An Adjustable Face Recognition System for Illumination Compensation Based on Differential Evolution

Guilherme Felippe Plichoski
Graduate Program in Applied Computing
State University of Santa Catarina
Joinville, Santa Catarina, Brazil
Email: guilherme.plichoski@edu.udesc.br

Chidambaram Chidambaram
Department of Production Engineering
State University of Santa Catarina
Joinville, Santa Catarina, Brazil
Email: chidambaram@udesc.br

Rafael Stubs Parpinelli
Graduate Program in Applied Computing
State University of Santa Catarina
Joinville, Santa Catarina, Brazil
Email: rafael.parpinelli@udesc.br

***Abstract*—It is well known that face recognition (FR) systems cannot perform well under uncontrolled conditions, but there are no general and robust approaches with total immunity to all conditions. Hence, we present an adjustable FR framework with the aid of the Differential Evolution (DE) optimization algorithm. This approach implements several preprocessing and feature extraction techniques aiming to compensate the illumination variation. The main feature of the present work stands on the use of the DE which is responsible for choosing which strategies to use, as well as tunning the parameters involved. In this case study, we aim to address the illumination compensation problem applying on the well known Yale Extended B face dataset. According to the proposed FR framework, the DE can choose any combination of the following techniques and tune its necessary parameters achieving optimized values: the Gamma Intensity Correction (GIC), the Wavelet-based Illumination Normalization (WBIN), the Gaussian Blur, the Laplacian Edge Detection, the Discrete Wavelet Transform (DWT), the Discrete Cosine Transform (DCT), and the Local Binary Patterns (LBP). Our experimental analysis confirms that the proposed approach is suitable for FR using images under varying conditions. It is proved by the average recognition rate of 99.95% obtained using four different datasets.**

I. Introduction

With the exponential development of technology, biometrics identification has drawn attention from the research community expanding the range of possibilities for its application. Among them, face recognition (FR) received significant attention in security areas and other applications, especially in information security domain, becoming more popular in recent years [1]. With this, several algorithms have been proposed. Although most of them have made huge progress in various applications and research fields, they are still facing great challenges such as variations of illumination, pose, expression and occlusion [2] [3]. Recent FR methods are based on Convolutional Neural Networks (CNN), and have made a great progress but its performance is still not enough for real-world applications due to the problems such as data bias and overfitting [4]. Thus, FR remains an open research topic.

Regarding the aforementioned challenges found in FR systems, uncontrolled illumination is a critical factor for real-world applications due to the non-ideal imaging environments frequently encountered in practical cases [3]. In the review of illumination compensation approaches performed in 2015 [5], some authors presented the works with most significant results. Firstly, Cheng et al. [6] proposed the use of Discrete Wavelet Transform (DWT) technique for illumination normalization based on retinal modeling executed on low-frequency band. In Baradarani's work [7], the authors suggest the use of a Double-Density Dual-Tree Complex Wavelet Transform (DD-DTCWT) with Principal Component Analysis (PCA) for dimensionality reduction followed by the classification task which is performed by an Extreme Learning Machine (ELM) routine. After, Baradarani and Wu [8] proposed the use of resonance based decomposition of images for illumination invariant face recognition using PCA and ELM. In [9], the authors used a fuzzy filter applied over the low-frequency Discrete Cosine Transform (DCT) coefficients for FR task. Hu's paper [10] proposes a DWT-based denoising technique and a multi-scale reflectance model to extract the illumination invariant features. Zhao et al [11] proposes to build a facial model using the stable point pairs of the Grayscale Arranging Pairs (GAP) feature. Besides these, recent works still attempt to address the illumination related issues such as [12] presented a 2D image-matrix-based error model, namely, nuclear norm based matrix regression (NMR) and an algorithm to solve the approximate NMR model. The authors show that their method has overcome the state-of-the-art regression-based methods. Last, Varadarajan et al [13] proposed a FR system using a combination of Chirp Z-Transform (CZT) and Goertzel algorithm for illumination normalization and Binary PSO optimization algorithm for feature selection. Their work demonstrates better performance compared to existing methods.

All works reviewed propose different methodologies aiming to overcome the illumination variation through the compensation. In our work, we present an adjustable FR framework with the aid of Differential Evolution (DE) optimization algorithm. The proposed methodology provides several illumination compensation techniques for preprocessing and for feature extraction which includes Gamma Intensity Correction (GIC),

Wavelet-based Illumination Normalization (WBIN), Gaussian Blur, Laplacian Edge Detection, DWT, DCT and Local Binary Patterns (LBP) and the optimization process is carried out by the DE algorithm. The DE algorithm was chosen since it is well known and efficient optimizer when applied to problems in continuous domains [14]. This algorithm is responsible to choose which technique will be part of the training and test classification model and furthermore it optimizes the parameters of selected techniques. Thus, the performance of the proposed FR system strongly depends on the chosen techniques and their parameter tunning. The well known Yale Extended B face database [15] is employed in experiments aiming to work on the illumination variations contained in this dataset. For performance evaluation, we compare our results with the works found in the literature.

The remaining sections of this work are organized as follows: Section II briefly describes the set of techniques for preprocessing and for feature extraction as well as the DE optimization algorithm; Section III presents the proposed FR system in details; Section IV discusses and exposes the experimental results; finally, Section V concludes this paper and gives some future directions.

II. Background

In this Section, we introduce the implemented techniques that compose the proposed FR system, as well as the DE optimization algorithm.

A. Preprocessing Techniques

The technique proposed by [16], namely *Wavelet-based Illumination Normalization* (WBIN), has been successfully applied in the illumination normalization problem due to its ability to provide a multi-resolution analysis of the image and the capability of decomposing the image into sub-bands in both time and frequency domains. In the 2-Dimensional Discrete Wavelet Transform (2D DWT), the image is represented in terms of translations and dilatations of a scaling function and a wavelet function using a 2D filter bank consisting of low-pass and high-pass filters [16]. The different wavelet families make different trade-offs between how compactly the basis functions are localized in space and how smooth they are, which are biorthogonal (bior), Coiflets (coif), Daubechies (db), Haar, reversed biorthogonal (rbio) and Symlets (sym). Each family contains several variations according to its smoothness [17]. High pass filtering produces detail information (such as edges) and low pass filtering with scaling produces coarse approximations. The 2D-DWT produces four components, which are approximation (LL), horizontal (HL), vertical (LH), and diagonal (HH), as shown in Figure 1.

For the illumination normalization procedure, histogram equalization is performed in the approximation coefficients (LL) to achieve contrast enhancement. The fine details in the image can be controlled by multiplying each element of the detail coefficient matrices (HL, LH, and HH) with a scale factor (*Sf*). Then, the enhanced image is reconstructed using the inverse wavelet transform. The considered parameters

Fig. 1. DWT decomposition levels. Left: original image. Middle: Decomposed components. Right: Components specification.

within this technique are: the wavelet family (wf) that can assume 105 possible configurations and the scaling factor (sf) that can assume any value greater than 0.

The *Gamma Intensity Correction* [18], also known as power-law transformation, can control overall brightness of an image by changing the parameter γ. Its formula is shown in Equation 1.

$$O(u,v) = c \times I(x,y)^{\gamma} \tag{1}$$

where *O(u,v)* is the gamma corrected image, *c* is a constant, *I(x,y)* is the input image and γ is the correction factor. The only parameter to adjust within this technique is the parameter γ. If it assumes values lower than 1 will shift the image towards the darker end of the spectrum while γ values greater than 1 will make the image brighter.

The *Gaussian Blur* uses a 2D convolution operator with a Gaussian distribution to blur and remove details and noises from the image [19]. The 2D Gaussian distribution has the form as shown in Equation 2.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{2}$$

where σ is the standard deviation of the distribution. The parameters within this technique are: Gaussian kernel size, which must assume positive and odd values, and Gaussian kernel standard deviation in X (sdX) and Y (sdY) directions [20].

An edge is a set of pixels in the image where intensity abrupt changes, representing the shape of face, eyes, nose and other discriminant features. In [21], a review of edge detection techniques was performed which featured the *Laplacian Edge Detection* technique with better results for visual perception and edge quantity. The Laplacian $L(x,y)$ of an image with pixel intensity values $I(x,y)$ is given by Equation 3 and can be calculated using the convolution filter shown in Equation 4.

$$L(x,y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \tag{3}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{4}$$

The parameters that have been considered in this technique are: the aperture size used to compute the second-derivative filters which must be positive and odd, a scale factor (sf) for the computed Laplacian values and a delta value that is added to the results prior to storing them in the resulting image [20].

B. Feature Extraction Techniques

The first feature extraction technique included in the FR framework is the 2D-DWT, which was previously introduced. As mentioned in Section II-A, the 2D-DWT produces four sub-bands (Figure 1) containing the approximation component (LL) and three detail components (HL, LH and HH). According to [18], the information in the approximation component have the most discriminant features. Hence, only this sub-band was used. This procedure may have *n* levels of decomposition. The considered parameters within this technique are: the wavelet family (wf) that can assume 105 possible configurations as mentioned in Section II-A and the number of levels of decomposition.

The *Discrete Cosine Transform* (DCT) was also added to the framework. The DCT technique is popular for image processing because of its energy compaction property. It transforms the image from spatial to frequency domain using cosine functions. In frequency domain, the relevant features tend to be concentrated in low-frequency components or at the corner of the spectrum [22]. Equation 5 shows the DCT procedure for an image pixel value of $f(x, y)$ of dimension NxM in the spatial domain. The output image in the frequency domain is represented by $F(u, v)$.

$$F(u,v) = \alpha_u \alpha_v f(x,y) \times \cos\left(\frac{\pi(2x+1)u)}{M}\right) \times \cos\left(\frac{\pi(2y=1)v)}{N}\right) \quad (5)$$

where,

$$\alpha_u \begin{cases} \frac{1}{\sqrt{N}}, u = 0 \\ \sqrt{\frac{2}{N}}, 1 \leq u \leq N-1 \end{cases} \qquad \alpha_v \begin{cases} \frac{1}{\sqrt{M}}, u = 0 \\ \sqrt{\frac{2}{M}}, 1 \leq u \leq M-1 \end{cases}$$

Texture is an important characteristic of the appearance of surfaces, thus an important component of many FR systems. When choosing a texture descriptor, two competing goals should be considered: low computational complexity, and capturing the most representative texture information. According to [23], LBP has some advantages that are: easy to implement, invariance to monotonic illumination changes, and low computational complexity. Its canonical version is used in this work from which the features are extracted directly from the input image. Figure 2 illustrates LBP's procedure with a kernel size (*ksize*) of 3x3.

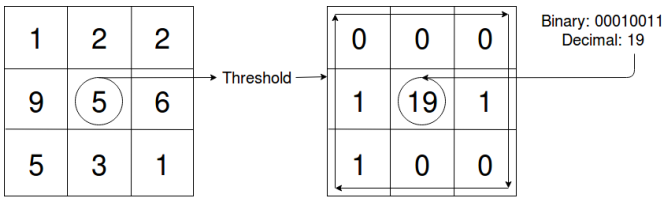


Fig. 2. LBP texture descriptor procedure.

The neighborhood structure is a set of pixels taken from a square neighborhood, which are compared against the value of the central pixel (thresholded) resulting in an 8 bits vector which is converted to decimal. Then, the resultant value is used to represent the pixel.

C. Classification

For the classification task, each image to be classified is compared to each image in the database in terms of similarity returning the most similar one, denoted by M. Then, a match occurs if both images are from the same person. This classification is carried out by the well known k-NN classifier, where $k = 1$ [24]. Equation 6 represents this comparison, where x is the searching image, y is the database image and $DistMeasure$ is the distance measure used. In this work, the distance measure is considered as a parameter, thus, we included it in our optimization task.

$$M = \underset{1 \leq y \leq N}{\operatorname{argmin}} \left(DistMeasure(x,y)\right) \quad (6)$$

Four different distance measures are considered in this work. The first distance measurement implemented is the Manhattan distance (L_1-norm), that is the distance between two points alongside the axis at right angles, as shown in Equation 7. Similarly, the well-known Euclidean distance (L_2-norm) was also employed according to Equation 8. In addition to these, the L_{2SQR}-norm (Equation 9) and L_∞-norm (Equation 10) were also implemented.

$$DistMeasure(x,y) = \sum_{i=1}^{N} |x_i - y_i| \quad (7)$$

$$DistMeasure(x,y) = \sqrt{\sum_{i=1}^{N} (x_i - y_i)^2} \quad (8)$$

$$DistMeasure(x,y) = \sum_{i=1}^{N} (x_i - y_i)^2 \quad (9)$$

$$DistMeasure(x,y) = \max |x_i - y_i| \quad (10)$$

D. Differential Evolution (DE)

Differential Evolution [25] (DE) is a global optimization algorithm that encodes candidate solution vectors (or individuals) in real-valued numbers. Initially, the individuals are generated randomly in the search space. New individuals are created adding the weighted difference between two individuals to a third one, namely target. This routine is called mutation. Then, the target individual is combined with a randomly preselected individual resulting the trial individual and representing crossover procedure. The decision if the trial individual replaces the target in the next generation is based on an objective function and its constraints (fitness). If trial vector is worse than the target's fitness value, the trial is discarded. Otherwise, the trial vector replaces the target. This operation is called greedy selection. This process repeats until a stop criterion is reached. The pseudo-code for DE is shown in Algorithm 1 [26]. DE has a nomenclature that describes the adopted configuration. This takes the form of $DE/x/y/z$, where x represents the solution to be perturbed, such as random or best. The y identifies the number of difference vectors used in the perturbation of x. Last, z is the recombination operator performed such as bin for binomial and exp for exponential [26]. Algorithm 2 shows the procedure to generate a trial solution. In this work, we used the popular configuration DE/rand/1/bin. The population size, the crossover probability (CR) and the mutation factor (F) are some parameters to tune.

Algorithm 1 Pseudo-code for Differential Evolution algorithm [26]

Input: $Population_{size}, Problem_{size}, F, CR$
Output: S_{best}

$Population \leftarrow InitializePopulation(Population_{size}, Problem_{size})$
$EvaluatePopulation(Population)$
$S_{best} \leftarrow GetBestSolution(Population)$
while *stop condition is not reached* **do**
  $NewPopulation \leftarrow \oslash$
  for $P_i \in Population$ **do**
    $S_i \leftarrow NewSample(P_i, Population, Problem_{size}, F, CR)$
    if $Cost(S_i) \leq Cost(P_i)$ **then**
      $NewPopulation \leftarrow S_i$
    else
      $NewPopulation \leftarrow P_i$
    end if
  end for
  $Population \leftarrow NewPopulation$
  $EvaluatePopulation(Population)$
  $S_{best} \leftarrow GetBestSolution(Population)$
end while
return S_{best}

III. Proposed Methodology

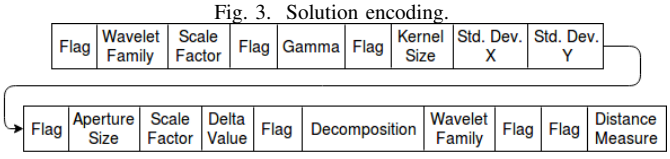
In this section, we detail each step of the proposed approach. The optimization stage is carried out by the DE algorithm. As well as its parameters, techniques are encoded in the solution vector (shown in Figure III) which consists of the WBIN, the GGIC, the Gaussian Blur, the Laplacian Edge Detection, the DWT, the DCT, the LBP, and the distance metric of k-NN. The parameter $flag$ represents if the technique will be used or not. As DE operates in continuous domain, the parameter $flag$ is discretized to 0 or 1. If the $flag$ is greater than 0.5 the value is set to 1, then the specific technique will be used by that candidate solution. Otherwise, it will not be used in the optimization process. Likewise, some parameters that should be tuned have bound constraints and, if a parameter assumes a value out of its bounds, then the value will be adjusted to the closest bound. It is worth to point out here that the distance measure used in classification stage is encoded in the solution vector itself.

Algorithm 2 Pseudo-code for $NewSample$ function [26]

Input: $P_o, Population, Problem_{size}, F, CR$
Output: S

repeat
  $P_1 \leftarrow RandomMember(Population)$
until $P_1 \neq P_o$
repeat
  $P_2 \leftarrow RandomMember(Population)$
until $P_2 \neq P_o \vee P_2 \neq P_1$
repeat
  $P_3 \leftarrow RandomMember(Population)$
until $P_3 \neq P_o \vee P_3 \neq P_2 \vee P_3 \neq P_1$
$CutPoint \leftarrow RandomPosition(Problem_{size})$
$S \leftarrow 0$
for i**to**$Problem_{size}$ **do**
  if $i \equiv CutPoint \wedge Rand() < CR$ **then**
    $S_i \leftarrow P_{3_i} + F \times (P_{1_i} - P_{2_i})$
  else
    $S_i \leftarrow P_{o_i}$
  end if
end for
return S

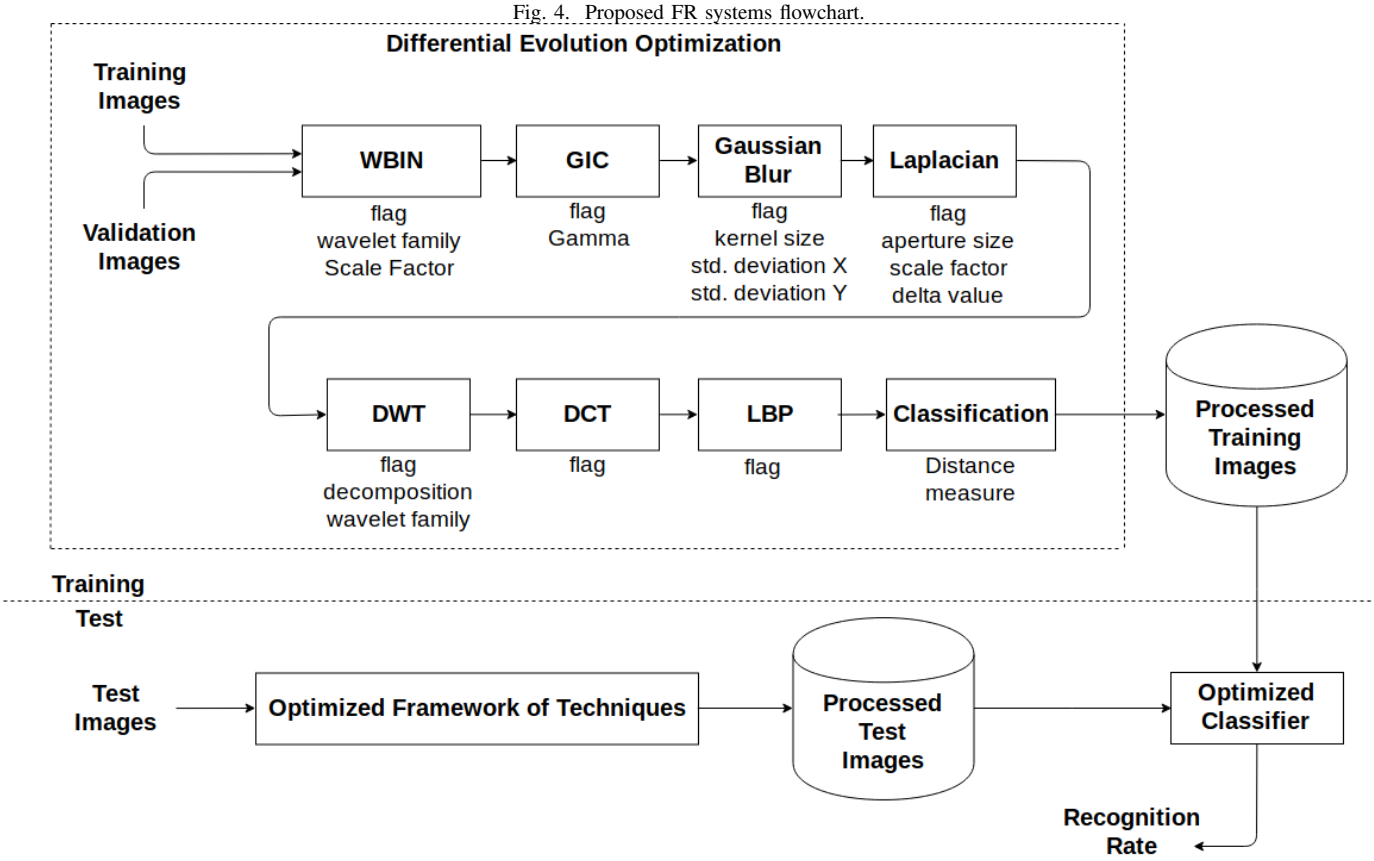
Fig. 3. Solution encoding.



The flowchart representing the proposed FR system is presented in Figure 4. The DE algorithm starts initializing the population with random solution vectors and, as the generation cycle begins, the solutions are modified accordingly to the DE/rand/1/bin strategy (Section II-D).

The dataset should be split into three subsets: training, validation and test. The training subset is composed by the base images, which will be used to classify the validation and test images. The validation and test subsets are composed by the remaining images divided into two parts equally.

Fig. 4. Proposed FR systems flowchart.



The training subset is used to evaluate the candidate solutions of DE based on validation subset and the test subset is used in fact to test the optimized model. In both training and test stages the evaluation (or fitness evaluation) is done based on the recognition rate. A predefined number of iterations is established as a stopping criterion. Consequently, when it is reached, the setting that achieved the best recognition rate is selected as an optimized solution. To test the classification model generated in training and validation stages, the test images set is transformed using the optimized framework of techniques and thus, generating the processed test images dataset. Finally, the processed test images are classified by the optimized k-NN classifier using the processed training images dataset as a reference.

IV. Experiments and Results

To evaluate our approach, the well known *Yale Extended B database* is employed. This database consists of 38 subjects under different conditions varying pose and illumination (9 poses x 64 illuminations). As our purpose is to compensate the illumination variation, only the frontal face images were selected. This database is commonly divided into five subsets according to the angle between the light source direction and camera axis [5]. Table I shows the angle variations, as well as the number of subjects, and Figure 5 illustrates the sample images from subsets 1 to 5. The well-illuminated images from S1 are used as training set. The remaining subsets are randomly divided into half as a validation and a test sets.

TABLE I
Five subsets according to the light angle source directions.

	S1	S2	S3	S4	S5
angle	0^o-12^o	13^o-25^o	26^o-50^o	51^o-77^o	$>78^o$
number of images	6	12	11	12	16

Fig. 5. Example of Yale Extended B database subsets from S1 to S5, respectively.

The DE parameters are the population size, the crossover probability (CR), and the mutation factor (F). For this case of study, we set population size empirically to 20, CR to 0.9 and F to 0.8, both based on literature [26]. The stop criterion was defined empirically as 100 iterations. The experiments were run 10 times with confidence interval of 95% [27].

All experiments were run in an Intel® Core i7-4770 computer with 16 GB RAM memory using programming language python with multiprocessing library for DE parallelism. Parallelism is achieved distributing the fitness evaluation of DE population in 7 out of 8 available processing cores.

In Table II, we present the techniques chosen by the optimization step for each experiment, as well as their specific parameters. From the experimental analysis based on the optimization approach, the GIC and Laplacian Edge Detector techniques were chosen in every final solution, representing an important role in the illumination compensation task. In most of the experiments, the L_{2sqr} distance measure was chosen, although same results are accomplished with L_1-norm. In addition to this, the LBP texture descriptor was never chosen. LBP as well as all other techniques proposed in the framework somehow attempt to compensate illumination variation. According to the similar recognition rates obtained from the different sets of techniques, it can be possibly said that the problem space in this work is multi-modal, i.e. where different possible solutions lead to the almost same values.

The experiments are divided into to two parts: training and test. The recognition rates for training and test stages in each experiment are reported in Table III. It is possible to notice that similar results are obtained in both stages, which means that precision on training reflects well on test. In addition to this, the low standard deviation values imply the low sensitiveness to random variables in an optimization approach.

The average training time was 5 hours and 48 minutes with standard deviation of 1 hour and 11 minutes. However, after the training task is completed with the definition of the optimized settings, the average time for testing an image is 0.0291 seconds with a standard deviation of 0.0052.

Among all 10 runs, the experiment 8 stood out with 100% overall recognition rate on test. In order to promote a fair comparison with the related works presented in Section I, the average recognition rate of training and test obtained using the optimized settings in experiment 8 was used. Table IV presents the FR rates obtained for each subset and its average recognition rate that belong to the related works and to our approach, as well as the results achieved by using the literature settings which refer to the recognition rates achieved by using the most common parameters found in the literature using the optimized techniques defined by our approach. They are: for the wavelet family, which is used in DWT and WBIN techniques, the Haar wavelet is the most common. In DWT, only one level of decomposition was employed. The scale factor was set to 3 based on Manikantan et al [18] work. The parameter γ was set to 2. For Gaussian Blur, kernel size was set to 3x3 and the standard deviation based on kernel size. For Laplacian Edge Detector, the kernel size was also 3x3 and no scale or delta value were employed.

In Table IV we present the experimental results of the related works. These works achieved high recognition rates with their proposed methodologies. Most of them presented a precision of 100% in subsets 2 and 3, which have better illumination. Consequently, this condition facilitates the illumination compensation. Meanwhile, the compensation on the subsets 4 and 5 are not that trivial since the present the darkest images where the face is barely visible. However, most works achieved recognition rates higher than 98% for both subsets. Among these, we can highlight the wGAP and our approach. Considering the average recognition rate, it can be said that all approaches are competitive since they have achieved rates higher than 98.8%. It is important to mention that our method stood out with 99.95% recognition rate. In the experiment employing literature settings, the values of parameters for the optimized techniques were set to the literature recommendation. Though we have used the optimized techniques defined by our approach with literature settings, the recognition rates are not competitive. From this, we can deduce that how important is the parameter tuning within the optimized techniques and how efficient was the DE in optimizing the parameters.

Finally, Figure 6 presents the average convergence and diversity charts of DE optimization task for all experiments. The convergence chart represents the recognition rate of the best solution at each iteration from which we can visualize the algorithm's convergence. Close to the 40th iteration, the best solution improves in a slow rate but continuously improving when the stop criteria is reached. However, in most experiments, almost maximum recognition rate is reported justifying 100 iterations as stop criteria. The population's diversity chart is obtained using the formula PCM_{13} presented in Corriveau's work [28]. Due to the multi-modal search space nature of the problem, the diversity in our experiments remained high during all optimization step as we can see in the diversity chart. Therefore, the use of different strategies with different values can lead to similar (or same) recognition rates.

V. Conclusions and Future Work

To the best of our knowledge, from early studies to the most recent ones, it can be said that there is no generic technique that is immune to all illumination conditions, indicating that the illumination compensation is still an open research topic. Hence, in this paper, we proposed an adjustable FR framework with the aid of the DE optimization algorithm. This approach implements several illumination compensation techniques for preprocessing and for feature extraction. The main aspect of this approach is the definition of which set of techniques and its ideal parameters values should be used for the FR using DE. Hence, the DE is responsible for selecting techniques, as well as to tune their parameters which, in turn, can lead to the best FR rates. The well known Yale Extended B face dataset was used to evaluate the proposed approach, since it permits to compare the results with relevant works found in the literature.

In this work, we have shown that our approach is suitable for compensating illumination variation using Extended B face dataset and it presents competitive results in comparison with the works proposed in the literature. The subsets 2 and 3 of the dataset have better illumination compared to 4 and 5, therefore it is trivial to compensate them. On the

TABLE II
SETTINGS FROM TRAINING STEP

Experiment	WBIN		GIC	Gaussian Blur			Laplacian Edge Detector			DWT		DCT	LBP	Distance Measure
	wf	sf	γ	$Kernel$	sdX	sdY	$Kernel$	sf	$Delta$	wf	$Level$			
1	-	-	2.5108	5	5	5	7	2.7198	3	sym9	1	-	-	L_{2sqr}
2	-	-	3	-	-	-	13	1.3637	2.8603	coif12	2	-	-	L_{2sqr}
3	db5	1.2606	2.8723	3	1	3.5963	15	2.0067	1	-	-	Yes	-	L_{2sqr}
4	-	-	2.0171	-	-	-	15	1.8360	1.9107	sym10	1	-	-	L_{2sqr}
5	-	-	3	-	-	-	15	3	2.1105	-	-	-	-	L_{2sqr}
6	-	-	2.3487	9	1	3.3040	13	3	1	-	-	-	-	L_{2sqr}
7	-	-	1.7961	-	-	-	15	2.8252	3	-	-	Yes	-	L_{2sqr}
8	bior3.7	0.6483	2.2690	15	1.7642	5	7	1	3	bior1.1	1	Yes	-	L_1-norm
9	bior1.1	0.0343	2.7519	3	5	1	7	1	1	-	-	-	-	L_1-norm
10	-	-	2.9877	3	5	2.4727	15	1.2302	1.1656	-	-	-	-	L_2-norm

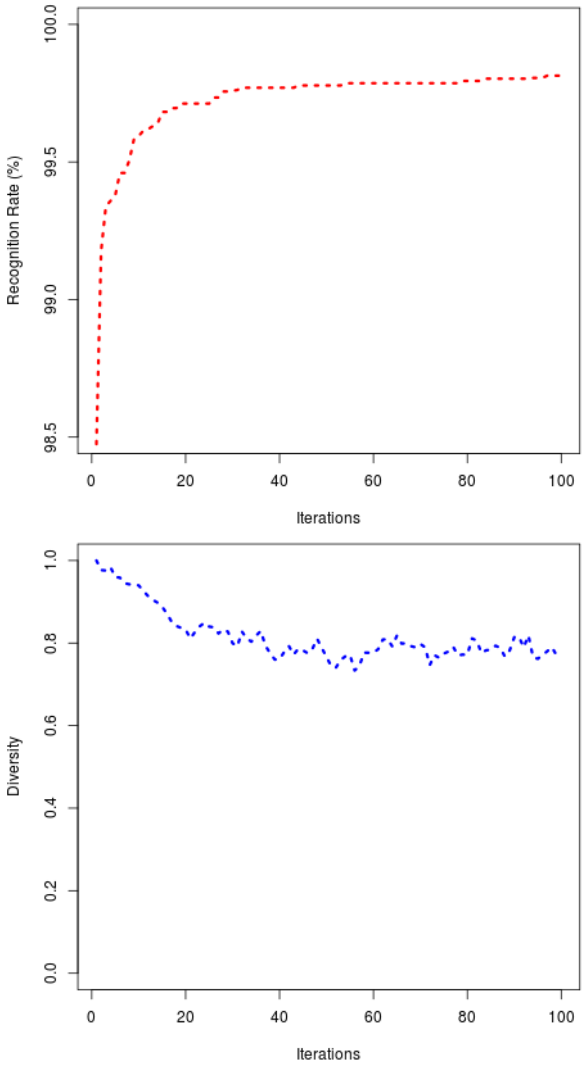
TABLE III
RECOGNITION RATES (%) FOR EACH EXPERIMENT.

	Recognition Rates (%)	
Experiment	**Training**	**Test**
1	99.80	99.80
2	99.80	99.58
3	99.91	99.23
4	99.78	99.72
5	99.80	99.61
6	99.91	99.50
7	99.89	99.61
8	99.91	100
9	99.80	99.52
10	99.69	99.91
Average	99.83 $\pm$ 0.07	99.65 $\pm$ 0.22

TABLE IV
RECOGNITION RATES (%) OF PROPOSED APPROACH AND RELATED WORKS.

Method	S2	S3	S4	S5	All
DTCWT [7]	100	100	98.68	99.03	99.42
resonance [8]	100	100	98.68	99.16	99.46
DWT [10]	100	100	98.86	99.35	99.55
WT-Retinal [6]	99.78	99.44	98.47	97.78	98.86
fuzzy + LFDCT [9]	-	100	98.87	98.06	-
wGAP [11]	100	100	99.51	100	99.82
NMR [12]	-	-	90.2	47.9	-
CZT and Goertzel [13]	-	-	-	96.25	-
Our method	**100**	**100**	**100**	**99.83**	**99.95**
Literature settings	84.20	68.65	21.48	48.02	55.59

Fig. 6. Convergence and Diversity charts for DE optimization task, respectively.



other hand, the challenge remains on the subsets 4 and 5 in which we can note the presence of the dark and partially illuminated images. According to the results obtained from the experiments using these subsets, our approach reached almost 100% of recognition rate. It strongly demonstrates the robustness of our proposed. It was possible to realize in our experiments that optimizing the set of techniques for illumination compensation and for feature extraction as well as their parameters is important to enhance the precision in a FR system. Hence, the DE algorithm has shown to be suitable for this kind of task. This is evident when taking a look at the experiment using literature settings in which the use of the same set of techniques with optimized parameter values achieved much higher recognition rate.

Also, through the experiments, it can be observed that the present FR problem is highly multi-modal since the different sets of techniques can achieve similar (or almost the same) results.

In any meta-heuristic, including the DE, the adjustment of parameters plays an important role and it can directly affect the quality of the final solution. This adjustment can influence

the behavior of the algorithm during the search to find the promising regions in the solution space of the problem [29]. Also, the ideal values for the DE parameters may change for different datasets. Hence, we strongly believe that the use of on-line control of parameters in the DE as a future research.

Another future research direction can be the use of other datasets with occlusion, variation in pose, and other challenges to ensure the robustness of our approach. To attack these specific issues in different datasets, other techniques could be added to the present FR framework.

ACKNOWLEDGMENT

This research work is supported by the federal government agency CAPES and the Santa Catarina state government agency FAPESC from Brazil.

REFERENCES

[1] X. Liu, C. Charrier, M. Pedersen, and P. Bours, "Study on color space for the performance of degraded face image recognition," in *Media Watermarking, Security, and Forensics Conference, IS&T Electonic Imaging 2018*, 2018.

[2] X. Liu, L. Lu, Z. Shen, and K. Lu, "A novel face recognition algorithm via weighted kernel sparse representation," *Future Generation Computer Systems*, vol. 80, pp. 653–663, 2018.

[3] W. Zhang, Z. Xi, J.-M. Morvan, and L. Chen, "Improving shadow suppression for illumination robust face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018.

[4] G. Wen, H. Chen, D. Cai, and X. He, "Improving face recognition with domain adaptation," *Neurocomputing*, vol. 287, pp. 45–51, 2018.

[5] M. A. Ochoa-Villegas, J. A. Nolazco-Flores, O. Barron-Cano, and I. A. Kakadiaris, "Addressing the illumination challenge in two-dimensional face recognition: a survey," *IET Computer Vision*, vol. 9, no. 6, pp. 978–992, 2015.

[6] Y. Cheng, Z. Li, and L. Jiao, "Illumination normalization for face recognition under extreme lighting conditions," in *International Conference on Intelligent Science and Intelligent Data Engineering*. Springer, 2012, pp. 491–497.

[7] A. Baradarani, Q. J. Wu, and M. Ahmadi, "An efficient illumination invariant face recognition framework via illumination enhancement and DD-DTCWT filtering," *Pattern Recognition*, vol. 46, no. 1, pp. 57–72, 2013.

[8] A. Baradarani and Q. J. Wu, "Illumination invariant human face recognition: frequency or resonance?" in *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*. IEEE, 2013, pp. 1–6.

[9] V. P. Vishwakarma, "Illumination normalization using fuzzy filter in dct domain for face recognition," *International Journal of Machine Learning and Cybernetics*, vol. 6, no. 1, pp. 17–34, 2015.

[10] H. Hu, "Variable lighting face recognition using discrete wavelet transform," *Pattern Recognition Letters*, vol. 32, no. 13, pp. 1526–1534, 2011.

[11] X. Zhao, Z. He, S. Zhang, S. Kaneko, and Y. Satoh, "Robust face recognition using the gap feature," *Pattern Recognition*, vol. 46, no. 10, pp. 2647–2657, 2013.

[12] J. Yang, L. Luo, J. Qian, Y. Tai, F. Zhang, and Y. Xu, "Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 1, pp. 156–171, 2017.

[13] K. Varadarajan, P. Suhasini, K. Manikantan, and S. Ramachandran, "Face recognition using block based feature extraction with czt and goertzel-algorithm as a preprocessing technique," *Procedia Computer Science*, vol. 46, pp. 1458–1467, 2015.

[14] S. Das, S. S. Mullick, and P. N. Suganthan, "Recent advances in differential evolution–an updated survey," *Swarm and Evolutionary Computation*, vol. 27, pp. 1–30, 2016.

[15] A. Georghiades, P. Belhumeur, and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intelligence*, vol. 23, no. 6, pp. 643–660, 2001.

[16] S. Du and R. Ward, "Wavelet-based illumination normalization for face recognition," in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, vol. 2. IEEE, 2005, pp. II–954.

[17] A. Graps, "An introduction to wavelets," *IEEE computational science and engineering*, vol. 2, no. 2, pp. 50–61, 1995.

[18] K. Manikantan, M. S. Shet, M. Patel, and S. Ramachandran, "DWT-based illumination normalization and feature extraction for enhanced face recognition," *International Journal of Engineering & Technology*, vol. 1, no. 4, pp. 483–504, 2012.

[19] R. C. Gonzalez and R. E. Woods, "Digital image processing, pretice hall, 2008," *ISBN-13*, 2008.

[20] *The OpenCV Reference Manual*, 3rd ed., Itseez, February 2018.

[21] A. Kushwah, K. Gupta, A. Agrawal, G. Jain, and G. Agrawal, "A review: Comparative study of edge detection techniques," *International Journal*, vol. 8, no. 5, 2017.

[22] S. Rao and M. B. Rao, "A novel triangular dct feature extraction for enhanced face recognition," in *Intelligent Systems and Control (ISCO), 2016 10th International Conference on*. IEEE, 2016, pp. 1–6.

[23] L. Liu, P. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Local binary features for texture classification: Taxonomy and experimental study," *Pattern Recognition*, vol. 62, pp. 135–160, 2017.

[24] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21–27, 1967.

[25] R. Storm, "Price, k.,," *Minimizing the Real Functions of the ICEC 96 contest by differential Evolution*, pp. 842–844, 1997.

[26] J. Brownlee, *Clever algorithms: nature-inspired programming recipes*. Jason Brownlee, 2011.

[27] D. J. Lilja, *Measuring computer performance: a practitioner's guide*. Cambridge university press, 2005.

[28] G. Corriveau, R. Guilbault, A. Tahan, and R. Sabourin, "Review of phenotypic diversity formulations for diagnostic tool," *Applied Soft Computing*, vol. 13, no. 1, pp. 9–26, 2013.

[29] R. S. Parpinelli, G. F. Plichoski, R. S. da Silva, and P. H. Narloch, "A review of techniques for on-line control of parameters in swarm intelligence and evolutionary computation algorithms," *International Journal of Bio-Inspired Computation*, Accepted for publication.