8] Bouchemha, A., Nait-Ali, A. and Doghmane, N. (2010). A robust technique to characterize the palmprint using radon transform and Delaunay triangulation. International journal of computer applications, 10(10), 35-42.

[9] Gainza, P., Sverrisson, F., Monti, F., Rodola, E., Boscaini, D., Bronstein, M. M. and Correia, B. E. (2020). Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. Nature Methods, 17(2), 184-192.

[10] Gianaria, E., Balossino, N., Grangetto, M. and Lucenteforte, M. (2013, September). Gait characterization using dynamic skeleton acquisition. In 2013 IEEE 15th International Workshop on Multimedia Signal Processing (MMSP) (pp. 440-445). IEEE.

[11] Gu, T., Chen, S., Tao, X. and Lu, J. (2010). An unsupervised approach to activity recognition and segmentation based on object-use fingerprints. Data and Knowledge Engineering, 69(6), 533-544.

[12] Jain, A. K., Ross, A. and Prabhakar, S. (2004). An introduction to biometric recognition. IEEE Transactions on circuits and systems for video technology, 14(1), 4-20.

[13] Jain, A. K., Ross, A. A. and Nandakumar, K. (2011). Introduction to biometrics. Springer Science and Business Media.

[14] Lamos-Sweeney, J. D. (2012). Deep learning using genetic algorithms. Rochester Institute of Technology.

[15] Liu, L. F., Jia, W. and Zhu, Y. H. (2009, September). Survey of gait recognition. In International Conference on Intelligent Computing (pp. 652-659). Springer, Berlin, Heidelberg.

[16] Maharjan, P., Shrestha, K., Bhatta, T., Cho, H., Park, C., Salauddin, M., ... and Park, J. Y. (2021). Keystroke Dynamics based Hybrid Nanogenerators for Biometric Authentication and Identification using Artificial Intelligence. Advanced Science, 2100711.

[17] Rani, M. P. and Arumugam, G. (2010). An efficient gait recognition system for human identification using modified ICA. International journal of computer science and information technology, 2(1), 55-67.

[18] Rashid, R. A., Mahalin, N. H., Sarijari, M. A. and Aziz, A. A. A. (2008, May). Security system using biometric technology: Design and implementation of Voice Recognition System (VRS). In 2008 International Conference on Computer and Communication Engineering (pp. 898-902). IEEE.

[19] Ross, A., Nandakumar, K. and Jain, A. K. (2008). Introduction to multibiometrics. In Handbook of biometrics (pp. 271-292). Springer, Boston, MA.

[20] Sanchez-Reillo, R., Sanchez-Avila, C. and Gonzalez-Marcos, A. (2000). Biometric identification through hand geometry measurements. IEEE Transactions on pattern analysis and machine intelligence, 22(10), 1168-1171.

[21] Shahin, M., Badawi, A. and Kamel, M. (2007). Biometric authentication using fast correlation of near infrared hand vein patterns. International Journal of Biological and Medical Sciences, 2(3), 141-148.

[22] Sun, J., Wang, Y., Li, J., Wan, W., Cheng, D. and Zhang, H. (2018). View-invariant gait recognition based on kinect skeleton feature. Multimedia Tools and Applications, 77(19), 24909-24935.

[23] Uddin, M. Z., Khaksar, W. and Torresen, J. (2017, November). A robust gait recognition system using spatiotemporal features and deep learning. In 2017 IEEE international conference on multisensor fusion and integration for intelligent systems (MFI) (pp. 156-161). IEEE.

[24] Wang, J., She, M., Nahavandi, S. and Kouzani, A. (2010, December). A review of vision-based gait recognition methods for human identification. In 2010 international conference on digital image computing: techniques and applications (pp. 320-327). IEEE.

[25] Wu, J. D. and Ye, S. H. (2009). Driver identification using finger-vein patterns with Radon transform and neural network. Expert Systems with Applications, 36(3), 5793-5799.