
David Zabala-Blanco 1,2,†, Ruber Hernández-García 1,2,*† and Ricardo J. Barrientos 1,2

1 Department of Computing and Industries, Faculty of Engineering Sciences, Universidad Católica del Maule, Talca 3480112, Chile; dzabala@ucm.cl (D.Z.-B.); rbarrientos@ucm.cl (R.J.B.)
2 Laboratory of Technological Research in Pattern Recognition (LITRP), Universidad Católica del Maule, Talca 3480112, Chile

* Correspondence: rhernandez@ucm.cl
† These authors contributed equally to this work.

Abstract: Contactless biometric technologies such as palm vein recognition have gained more relevance in the present and immediate future due to the COVID-19 pandemic. Since certain soft biometrics like gender and age can generate variations in the visualization of palm vein patterns, these soft traits can reduce the penetration rate on large-scale databases for mass individual recognition. Due to the limited availability of public databases, few works report on the existing approaches to gender and age classification through vein pattern images. Moreover, soft biometric classification commonly faces the problem of imbalanced data class distributions, representing a limitation of the reported approaches. This paper introduces weighted extreme learning machine (W-ELM) models for gender and age classification based on palm vein images to address imbalanced data problems, improving the classification performance. The highlights of our proposal are that it avoids using a feature extraction process and can incorporate a weight matrix in optimizing the ELM model by exploiting the imbalanced nature of the data, which guarantees its application in realistic scenarios. In addition, we evaluate a new class distribution for soft biometrics on the VERA dataset and a new multi-label scheme identifying gender and age simultaneously. The experimental results demonstrate that both evaluated W-ELM models outperform previous existing approaches and a novel CNN-based method in terms of the accuracy and G-mean metrics, achieving accuracies of 98.91% and 99.53% for gender classification on VERA and PolyU, respectively. In more challenging scenarios for age and gender–age classifications on the VERA dataset, the proposed method reaches accuracies of 97.05% and 96.91%, respectively. The multi-label classification results suggest that further studies can be conducted on multi-task ELM for palm vein recognition.

Keywords: soft biometrics; gender classification; age estimation; extreme learning machine; palm vein images

1. Introduction

Biometric recognition is now present in various application areas, such as security, health, and banking [1]. In this context, increasingly efficient techniques are required to allow the mass identification of people. Particularly, for the recognition of individuals in massive databases, the use of specific filters can reduce the penetration rate of the database. In a biometrics context, the penetration rate refers to the proportion of the database that must be successfully accessed or searched during a particular query or matching process. Thus, the database can be segmented based on soft biometrics, which allows the characterization of individuals, although these biometrics do not uniquely identify people. Soft biometrics include gender, age, height, weight, and skin color, among others. As a definition, in order to avoid confusing and controversial terms, in soft biometrics, the term “gender” is used to refer to a binary sex categorization (male/female).
Contactless biometric technologies such as palm vein recognition have gained more relevance at present and indeed for the immediate future due to the COVID-19 pandemic. The acquisition of palm vein images uses near-infrared (NIR; a table of abbreviations used in this manuscript is at the end of the paper) devices. Although no statistically significant biases were detected in the score distributions regarding sex and age [4], the obtained palm vein pattern depends on various factors, such as gender, age, body build, skin color, body temperature, and lighting conditions [5–7]. The individual’s gender can be differentiated by the anthropometric measurements of the hand [8], the color and texture of the palm [5], the vessels’ diameter [9], or the hemoglobin levels [10]. Thus, there are some differences between palm vein images captured from male and female subjects.

The palms of males typically exhibit a darker partial shade and more noticeable texture differences, which can be attributed to the distinctions above. Furthermore, [6] indicated that vein-topology-related features, such as the vein pattern gradient, can provide insights into gender information. Additionally, Kuzu et al. [11] used gender-specific features for palm vein recognition. Moreover, the visualization of palm veins is influenced by age and body build due to variations in light intensity caused by differences in hemoglobin levels, water content in the blood, and skin properties [12]. The impact of NIR radiation on hemoglobin over different hues has also been demonstrated by the authors of [13]. In certain scenarios, background absorption dominates the spectra, leading to a reduction in the spectral characteristics of hemoglobin.

Despite the above, it is noteworthy that only the VERA dataset [14] includes metadata about the age and gender of individuals. To the best of our knowledge, the existing works include few approaches on gender and age classification through palm vein images [6,11,15,16]. Furthermore, the PolyU Multispectral Palmprint dataset [17] comprises palmprint images with gender labeling. The infrared images from this database show vascular patterns so that they can be used for soft biometric tasks. However, soft biometric classification commonly faces the problem of imbalanced data class distributions, representing a limitation of the reported approaches. Therefore, the motivation of the proposed methodology is to address these limitations of soft biometric classification in palm vein recognition systems.

Data imbalance typically denotes an unequal distribution of classes within a dataset, where certain classes (minority classes) possess much fewer instances than other classes (majority classes). As most conventional learning algorithms aim to decrease training errors for all classes, an imbalanced class distribution may result in the development of a classifier that provides unsatisfactory predictive performance for the minority classes. Different techniques are available to solve the class imbalance problem [18], most of which are inherently designed for two-class imbalances. However, multi-class data are more common in real-world applications, such as soft biometrics, where learning from multi-class imbalanced data is more challenging. The reason for this is twofold. Firstly, as the number of classes increases, the complexity of the data’s underlying concepts also grows. However, an even more significant factor is the presence of multiple minority classes (referred to as multiminority classes) and multiple majority classes (referred to as multimajority classes), which further complicates the class imbalance problem [19]. For palm vein classification, particularly for the multi-class tasks such as age and gender–age addressed in our work, we face this more challenging multi-class imbalance problem. Unfortunately, existing oversampling solutions have been shown some deficiencies, such as causing severe overgeneralization or not actively improving the class imbalance in the data space [18].

Recently, various approaches have been proposed using Extreme Learning Machines (ELMs) for biometric identification [20,21] and soft biometric classification [22–24]. In this context, Zabala-Blanco et al. [15] evaluated the standard and the regularized ELM to reduce the penetration rate and computational complexity of identification systems based on gender and age classification through palm vein images. Although good results were obtained, both ELM models do not adequately address the problem of training unbalanced data on the VERA dataset [14]. The issue of imbalanced data arises due to the uneven costs
Electronics 2023, 12, 3608

of misclassifying elements from different classes and disparities between the distributions of the training and testing sets [21]. To address this problem, Zong et al. [25] introduced a weighted Extreme Learning Machine (W-ELM) model specifically designed for handling imbalanced data class distributions. The W-ELM approach minimizes the weighted error by incorporating a weight matrix into the optimization problem of ELM, utilizing concepts from structural and empirical risk minimization.

This work introduces weighted ELM (W-ELM) models for gender and age classification based on palm vein images to address imbalanced data problems and improve classification performance. The proposed methodology evaluates two W-ELM approaches on VERA [14] and PolyU [17] datasets, comparing their results against existing methods [6,15], a novel convolutional neural network (CNN)-based approach for palm vein recognition [26], and three deeper CNN architectures evaluated in [11] for gender recognition. The proposed SoftVein-WELM model avoids the extraction of texture descriptors for gender and age single/multi-label classification on palm vein images, which guarantees the robustness and simplicity of the model. We performed a dynamic pre-processing procedure to obtain the region-of-interest (ROI) of palm vein images and numerical optimization of W-ELM hyper-parameters to improve the overall performance. In contrast to the baseline approach [6], utilizing a distinct class distribution for soft biometrics on the VERA dataset enables the adoption of a novel multi-label scheme that can identify gender and age simultaneously (gender-age). Additionally, this approach leads to a reduction in the penetration rate when dealing with extensive databases. Palm vein classification can be a method to increase the recognition performance on large-scale databases. Thus, people are grouped into different classes so that an input palm vein is contrasted only with those belonging to the predicted class, reducing the penetration rate of the matching process [15]. The performed evaluation reveals that the W-ELMs reach the highest performance in terms of accuracy and G-mean metrics.

The SoftVein-WELM model can compensate for the imbalanced nature of the data by the weighted component of the model, which guarantees its application in realistic scenarios and is the main novelty of the proposed methodology. The main contributions of this paper can be summarized as follows:

i. A novel weighted soft biometrics classification model called SoftVein-WELM is proposed that avoids the extraction of texture descriptors from palm vein images and improves classification performance;

ii. A more efficient data-driven method based on W-ELM is proposed that incorporates a weight matrix when optimizing the ELM model by exploiting the imbalanced nature of the soft biometric data, which guarantees its application in realistic scenarios;

iii. An experimental study to measure the performance of different soft biometric traits is conducted on palm vein images outperforming previous state-of-the-art works on VERA and PolyU datasets, including single and multi-label classification tasks and suggesting that further studies on ELM multi-task palm vein recognition can be conducted.

The remaining sections of this article are organized as follows: Section 2 presents the related works; Section 3 formally describes our proposed methodology; Section 4 discusses the experimental results; and, finally, Section 5 summarizes our findings and future works.

2. Related Works

Among the different biometric applications, palm vein recognition has mainly been adopted for authentication or identity verification. On the other hand, for human identification, where exhaustive comparisons are usually made between a query and the biometric database, it is very helpful to have other individual traits to reduce the penetration rate in the database and, therefore, the execution time. From a practical point of view, gender classification and age estimation offer multiple advantages as soft biometrics are used to filter the database. In this regard, different authors have proposed different recognition approaches based on other biometric traits such as fingerprints [27,28], palmprints [29,30],
or the iris \cite{31,32}, face \cite{33,34}, and gait \cite{35,36}. However, to the best of our knowledge, few works in the literature evaluate gender- and age-recognition tasks on vein pattern images. Particularly, we have only found four works for palm vein images \cite{6,11,15,16}, whereas other authors have experimented with dorsal hand images \cite{12,37–39} and the finger veins \cite{11,40}.

In this context, Damak et al. \cite{6} conducted research on gender classification and age group estimation (adolescent, child, and adult) using texture descriptors such as Local Binary Patterns (LBPs) and Center Symmetric Local Binary Patterns (CSLBPs). To address the class imbalance problem in the VERA dataset \cite{14}, the authors compared the performance of K-Nearest Neighbors (KNN) and Weighted K-Nearest Neighbor (WKNN) classifiers. In the gender classification and age estimation tasks, the WKNN classifier achieved better results of 95.8\% and 94.2\% for the F measure and 95.9\% and 94.4\% for the geometric mean, respectively. However, it should be noted that the methodology used in this work presents a bias in the selection of training and testing samples, raising the level of precision in the results. The authors used five cross-validation splits for training, taking 80\% of the data for training and 20\% for testing. This data division makes it possible to take samples from both acquisition sessions for training; thus, the classifier could know the a priori characteristics of the testing set. Notice that details about the VERA database are exposed in Section 3.1.

On the other hand, Zabala et al. \cite{15} evaluated the standard and regularized ELMs for gender and age classification on the VERA dataset \cite{14}. The authors propose a novel approach in which a feature extraction process is unnecessary but instead uses the raw image of the ROI of the palm. After a numerical optimization of their hyper-parameters, the results demonstrate the superiority of the regularized ELM. For this model, an accuracy of 93.40\% and 90.68\% was obtained with a G-mean of 92.56\% and 78.90\% for gender and age classification, respectively. To solve the aforementioned problem of the baseline method proposed in \cite{6}, the authors proposed to separate samples from each subject, using one half for training and the second for testing. Thus, compared to the methods based on LBPs \cite{6}, the evaluated ELM models demonstrated their superiority in terms of performance (accuracy, F-measure, and G-mean), as well as complexity (training time). However, the learning process could be biased or inaccurate in this study because the ELM training stage has information on all individuals in the database. Also, the imbalanced data problem needed to be adequately addressed in the training process of both models.

Due to the robustness of CNN-based feature learning, recent works \cite{11,16} have analyzed the performance of CNN architectures for gender and age classification based on palm vein images. In \cite{16}, a shallow convolutional neural network architecture is introduced to automatically learn discriminative features and a KNN classifier, achieving G-mean values of 96.3\% and 96.7\% for gender and age classification, respectively. On the contrary, Kuzu et al. \cite{11} used a deeper architecture, namely Densenet-161, to improve verification performance by exploiting gender information. In an imbalanced open-set scenario, the results on the VERA database reached an accuracy of 96.82\% for gender classification. Although these CNN-based approaches have shown promising results, they have the drawback of requiring much larger databases for training, evidenced by the higher performance for male subjects. Both in public databases such as VERA and real scenarios, the latter also represents an additional limitation introduced by the significant data imbalance.

Some previous studies based on ELM models have achieved significant performance improvements in terms of accuracy and efficiency for gender and age classification on facial images \cite{22–24}, and preliminary results on palm vein were presented in \cite{15}. From the above literature review, we expanded the application of the ELM model by employing weighted ELMs for gender and age classification using palm vein images and taking into account the inherent class imbalance of the data. Contrary to previous work, the proposed method incorporates a weight matrix in optimizing the ELM model and the extraction of texture descriptors on palm vein images. In the subsequent sections, we present the fundamental principles of the proposed methodology, aiming to address the imbalanced data issues for gender and age classification through palm vein patterns.
In the following sections, we introduce the foundations of the proposed methodology to address the imbalanced data problems, evaluating the overall performance of the weighted ELM for gender and age classification based on palm vein patterns.

### 3. Proposed Methodology

The conceptual overview of the proposed SoftVein-WELM model is shown in Figure 1. The method is based on a weighted ELM model for gender and age single/multi-label classification without a feature extraction process, ensuring the simplicity of the biometric system. The overall procedure comprises three main stages: (i) a palm vein image pre-processing to extract ROIs and enhance vein pattern details; (ii) heuristic optimization of the W-ELM model to compute its hyper-parameters; and (iii) the classification process based on soft biometric classes to obtain performance metrics. In this section, we describe the different components of our proposal according to the experimental design. The results of the model optimization are presented in Section 4.1.

![Figure 1. Overview of the proposed SoftVein-WELM model for gender and age single/multi-label classification on palm vein images using weighted ELM.](image)

#### 3.1. Palm Vein Database and Validation Scheme

In the experiments, we used the VERA-Palmvein database [14], which is the only palm vein image dataset that provides information about individuals’ genders and ages. Due to the lack of public datasets containing soft biometrics information, we perform additional experimentation on the NIR spectrum of PolyU Multispectral Palmprint dataset [17] in order to validate the proposed methodology in at least two public databases. Table 1 summarizes the details of both datasets. It is noteworthy that the imbalance ratio between majority and minority gender classes is significantly greater for the PolyU dataset, which is useful for testing the robustness of the proposed method.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Number of Samples</th>
<th>Male’s Subjects</th>
<th>Female’s Subjects</th>
<th>Male/Female Ratio</th>
<th>Age Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>VERA-Palmvein database [14]</td>
<td>110 × 2 × 2 × 5</td>
<td>70</td>
<td>40</td>
<td>1.75</td>
<td>18–60</td>
</tr>
<tr>
<td>PolyU Multispectral Palmprint (NIR) [17]</td>
<td>250 × 2 × 2 × 6</td>
<td>206</td>
<td>44</td>
<td>4.68</td>
<td>N/A</td>
</tr>
</tbody>
</table>

VERA dataset [14] comprises 110 volunteers, of whom 70 are male and 40 are female, between 18 and 60 years old. For each subject, five samples were obtained from each hand in two acquisition sessions separated by at least 5 min, totaling ten samples per individual and 2200 palm vein images in the database (110 subjects × 2 hands × 2 sessions × 5 samples). The images were acquired using a contactless device at a wavelength of 940 nm and stored with a resolution of 480 px in PNG 8-bit gray-level format. The class selection was based on the specific classification tasks to be performed, namely gender, age, and gender–age. Figure 2 depicts the overall distribution of age and gender groups in the...
VERA database [14]. To address the significant imbalance between age groups proposed in [6], we use a new class distribution for the age classification task by grouping subjects into 10-year age intervals, which is a more natural classification. This approach effectively reduces the penetration rate in the VERA database. Also, it enables the assessment of a multi-label scheme that can simultaneously identify gender and age (gender–age), which is impossible with the baseline distribution due to the lack of male subjects in the Child class. As a result, we combine the gender and age groups, resulting in eight multi-label classes.

![Age and sex groups distribution](image)

**Figure 2.** Overall distribution of age and sex groups from the VERA database.

PolyU Multispectral Palmprint dataset [17] was collected from 15 different spectra, from which we only used the NIR spectrum (880 nm) where vascular patterns are visible. The database includes 250 people, comprising 206 males and 44 females. Totally, we used 6000 images (250 subjects × 2 hands × 2 sessions × 6 samples). The inter-session time interval was between 5–15 days. The acquisition device consisted of a semi-closed environment with a circle LED array to produce uniform illumination. Regarding soft biometrics information, the dataset only provides the volunteers’ gender, so we only can perform the gender classification task. Figure 3 shows samples from the PolyU dataset (NIR), where the differences between the two genders can be seen.

![Palmprint images samples](image)

**Figure 3.** Palmprint images samples from the NIR spectrum of PolyU dataset.

We implemented a validation scheme based on the individuals in the database. For individual identification, the recommended protocol is to use one section for training and the other for testing. However, for gender and age classification, the identification splitting protocol is not recommended because the model can know the a priori characteristics of the testing set. Therefore, the protocol should split the subjects rather than samples or
sections for soft biometric tasks. Thus, we use the subject-exclusive protocol, which does not allow samples of the same person to appear in training and testing subsets. Therefore, we used 80% of individuals for training and 20% for testing in a 5-fold cross-validation scheme in a realistic open-set scenario to avoid overfitting behavior. Thus, both datasets are divided into five folds, each having 20% of the individuals. The ELM model is trained by employing 80% of the palm vein images from the other folds while testing is performed on the actual fold. Hence, the reported performance metrics are the average of the five evaluation iterations.

3.2. Pre-Processing of Palm Vein Images

Images of palm veins are typically captured using a contactless acquisition device with near-infrared illumination. Commonly, the captured image contains the entire hand, including the fingers, and has typical anomalies of a contactless system, such as rotations, translations, close-ups, or illumination differences. Additionally, vein images have low contrast due to the NIR illumination method, which decreases the visualization of vascular patterns. Therefore, image pre-processing is essential as an initial step of palm vein recognition systems.

Initially, we perform an ROI detection procedure on the raw images. In the context of palm vein images, the ROI represents the hand region that contains the most valuable information for the recognition process. Therefore, we attempted to capture most of the palm region to extract the ROI, selecting a square with a side between the center points of the index and pinky fingers to the base of the palm. For this purpose, we adopted the automatic ROI detection method proposed in [41], avoiding the extraction of the hand background and adjusting the ROI dynamically. Later, a gray-level normalization is carried out to reduce illumination variations in the images. Thus, a local normalization is applied in blocks of $32 \times 32$ with an overlap of 3 pixels, considering the average of the pixels of the local area and subtracting the estimated background illumination. Finally, the ROI images were resized to $16 \times 16$ pixels for the ELM models and $128 \times 128$ pixels for the existing models, aiming to standardize the image size, reduce processing time, and maximize the performance results. It should be noted that all the evaluated approaches, including the existing and proposed ones, followed the same pre-processing procedure so that the comparison was made on equal conditions in this regard.

For the training and evaluation of ELM-based models, the obtained images were vectorized as the network input without performing a feature extraction process. Thus, the network input consisted of a vector with 256 positions ($16 \times 16$), normalized in the range $[-1, 1]$ to ensure and maximize the learning of the ELM algorithm. Furthermore, considering the limited number of samples in the VERA database, we implemented an image augmentation procedure to train the CNN-based approach. Thus, we adopt the proposal of [26] as shown Figure 4. In the original image, a central sliding window is set with a size of $3/4$ of the ROI. Later, we perform 5-pixel translations on both axes until reaching the ROI’s border, in addition to performing 5-degree rotations with respect to the center of the ROI until $\pm15^\circ$. By adopting this approach, each original image was augmented with 86 new samples. Although this augmentation technique helps enhance the performance of the CNN model, it is important to note that it introduces additional processing, leading to increased computational time.
Figure 4. Sample augmentation process adopted from [26]. A central sliding window is translated by 5 px on both axes until the ROI border and also rotated by $5^\circ$ with respect to the center of the palm until $\pm15^\circ$.

3.3. SoftVein-WELM Model

The ELM model was proposed by Huang et al. [42], inspired by randomization-based feedforward neural networks [43,44] as a novel learning rule for single-hidden-layer feedforward networks. Compared to traditional neural networks, namely back-propagation and multi-perceptron neural networks, the ELM gives better generalization performance and much faster training efficiency [45]. In the ELM model, the input weights and hidden biases are arbitrarily originated, and the norm of output weights is minimized, which are the reasons behind its advantages. Consequently, the output weights are optimized by following the Moore–Penrose pseudo-inverse theory. Hence, the parameters between the input and hidden layers do not need to be adjusted.

The proposed methodology introduces the SoftVein-WELM model based on the weighted Extreme Learning Machine (W-ELM) for dealing with imbalanced data class distributions on soft palm vein biometrics. The class with a more significant number of instances comes to be the majority, while the minority class represents fewer samples. Hence, the W-ELM model assigns the same weight value to all samples in the minority class, and the same weight value is given to all majority class samples. As a result, the accuracy of the minority class classification is enhanced by properly weighted settings.

The ELM architecture is characterized by only three layers, input, hidden, and output, composed of $u$, $N$, and $v$ neurons, respectively. Given a training set $(X, T) = (x_j, t_j) \mid 1 \leq j \leq M$ of $M$ samples, where the input and output information are represented by $x_j = [x_{j1}, x_{j2}, ..., x_{ju}]^T$ and $t_j = [t_{j1}, t_{j2}, ..., t_{jv}]^T$, respectively, the current output of the ELM acquires the form of [42]

$$f(x_j) = \sum_{i=1}^{N} \beta_i g(w_i \cdot x_j + b_i),$$

where $\beta_i = [\beta_{i1}, \beta_{i2}, ..., \beta_{iv}]^T$ is the output weights, $g(\cdot)$ is the activation function of the hidden neuron, $w_i = [w_{i1}, w_{i2}, ..., w_{iu}]^T$ denotes the input weights, $b_i$ represents the biases of the hidden layer, and $w_i \cdot x_j$ is the inner product between $w_i$ and $x_j$. Considering that, in ELM approaches, $w_i$ and $b_i$ parameters can be randomly originated based on a probability density function and do not require more adjustments, the output function can be written in the following form:

$$t_j = \sum_{i=1}^{N} \beta_i g(w_i \cdot x_j + b_i) = \sum_{i=1}^{N} \beta_i h_i(x_j) = h(x_j)B,$$

being $B = [\beta_1^T, ..., \beta_N^T]^T$ the matrix of the output weights of dimensions $v \times N$ and $h(x_j) = [h_1(x_j), ..., h_N(x_j)]$ the nonlinear feature mapping. For simplicity purposes, Equation (2) can be expressed as
\[ T = HB, \]  
where \( T = [t_1^T, ..., t_M^T]^T \) represents the matrix of target data of dimension \( v \times M \) and \( H \) denotes the output matrix of the hidden layer with \( N \times L \) dimensions, which is given by:

\[
H = \begin{bmatrix}
g(w_1 \cdot x_1 + b_1) & \cdots & g(w_N \cdot x_1 + b_N) \\
\vdots & \ddots & \vdots \\
g(w_1 \cdot x_M + b_1) & \cdots & g(w_N \cdot x_M + b_N)
\end{bmatrix},
\]  

Consequently, the output weights can be trained by exploiting the least squares method \cite{45}

\[
\min_{B \in \mathbb{R}^N} \Phi_{ELM} = \{ \| T - HB \|_2^2 \},
\]

where \( B = H^T T \) represents the solution of Equation (5), \( H^T \) being the generalized Moore–Penrose inverse of \( H \), and is known as the standard ELM. \( H^T \) can be computed using the singular value decomposition (SVD) technique to simplify computational cost, avoid numerical noise, and improve performance results. The SVD of the output weights \( B \) is given by \cite{20}

\[
HB = \sum_{i=1}^{M} z_i z_i^T T,
\]

where \( z \) are the eigenvectors of \( HH^T \).

The least squares problem presented in Equation (5) is poorly formulated by producing an unstable or unfeasible ELM \cite{46}. In addition, it can cause overfitting issues (good performance with the training data but not so with the testing set). The regularizing condition is hence introduced via the following condition \cite{47}:

\[
\min_{B \in \mathbb{R}^N} \Phi_{ELM} = \{ C \| T - HB \|_2^2 + \| B \|_2^2 \},
\]

where the first and second terms represent the error vectors and the regularization condition against overfitting, respectively, and \( C > 0 \) denotes the regularization parameter. The gradient of the ELM with respect to \( B \) is set equal to 0 to maximize the learning capability, based on the results of Equation (7) \cite{15}

\[
\nabla \Phi_{ELM} = B + CH^T (T - HB) = 0.
\]

By employing the Karush—Kuhn—Tucker (KKT) theorem \cite{21}, the previous optimization problem can be found. In case the hidden matrix \( (H) \) is of full-column rank (the number of samples is greater than the number of hidden neurons), the matrix of the output weights acquires the form of \cite{25}

\[
B = (H^T H + I/C)^{-1} H^T T,
\]

where \( I \) corresponds to the identity matrix with \( N \) dimensions. On the contrary, when the number of samples is lower than hidden neurons, \( B \) is given by \cite{25}

\[
B = H^T (H^T H + I/C)^{-1} T,
\]

where \( I \) has dimension \( M \). Equations (9) and (10) represent the R-ELM algorithm, where the hyper-parameter \( C \) must be highlighted.

For imbalanced learning, weighted ELMs (W-ELM) \cite{48} were proposed by adding cost-sensitive ideas into the standard ELM algorithm to balance the data distribution. The W-ELM considers the sample weights to reinforce the impact of the minority classes and the opposite in the minority classes. Based on the R-ELM that minimizes the training error as well as the output weight norm, the optimization function of W-ELM can be represented by \cite{48}:
\[
\min_{B \in \mathbb{R}^N} \Phi_{\text{ELM}} = \{ CW \| T - HB \|^2_2 + \| B \|^2_2 \},
\]
where \( W \) is a \( N \times N \) diagonal matrix whose diagonal elements are the weights assigned to the training samples. The two most popular weighting schemes introduced in [25] for W-ELM are the following:

- **W1-ELM**: \( W_{ii} = 1/N(t_i) \), (12)
- **W2-ELM**: \( W_{ii} = \begin{cases} 
0.618/N(t_i) & \text{if } N(t_i) > \text{mean}[N(t_i)] \\
1/N(t_i) & \text{otherwise} 
\end{cases} \), (13)

where \( W_{ii} \) is the \( i \)-th diagonal element of \( W \), and \( N(t_i) \) corresponds to the number of instances of the \( i \)-th class. It can be seen that samples from the majority classes have smaller weights than those from minority classes; namely, these ELM approaches take into account more of the minority class samples. In addition, it can be noticed that the golden ratio frequently appearing in nature is presented in the second scheme (W2-ELM) to reinforce the decline in the weights of the majority classes. Namely, thanks to this number, the boundary in a binary classification (for instance) is pressed marginally backward from the minority class to the misclassification cases and the compromise on the majority side is alleviated.

According to Lagrange multiplier theory and KKT optimization conditions, the output weights can be calculated as [15]:

\[
B = \begin{cases} 
(H^TWH + I/C)^{-1}H^TW & \text{if } M > N \\
H^T(WH^TH + I/C)^{-1}WT & \text{otherwise} 
\end{cases}
\]

Once again, \( B \) can be determined by a right or left pseudo-inverse. In the case of a small amount of data, the right pseudo-inverse is suitable since it is much easier to compute an \( N \times N \) inverse matrix.

The sigmoid activation function is adopted throughout the paper for the ELM-based models since it is the most adopted in the literature for classification tasks [45]. In addition, the sigmoid function guarantees a universal approximation capability [42,49]. The input weights and biases obey a uniform distribution from \(-1\) to \(1\) [25]. Regarding the basic ELM models (standard and regularized), we adopt the hyper-parameters reported in Zabala-Blanco et al. [15] to increase the overall (accuracy) and particular (G-mean) performances for gender and age classifications without a high computational cost. The number of hidden nodes for the standard and regularized ELMs is 7000 and 6000, respectively. The regularized parameter for the improved ELM algorithm is fixed to \(2^6\).

In order to reinforce the understanding of the weighted ELM model, Algorithm 1 describes the training phase step by step:

**Algorithm 1** Algorithm of the weighted ELM.

For the training set \( \Omega = \{(x_j,t_j) \mid x_j \in \mathbb{R}^u, t_j \in \mathbb{R}^v, j = 1, ..., M \} \), the activation function \( g(\cdot) \), the regularization parameter \( C \), and \( N \) hidden nodes.

1. By following a distribution function, originate the weights \( w_i \) and the biases \( b_i \) of the hidden layer.
2. Determine the output matrix of the hidden layer \( H \), Equation (4), where \( x_j \) inputs are considered.
3. Compute the weighted matrix \( W \), Equations (12) or (13) for W1-ELM and W2-ELM, respectively.
4. Find the weights of the output layer \( \beta \) taking into account the \( t_j \) outputs, see Equation (14).
3.4. Experimental Setup and Metrics

To evaluate the proposed methods, we compare the ELM-based models described in Section 3.3 against the approaches presented in [6,16]. Thus, we implemented the same methods and configurations proposed by their authors.

For the extraction of the LBP and CSLBP descriptors, the ROI image was divided into $4 \times 4$ non-overlapping cells, and a 16-bin histogram was calculated for each region. Consequently, the feature descriptor was formed by concatenating the four histograms to create a 256-feature vector ($4 \times 4 \times 16$), normalized using the min-max technique. To compare the proposed W-ELM models with a more robust feature learning method than LBP and CSLBP, we utilized the CNN-based model presented in [16]. This CNN architecture was chosen due to its high performance in palm vein recognition and its simplicity in terms of computational complexity while still achieving learning outcomes comparable to more complex CNN architectures. The design of the CNN is inspired by the renowned AlexNet model. It comprises five convolutional layers (Conv1-Conv5), two fully connected layers (Full1, Full2), and a final Softmax classifier with cross-entropy loss, as detailed in Table 2. The feature map obtained from the Conv5 layer served as the feature descriptor for the classification process, resulting in a 1024-position vector.

Table 2. Adopted CNN architecture [26], implemented for feature learning on palm vein images. The Softmax layer corresponds to the classifier, where $N$ represents the number of classes.

<table>
<thead>
<tr>
<th>Input</th>
<th>Conv1</th>
<th>Conv2</th>
<th>Conv3</th>
<th>Conv4</th>
<th>Conv5</th>
<th>Full1</th>
<th>Full2</th>
<th>Softmax</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 x 64</td>
<td>7 x 7 x 96</td>
<td>5 x 5 x 192</td>
<td>3 x 3 x 284</td>
<td>3 x 3 x 512</td>
<td>2 x 2 x 1024</td>
<td>1024</td>
<td>512</td>
<td>N</td>
</tr>
<tr>
<td>Strides 2 x 2</td>
<td>BatchNorm</td>
<td>BatchNorm</td>
<td>BatchNorm</td>
<td>BatchNorm</td>
<td>Dropout 0.35</td>
<td>Dropout 0.35</td>
<td>Cross-entropy</td>
<td></td>
</tr>
<tr>
<td>BatchNorm</td>
<td>MaxPool 2 x 2</td>
<td>MaxPool 2 x 2</td>
<td>MaxPool 2 x 2</td>
<td>GlobalAvgPool</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We trained the CNN model using 100 epochs with an early stopping criterion (7 epochs), $\text{batchsize} = 32$, an initial learning rate $lr = 0.01$ ($\text{ReduceLROnPlateau} = 0.2$), SGD optimization, $\text{momentum} = 0.9$, and $\text{weight decay} = 1 \times 10^{-5}$. Additionally, we configured 25 extra training epochs where training and validation data were joined to reduce the lack of training samples. For the testing stage in both approaches, we used a WKNN classifier proposed in [6], and a promising alternative for the Softmax layer involves utilizing the feature vectors extracted by the CNN model. To implement the WKNN, we employed the Euclidean distance with $k = 7$, which was determined through experimentation. The WKNN weights were computed based on the number of samples of the $i$-th class ($|C_i|; i = 1...N$) as follows:

$$W(C_i) = \left( \frac{\min(|C_i|; i = 1...N)}{|C_i|} \right)^{\frac{1}{2}},$$

where $N$ is the total number of classes. Equation (15) is mathematically equivalent to the approach utilized in [6,15]. As a result, the weights fall within the $[0,1]$ range inversely proportional to the number of samples in the class. Consequently, the higher the number of samples in a class, the lower the weight assigned to it, while the minority class retains a weight value of 1.

For the performed evaluations, we employ accuracy and G-mean metrics to assess the performance of the proposed approach. Accuracy is selected because it is the most commonly reported model evaluation metric in machine learning literature. The accuracy relates the number of correctly classified samples to the total number of samples, and it is the most popular metric in machine learning literature [45]. Due to it being oriented to overall predictions, the problem of imbalanced classes could not be fairly evaluated. In this sense, the G-mean is considered, as it represents the geometric mean of sensitivity and specificity, which correspond to positive and negative samples. This metric aims to achieve maximum accuracy for each class while maintaining these accuracies balanced. A low G-Mean indicates poor performance in classifying positive cases, even if negative cases
are correctly classified [50]. This metric aims to achieve maximum precision for each class while maintaining uniformity in precision scores. Furthermore, this measure is widely used in weighted ELM-based works [21,25,48]. For any classification problem, the definitions of the two indicators are as follows [48].

\[
\text{Accuracy}_j = \sqrt{\frac{1}{N_s} \sum_{j=1}^{N_s} (t_{ij} - \tilde{t}_{ij})^2},
\]

\[
\text{G-mean} = \left( \prod_{j=1}^{cl} \text{Accuracy}_j \right)^{1/cl},
\]

where \(N_s\) denotes the number of samples in the class \(j\); \(t_{ij}\) and \(\tilde{t}_{ij}\) are the true value and predicted value, respectively, for the sample \(i\) of of the class \(j\) for a certain input and output; and \(cl\) represents the number of classes. For both metrics, the best value is 1, and the worst value is 0. Notice that in contrast to previous works [6,15], we discard the F-measure since it has many definitions for multi-class tasks [51], which could lead to erroneous comparisons.

4. Results and Discussion

This section shows the obtained results of the evaluated methods for gender and age classification on the VERA and PolyU datasets. Firstly, we present the hyper-parameter optimization process for the proposed SoftVein-WELM model (Section 4.1). After that, we report the experimental results compared against existing approaches on gender and age classification through palm vein images (Sections 4.2 and 4.3). Later, we also compare the results to the state-of-the-art methods based on different biometric traits (Section 4.4). Finally, we summarize and discuss our main findings (Section 4.5).

In the experiments, we used a dedicated server consisting of two Intel Xeon Gold 6140 (36 physical cores) CPUs, with 126 GB of RAM and four GPUs NVIDIA GeForce GTX 1080Ti and running on the Debian GNU/Linux 10 (buster) operating system (kernel 4.19.0-17-amd64 x86 64). The existing approaches based on LBP descriptors [6], CNN feature learning [16], and WKNN classifiers were developed with Python (v.3.7.3), using the libraries OpenCV (v.4.1), Tensorflow (v.1.15.0), Keras (v.2.3.1), Scikit-learn (v.0.24), CUDA 10.1 toolkit, and cuDNN 7.5. On the other hand, the ELM-based models were implemented with Matlab R2018b.

4.1. Optimization of the ELM Models’ Hyper-Parameters

In this subsection, a heuristic optimization for the hyper-parameters of the imbalanced ELM approaches is carried out. The hyper-parameters correspond to the hidden neurons and regularization parameters for both models. These should be fixed by considering the augmentation in performance without increasing the classification system’s complexity. The learning phase is useful for finding the weights between the hidden and output layers by following the ELM algorithm [48]. It is noteworthy that brute force optimization can only be performed because it does not require iterative processes (see Section 3.3).

The accuracy and G-mean metrics in terms of the hyper-parameters of the (a) W1-ELM and (b) W2-ELM for age, gender, and gender-age classifications are depicted in Figures 5, 6, 7 and 8, respectively. Both figures are represented by a three-dimensional surface with solid edge colors and varying face colors due to the ELM models having hidden neurons and regularization parameters as hyper-parameters. The color of the surface varies according to the heights based on accuracy or G-mean values. However, it is noteworthy that the values slightly vary within a quadrilateral (i.e., minor variations in the number of neurons and the regularization parameter) due to the effect of the regularization parameter. Both metrics are considered so that majority and minority classes have the same importance for the identification system. The exposed results correspond to the validation stage for all ELM models.
Figure 5. Performance in terms of the number of hidden neurons and regularization parameter of the (a) W1-ELM and (b) W2-ELM models for gender classification task (two classes) on the VERA dataset.

Figure 6. Performance in terms of the number of hidden neurons and regularization parameter by the (a) W1-ELM and (b) W2-ELM models for gender classification task (two classes) on the PolyU dataset.

Figure 7. Performance in terms of the number of hidden neurons and regularization parameter of the (a) W1-ELM and (b) W2-ELM models for age classification task (four classes) on the VERA dataset.

Figure 8. Performance in terms of the number of hidden neurons and regularization parameter of the (a) W1-ELM and (b) W2-ELM models for gender–age classification task (eight classes) on the VERA dataset.

As mentioned, our proposal is used for gender and age classification on the VERA database (see Figures 5, 7 and 8), and for gender classification on the PolyU dataset (see Figure 6). The same ELM architecture for all classification problems is adopted on the VERA database for generalization purposes. The three-dimensional surface plots of the W1-ELM and W2-ELM show quasi-rectangular planes where performance metrics reach their maximum values. In order to remain in these planes, the regularization parameter...
must be greater than $2^8$, and hidden neurons must be greater than 3000. Considering the computational cost and over-fitting issue, the following hyper-parameters are used: 6000 as the number of hidden neurons and $2^{12}$ as the regularization parameter. Regarding the PolyU dataset, the accuracy and G-mean reach maximum values as the regularization parameter and the number of hidden neurons increase. There is a zone (it can be identified with yellow color) where the performance metrics are sub-optimized regarding the ELM hyper-parameters. To keep the computational cost to a minimum as well as to ensure high performance for W1-ELM and W2-ELM, the number of hidden neurons is set to 4000, and the regularization parameter is fixed to $2^{18}$ in the rest of the paper. Notice that (i) we discard observations regarding the performance comparison for now, (ii) the accuracies and G-means slightly vary (as will be shown below with the standard deviations) due to the input weights and biases of the hidden nodes being randomly generated by the ELM algorithm, and (iii) given any subfigure, there is a local minimum when the number of hidden neurons is equal to the number of total samples, and the regularization parameter tends to infinity.

### 4.2. Gender Classification Results

Table 3 shows the gender classification results on both studied datasets. We report the average accuracy and G-mean for the 5-fold cross-validation scheme, together with their standard deviations (identified with parentheses). In order to compare our results against the reported methods in the literature, the table summarizes the W1-ELM and W2-ELM algorithms and existing approaches: LBP-based descriptors with a WKNN classifier [6], a CNN-based feature learning approach [16] with a WKNN classifier, and the standard and regularized ELMs [15]. The results indicate that the imbalance ratio (Male/Female) affects the classification performance, which is most noticeable for the PolyU database, where the majority class exceeds the minority class by 4.68 times. It can be seen that the proposals (W1-ELM and W2-ELM) are better for both metrics and datasets. Among the rest of the methods, the regularized ELM [15] obtains a similar performance, especially for the VERA database where the imbalanced problem is not so noticeable, see Figure 2. On the other hand, using a palm vein descriptor and a WKNN classifier (classical approaches) presents the worst accuracies and G-means. Finally, the results show improved generalization and classification performance of the proposed W-ELM models compared to the existing approaches, as higher the imbalance ratio such as in the PolyU dataset.

Table 3. Performance comparison of the existing and proposed approaches for gender classification on VERA and PolyU datasets. Standard deviations are shown in parentheses, and the higher results are identified in bold.
4.3. Age and Gender–Age Classification Results

Aiming to evaluate our method in more challenging scenarios, Table 4 compares the experimental results of the existing and proposed approaches for age and gender–age classifications on the VERA dataset. Both scenarios represent multi-class problems, which are more common in real-world applications. Thus, as the number of classes increases, the complexity of the data’s underlying concepts also grows. However, an even more significant factor is the presence of multiple minority classes (multiminority) and multiple majority classes (multimajority), further complicating the class imbalance problem.

Table 4. Performance comparison of the existing and proposed approaches for age and gender–age classifications on the VERA dataset. Standard deviations are shown in parentheses, and the higher results are identified in bold.

<table>
<thead>
<tr>
<th>Model</th>
<th>Age</th>
<th>Gender–Age</th>
<th>Age</th>
<th>Gender–Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>G-Mean</td>
<td>Accuracy</td>
<td>G-Mean</td>
</tr>
<tr>
<td>LBP and WKNN [6]</td>
<td>0.7571 (±0.0266)</td>
<td>0.7430 (±0.0329)</td>
<td>0.6295 (±0.0401)</td>
<td>0.6864 (±0.0382)</td>
</tr>
<tr>
<td>CSLBP and WKNN [6]</td>
<td>0.8753 (±0.0458)</td>
<td>0.8951 (±0.0532)</td>
<td>0.6679 (±0.0131)</td>
<td>0.7145 (±0.0123)</td>
</tr>
<tr>
<td>CNN with WKNN [16]</td>
<td>0.9486 (±0.0560)</td>
<td>0.9672 (±0.0228)</td>
<td>0.9272 (±0.0856)</td>
<td>0.9591 (±0.0142)</td>
</tr>
<tr>
<td>Standard ELM [15]</td>
<td>0.9700 (±0.0104)</td>
<td>0.9585 (±0.0105)</td>
<td>0.9591 (±0.0102)</td>
<td>0.9534 (±0.0148)</td>
</tr>
<tr>
<td>Regularized ELM [15]</td>
<td>0.9700 (±0.0072)</td>
<td>0.9568 (±0.0172)</td>
<td>0.9645 (±0.0044)</td>
<td>0.9574 (±0.0052)</td>
</tr>
<tr>
<td>W1-ELM (proposed)</td>
<td>0.9673 (±0.0121)</td>
<td>0.9660 (±0.0119)</td>
<td>0.9677 (±0.0030)</td>
<td>0.9600 (±0.0080)</td>
</tr>
<tr>
<td>W2-ELM (proposed)</td>
<td>0.9705 (±0.0103)</td>
<td>0.9694 (±0.0112)</td>
<td>0.9691 (±0.0047)</td>
<td>0.9614 (±0.0077)</td>
</tr>
</tbody>
</table>

Our proposed methods (particularly W2-ELM) outperform the others for both metrics. In contrast, techniques based on LBP descriptors perform the worst in all scenarios, especially the simplest LBP with WKNN. Standard and regularized ELM algorithms and CNN feature learning with WKNN almost achieve the W2-ELM’s level of performance as the penetration rate decreases and the minority classes become relevant. Another advantage of the ELMs over the other approaches is their learning stability (see low deviation standards), which is always desired for reliability reasons. It is worth mentioning that results for the proposed 10-age class distribution are better than the baseline age distribution [15], which reported 0.9068 as the best accuracy. The reasons behind the W2-ELM’s superiority are twofold: (i) it discarded a feature extractor, which can limit the classification capacity, and (ii) it includes the golden ratio for majority classes, which usually appears in real and complex datasets (such as the VERA database), regardless of the number of classes [25]. For instance, the relationship between men and women corresponds to $70/40 = 1.75$, which is near the golden ratio of $1/0.618 = 1.618$. To examine the class-wise impact of the W2-ELM model, the model with the best performance, Figure 9 shows examples of the confusion matrices for two multi-class tasks: age and gender–age classification. It can be observed that misclassifications are concentrated on consecutive age groups or same-gender classes for age and gender–age tasks since our model can almost always predict at least the adjacent age group.
4.4. Comparison against Other Biometrics Traits

The study of gender and age classification is not so extensive on palm vein images due to the low availability of datasets containing this information. For this reason, the comparison made in the previous section against related work is quite limited. Aiming to contrast the obtained results within a broader scope of studies, in the following, we compare some existing approaches for gender and age classification using different biometric traits.

Tables 5 and 6 summarize some existing approaches for gender and age classification, respectively. For the comparison in terms of accuracy, we considered reported works based on vein patterns (finger vein and hand-dorsal vein), fingerprints, palmprints, and the iris, face, and gait, with experiments on datasets with imbalanced distributions. To the best of our knowledge, we reported the most relevant work with the highest results. Due to the reported works focusing on different biometric traits, most were evaluated on other datasets. These results are shown only for exploratory purposes in order to give an overview of the state-of-the-art in these soft biometrics using different biometric traits. Nevertheless, it can be seen that the proposed approach achieves promising results with respect to other biometrics.

Among vein-based techniques [12,39,40], our proposal on palm vein images outperforms those on the finger and hand-dorsal veins without extracting hand-crafted features. For gender classification, our method performs similarly to the best results on palmprint [30], which contains rich information also capturing some palm vein structures in an ROI area similar to palm veins. Compared to more traditional biometric methods, such as fingerprints, the iris, and the face, our method demonstrates simplicity based on W2-ELM contrasting hybrid [31,33,52], multi-task [35,53] or deep [27,33] models. On the other hand, age group classification is a less challenging task in vein-based biometrics, obtaining higher results than other traits. Particularly, we have yet to find any relevant study for age classification on palmprint images due to the lack of datasets with this kind of information. Similar to the gender-classification approaches, the results of hybrid [28,32,54] and multi-task [35,55,56] models stand out for age classification, although with lower performance than ours. Finally, our results show the potential of palm vein patterns for gender and age classification over other biometrics that contain more information regarding these soft traits, such as face or gait. Based on the considered articles, the proposed SoftVein-WELM model is recommended as a good learning classifier for soft biometrics. As depicted in Tables 5 and 6, our proposed approach slightly outperforms the results on veins, fingerprints, face, and gait for gender classification, whereas it performed better than other methods on veins, fingerprints, iris, and face for age classification.
Table 5. Comparison against existing approaches in terms of accuracy for gender classification from different biometric traits. The best results for each biometric are highlighted in bold.

<table>
<thead>
<tr>
<th>Biometric Trait</th>
<th>Method</th>
<th>Dataset</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finger vein</td>
<td>LBP+KNN [40]</td>
<td>UTFVP [57] (44 male and 16 female)</td>
<td>94.89%</td>
</tr>
<tr>
<td>Vein-based</td>
<td>Densenet-161 CNN [11]</td>
<td>Self-created (102 male and 98 female)</td>
<td>98.20%</td>
</tr>
<tr>
<td>Palm vein</td>
<td>W2-ELM (proposed)</td>
<td>VERA [14] (70 male and 40 female)</td>
<td>98.91%</td>
</tr>
<tr>
<td>Vein-based</td>
<td>Fusion scheme+CNN [52]</td>
<td>Self-created (111 male and 128 female)</td>
<td>91.30%</td>
</tr>
<tr>
<td>Fingerprint</td>
<td>DDC-ResNet [27]</td>
<td>Self-created (98 male and 102 female)</td>
<td>96.50%</td>
</tr>
<tr>
<td>Palmprint</td>
<td>Fine-tuned VGGNet [60]</td>
<td>PolyU (blue) [17] (206 male and 44 female)</td>
<td>89.20%</td>
</tr>
<tr>
<td>Iris</td>
<td>Texture Stats+ANN [62]</td>
<td>IIT Delhi IRIS database [63] (176 male and 48 female)</td>
<td>96.40%</td>
</tr>
<tr>
<td>Iris</td>
<td>2DMWT+Radon+SVN [64]</td>
<td>LFW [66] (10,256 male and 2977 female faces)</td>
<td>93.90%</td>
</tr>
<tr>
<td>Iris</td>
<td>Hybrid Feature+SVN [31]</td>
<td>MORPH II [68] (46,645 male and 8487 female faces)</td>
<td>95.56%</td>
</tr>
<tr>
<td>Iris</td>
<td>DWT Stats+Ensemble DT [74]</td>
<td>OU-ISIR [69] (389 male and 355 female)</td>
<td>97.70%</td>
</tr>
<tr>
<td>Iris</td>
<td>PDM+Adaboost and SVM Fusion [32]</td>
<td>OU-MVLP [71] (5114 male and 5193 female)</td>
<td>94.27%</td>
</tr>
<tr>
<td>Gait</td>
<td>Multi-view PA-GCR+GaitSet [70]</td>
<td>OU-MVLP [71] (5114 male and 5193 female)</td>
<td>95.33%</td>
</tr>
</tbody>
</table>

Table 6. Comparison against existing approaches approaches in terms of accuracy for age classification from different biometric traits. The best results for each biometric are highlighted in bold.

<table>
<thead>
<tr>
<th>Biometric Trait</th>
<th>Method</th>
<th>Dataset</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finger vein</td>
<td>LBP+KNN [40]</td>
<td>UTFVP [57] (ages within 10–90 years, separated into four 20-year classes)</td>
<td>96.33%</td>
</tr>
<tr>
<td>Vein-based</td>
<td>Histograms Stats+KNN [12]</td>
<td>Self-created (two classes, 10 in ‘young’ class with &lt;30 years and 40 in ‘old’ class with &gt;60 years)</td>
<td>90.40%</td>
</tr>
<tr>
<td>Palm vein</td>
<td>W2-ELM (proposed)</td>
<td>VERA [14] (ages within 18–60 years, separated into four 10-year classes)</td>
<td>97.05%</td>
</tr>
<tr>
<td>Fingerprint</td>
<td>DWT+KNN [73]</td>
<td>Self-created (2830 fingerprints within 13–35 years, divided into three groups)</td>
<td>74.54%</td>
</tr>
<tr>
<td>Iris</td>
<td>DWT Stats+Ensemble DT [74]</td>
<td>Self-created (494 users separated into two groups of 283 young and 211 elderly people)</td>
<td>89.10%</td>
</tr>
<tr>
<td>Iris</td>
<td>PDM+Adaboost and SVM Fusion [32]</td>
<td>OU-MVLP [71] (5114 male and 5193 female)</td>
<td>83.70%</td>
</tr>
<tr>
<td>Iris</td>
<td>ASM+CNN [54]</td>
<td>Adience [76] (2284 subjects labeled by eight age groups)</td>
<td>92.62%</td>
</tr>
<tr>
<td>Face</td>
<td>Pretrained VGG16 [34]</td>
<td>FG-NET [77] (82 people with ages range 0–69 years, divided into four age groups)</td>
<td>94.59%</td>
</tr>
<tr>
<td>Face</td>
<td>Gated Residual Attention Network [55]</td>
<td>Adience [76] (2284 subjects labeled by eight age groups)</td>
<td>96.60%</td>
</tr>
<tr>
<td>Face</td>
<td>Multi-task Fusion CNN [35]</td>
<td>UTKFace [65] (20K face images ranging 0–116 years, divided into ten age categories)</td>
<td>93.70%</td>
</tr>
<tr>
<td>Gait</td>
<td>Pyramid GEIs+Parallel Fused CNN [56]</td>
<td>OU-ISIR [69] (744 subjects with ages range 2–78 years, divided into six age groups)</td>
<td>96.10%</td>
</tr>
<tr>
<td>Gait</td>
<td>Joint Angles Fourier series+KNN [36]</td>
<td>OU-Isir [69] (74 subjects with ages range 19–54 years, divided into four age groups)</td>
<td>98.20%</td>
</tr>
</tbody>
</table>
4.5. Discussion and Main Findings

This study introduces a weighted extreme learning machine (W-ELM) for gender and age classification based on palm vein images. The proposed SoftVein-W-ELM model avoids the feature extraction process and incorporates a weight matrix by exploiting the imbalance of soft biometric data. We evaluated our methodology on the VERA-Palmvein database. We used a new class distribution for the age classification task by dividing the subjects into 10-year age groups to reduce the significant imbalance between age groups. In addition, a subject-exclusive protocol was used to avoid using samples of the same subjects in training and testing subsets.

The experimental results demonstrate that the W-ELM models achieve accuracies of 98.91% and 97.05% for gender and age classification, respectively. We compared the obtained results against existing state-of-the-art approaches, with significantly improved performance over traditional methods [6,16] and slightly outperforming the previous ELM-based approaches [15]. In addition, the ELM multi-label classification results reached an accuracy of 96.91%, suggesting that further studies on ELM multi-task palm vein recognition can be conducted by combining subject identification and soft biometric classification.

Particularly, the results achieved using W2-ELM highlight its learning reliability and stability (i.e., low deviation standard deviation). It is noteworthy that classification results are affected by the increase in classes and data imbalance, while the differences between traditional methods and our proposed method are more significant. Therefore, incorporating the weight matrix as part of the model optimization process is an appropriate strategy to address the limitations of soft biometrics and imbalanced data. Moreover, misclassifications tend to focus on successive age- or gender-based classes, showing that our model can predict nearly every adjacent age group.

Regarding computational costs, it is worth noting that our proposal is characterized by its efficiency, which can be determined analytically. The proposed ELM-based models do not require a feature extraction process, and the training is carried out from the solutions of linear matrix systems. An ELM is a simple and efficient learning algorithm for single-hidden layer feedforward neural networks (SLFNs). Additionally, the solution for least squares in W-ELM is, in fact, an analytical solution, which is accomplished through a limited and fixed number of steps. The remarkable feature of this method is that it avoids using a computationally complex optimization procedure by using random input weight allocation. Although the training times of the WKNN classifier are practically negligible, such methods require feature extraction for all gallery samples, and, in the testing stage, they require exhaustive comparisons against all of them. Thus, the more samples in the gallery, the longer the time the WKNN classifier requires. Moreover, the CNN-based feature representation approach has the drawback of involving a computationally costly training process that requires a large number of samples. However, it is well-known that training is an offline process with less effect on processing times, highlighting the use of ELMs in practical solutions and low-cost devices. Regarding the classification process, no significant differences are shown between the compared models [15].

Our findings demonstrate that the potential of palm vein patterns for gender and age classification surpasses other biometric modalities that provide more extensive information regarding these characteristics, such as facial features or gait. As shown in Tables 5 and 6, our proposed methodology slightly outperforms gender classification methods based on the veins, fingerprints, face, and gait, whereas it exhibits superior performance to existing techniques for age classification based on the veins, fingerprints, iris, and face. After thoroughly analyzing the relevant literature, the SoftVein-W-ELM model is recommended as an effective learning classifier for soft biometrics.
The main limitation of our study is the need for soft biometric information on public palm vein databases. Therefore, further research should be carried out after creating new palm vein datasets with a more significant number of subjects along with their soft biometric information. Moreover, the lack of interpretability of the ELM models is a drawback of our proposal, which makes them less suitable for use than simpler algorithms like KNN to understand the underlying mechanisms and reasoning behind the classification model.

5. Conclusions and Future Works

Classifying soft biometrics such as gender and age is an essential recognition task to reduce penetration rates in large-scale biometric databases. In order to reduce the influence of the imbalance of these types of data, this paper introduced the use of W-ELM models for gender and age classification on palm vein images. The proposed SoftVein-WELM model uses the palm ROI images to train and classify new samples, avoiding a feature extraction process. For evaluating the proposed method, we used 10-year age groups and a new multi-label scheme identifying gender and age simultaneously. The obtained performance results show that both evaluated W-ELM models outperform the previous existing approaches and a novel CNN-based method. Our proposed method’s highlights are its simplicity and ability to incorporate a weight matrix to optimize the ELM model. Comparison against other biometrics traits suggests the potential of palm vein patterns and W-ELM models for gender and age classification.

However, the primary limitation of our study is the lack of soft biometric data in publicly available palm vein databases. Consequently, further research is needed involving the creation of new palm vein datasets with a more significant number of subjects along with soft biometric information. Based on the multi-label outcomes, exploring multi-task recognition models that integrate human identification and soft biometric classification is suggested. Moreover, as large-scale databases and real-world applications continue to develop, we intend to implement our proposed approach for massive palm vein identification to reduce the search space and enhance matching efficiency. Lastly, considering the promising results, weighted data-driven models like W-ELMs should be further investigated for other biometric traits.

Author Contributions: All authors contributed to the study conception and design. Conceptualization, methodology, and formal analysis were performed by D.Z.-B. and R.H.-G. The first draft of the manuscript was written by D.Z.-B. and the review and editing upon the draft version were carried out by R.H.-G. and R.J.B. Funding was acquired by R.H.-G. and R.J.B. All authors read and approved the final manuscript.

Funding: This work was partially funded by Agencia Nacional de Investigación y Desarrollo (ANID), Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, Gobierno de Chile, Research Project ANID FONDECYT Iniciación en Investigación 2022, Grant number No. 11220693.

Data Availability Statement: VERA-Palmvein database can be found on: https://www.idiap.ch/dataset/vera-palmvein (accessed on 22 August 2023). PolyU Multispectral Palmprint dataset is publicly available by requesting to the principal investigator. The source code implemented to obtain the presented results will be available at https://github.com/ruberhg/SoftVein-WELM (accessed on 22 August 2023).

Acknowledgments: Authors thank ANID FONDECYT Iniciación Project No. 11220693, ANID FONDECYT Regular Project No. 1211132, ANID FONDECYT Regular Project No. 1200810, and the Laboratory of Technological Research in Pattern Recognition (https://www.litrp.cl). Portions of the research in this paper used the VERA-Palmvein Corpus made available by the Idiap Research Institute, Martigny, Switzerland.

Conflicts of Interest: The authors declare no conflict of interest.
Abbreviations
The following abbreviations are used in this manuscript:

- CBP: Local binary pattern
- CNN: Convolutional neural network
- CSLBP: Center symmetric local binary pattern
- ELM: Extreme learning machine
- KKT: Karush–Kuhn–Tucker
- KNN: K-nearest neighbors
- NIR: Near-infrared
- R-ELM: Regularized extreme learning machine
- ROI: Region-of-interest
- SVD: Singular value decomposition
- SGD: Stochastic gradient descent
- W-ELM: Weighted extreme learning machine
- WKNN: Weighted K-nearest neighbor

References
10. Murphy, W.G. The sex difference in haemoglobin levels in adults—Mechanisms, causes, and consequences. *Blood Rev.* 2014, 28, 41–47. [CrossRef]


41. Hernández-García, R.; Barrientos, R.J.; Rojas, C.; Mora, M. Individuals identification based on palm vein matching under a parallel environment. *Appl. Sci.* 2019, 9, 2805. [CrossRef]


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