Scientific African 22 (2023) e01964

Contents lists available at ScienceDirect

Scientific African

journal homepage: www.elsevier.com/locate/sciaf

On multiple imputation-based reconstruction of degraded faces and recognition in multiple constrained environments

Joseph Agyapong Mensah [a,b], Eric Ocran [a], Louis Asiedu [a,*]

[a] Department of Statistics & Actuarial Science, School of Physical and Mathematical Sciences, University of Ghana, Legon, Accra, Ghana
[b] Department of Computer Science, Ashesi University, No. 1 University Avenue, Berekuso, Eastern Region, Ghana

ARTICLE INFO

Editor name: Aboul Ella Hassanien

Keywords:
Multiple imputation methods
Face recognition
Multiple constraints
Occlusions
Discrete cosine transform

ABSTRACT

Recognition of degraded frontal face images acquired under occlusion constraints remain challenging despite the plethora of reconstruction mechanisms. Though recent works have leveraged on some imputation mechanisms in this regard, their robustness in multiple constrained environments may not be guaranteed and may be affected by the choice of pre-processing mechanism. This paper proposes enhancement mechanisms that augment or complement the use of three (3) multiple imputation mechanisms for facial reconstruction in the presence of multiple constraints (10% and 20% occlusions and varying facial expressions). Specifically, we propose the use of a Discrete Cosine Transform-based (DCT) denoising or a Discrete Wavelet-based denoising following Histogram Equalization (HE-DWT) of the reconstructed face images prior to recognition. Experimental results showed that the proposed augmented enhancements improved significantly the recognition rates (90.63% & 91.15% and 86.98% & 85.94% for DCT and HE-DWT at 10% and 20% occlusion levels respectively for Missforest de-occluded face images) as compared with DWT in recognizing degraded frontal face images under moderately low levels of occlusions and varying expressions.

Introduction

With the birth of the digital age, access to digital images for planning and decision-making has dramatically improved. For instance, people are always seen taking photographs using their mobile and photographic devices, satellite images are picked up by remote sensors every microsecond for weather forecasting [1] and monitoring, surveillance and access-control systems constantly report human activities over 24-hour periods and raise red flags, when necessary, for national security interventions. For the latter, the use of biometrics (face images) for identification or verification is widespread due to the non-invasive nature of the face acquisition systems [2]. In health, [3] achieved characterization of abnormalities in breast cancer images using a nature-inspired metaheuristic optimized convolutional neural networks model. Also, [4] proposed multi-stage faster RCNN-based iSPLInception (MFRCNN-iSPLI) method to classify both benign and malignant tumors efficiently. The proposed classifier was used to overcome the overfitting problem associated with Convolutional neural networks (CNNs) when used to classify melanoma skin cancer images.

Their methods' accuracy (95.82%), precision (96.85%), recall (96.52%), and F1 score (0.95%) values were calculated, and the results are compared with the existing methods such as CNN, hybrid DL, Inception v3, and VGG19. The output analysis of each measure verified the prediction and classification effectiveness of their method.

In automatic face recognition systems, face images are acquired and analyzed for identification or verification purposes. The process involves four key stages which are: image acquisition using digital devices such as cameras; preprocessing of acquired

* Corresponding author.
E-mail address: lasiedu@ug.edu.gh (L. Asiedu).

https://doi.org/10.1016/j.sciaf.2023.e01964
Received 15 July 2023; Received in revised form 23 October 2023; Accepted 31 October 2023
Available online 7 November 2023

images to enhance their quality; selection of unique features that characterize the image by employing techniques that minimize information loss such as Principal Component Analysis (PCA); and image classification or recognition which involves matching the unique features of a probe face to those of known faces in the gallery, by means of a classifier such as the Euclidean or City Block distance (L1).

A key issue of concern in automatic face recognition is the sub-optimal performances of recognition algorithms as a result of the acquisition and use of poor-quality images acquired from uncontrolled environments, which affect the quality of features used for classification [5].

The effects of occlusions and varying facial expressions are known to be more pronounced, especially for recognition systems based on Principal Component Analysis (PCA) features [6]. The presence of random occlusions, particularly, lead to the loss of pertinent facial features or pixel values, thus, resulting in incomplete feature matrix representations of the affected images. Under this circumstance, the estimation of the vector of mean values and covariance matrix becomes difficult [7]. To overcome this challenge, one may consider estimating the covariance matrix, for example, from only the non-missing pixels in the affected face images. However, a covariance matrix which is not necessarily semi-positive definite may result, and PCA could give erratic results.

Some approaches to handling the occlusion challenges involve learning occlusion robust features for recognition. In [8], the authors proposed an approach invariant to occlusion in face recognition based on 2D PCA. They first detected occluded parts by applying a combined k-NN and 1-NN classifier. Then, they conducted partial matching only on the non-occluded parts after eliminating the occlusion effect. This approach has the advantage of using a few features for face representation and, therefore may be less computationally expensive. However, with moderate levels of occlusions, such an approach may be deficient.

In [9], the authors proposed a robust feature extraction under partial occlusions by combining Gabor transform, Local Binary Pattern (LBP) and local histograms. However, according to [10], Local Gabor Binary Pattern has a high feature dimension and therefore has a high computational requirement.

In [11], the authors also achieved robustness to occlusions by using a point set matching method that utilizes geometric and texture features. However, according to [10], the method employed lacks flexibility and reduces recognition performance when the occlusion area is large. One may consider learning occlusion robust features using deep learning techniques for classification. However, the complexity of such architectures may offset the gains made [12]. Image reconstruction, particularly using multiple imputation techniques, have been considered in recent years due to their flexibility and ease of implementation [13].

In [13], the relative merits of using the MICE multiple imputation algorithm to resolve 5% random occlusions in test faces were explored. The authors demonstrated that the use of the MICE-reconstructed faces for recognition resulted in lower recognition distances compared with the use of the occluded test faces for recognition. From their study, the MICE algorithm was recommended for reconstruction at 5% occlusion rate.

[14] proposed an efficient technique for face detection from still images from public dataset AR face dataset and Color FERET dataset captured under occlusion and non-uniform illumination (multiple constraints). They presented a face detection technique using a combination of YCbCr, HSV and $L \times a \times b$ color model.

[15] proposed a novel face de-occlusion model based on face segmentation and 3D face reconstruction, which is robust to arbitrary kinds of face occlusions. The proposed model consists of a 3D face reconstruction module, a face segmentation module, and an image generation module. They demonstrated that, with the face prior and the occlusion mask predicted by the first two, respectively, the image generation module can faithfully recover the missing facial textures.

In recent works, the performance of PCA-based algorithms were evaluated under multiple constraints (occlusions and varying facial expressions) for recognition [16,17]. Specifically in [17], numerical evaluation of their study algorithm gave reasonable average recognition rates of 77.31% and 76.85% for left and right reconstructed face images with varying expressions, respectively.

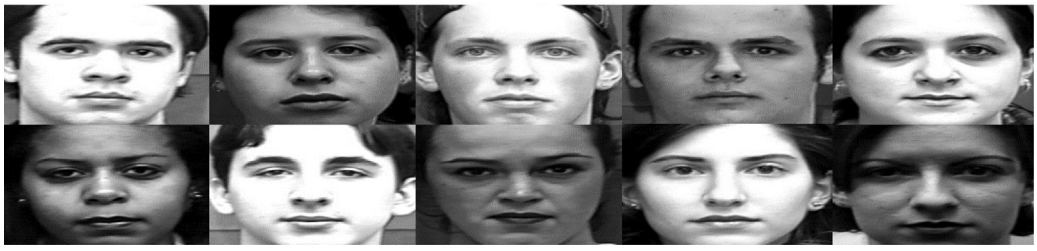
Although the above results are appreciable, it is evident that the PCA-based algorithms are still challenged when used for recognition under multiple constraints (presence of more than one environmental constraints such as occlusion, expressions, lighting etc.).

In this paper, we propose appropriate enhancement mechanisms that augment three (3) (MICE, MissForest and RegEM) multiple imputation algorithms for reconstruction in PCA-based facial recognition when test faces are acquired with 10% or 20% occlusions and varying facial expressions. The MICE and Missforest multiple imputation mechanisms were chosen in quest to assess the imputation performance of parametric multiple imputation mechanisms (eg. MICE), non-parametric imputation mechanisms (eg. MissForest) but the RegEM mechanism was chosen because it is well-known for having desired qualities such as minimal cost per iteration, simplicity of implementation, and reliable global convergence. The proposed enhancements leverage on transform domain-based filtering such as Discrete Cosine Transforms (DCT) and Contrast Enhancement using Histogram Equalization (HE).

It has been shown in [18] that image filtering mechanisms can further enhance the quality of degraded images, thereby improving the performance of face recognition algorithms. However, specifying the right combination of image enhancement mechanisms to achieve the optimum recognition rate is a challenging task.

The remaining parts of the paper are organized as follows: Section (Materials and Methods) presents the data acquisition process, an overview of the multiple imputation methods used for image reconstruction, adopted enhancement mechanisms, the research design and implementation. In Section (Results and Discussion), we evaluate the performance of recognition modules under the proposed augmented enhancement mechanisms. Section (Conclusion and Recommendation) summarizes the overall achievement of the study with some recommendations and direction for future studies.

(a) Subjects in the JAFFE database



(b) Sample of subjects in the CKFE database

Fig. 1. Train-image database.

Materials and methods

The performances of facial recognition algorithms are benchmarked using standard databases of face images. For this work, the Japanese Female Facial Expression (JAFFE) and the Cohn Kanade AU-coded Facial expression (CKFE) databases were used.

- **Dataset 1:** The Japanese Female Expression (JAFFE) dataset contains the face images of ten (10) Japanese female subjects captured with seven principal emotions (neutral, angry, disgust, fear, sad, surprise and happy).
- **Dataset 2:** The Cohn-Kanade AU-Coded Facial Expression (CKFE) database contains face images of twenty-two (22) subjects also captured along the above seven principal emotions.

Thirty-two (32) neutral expressions of subjects in both databases (Dataset 1 (JAFFE) and Dataset 2 (CKFE)) were captured into the train images (following face detection and cropping) for training the study algorithm. Fig. 1 shows the face images of some subjects in the train image database.

The face images of subjects acquired under other principal expressions were synthetically occluded at 10% and 20% occlusion rates and captured into test image databases 1 and 2, respectively, following face detection and cropping. The images in these two test image databases are, thus, characterized (constrained) by occlusions of varying magnitudes as well as varying facial expressions and are hereafter referred to as multiple-constrained test faces. Figs. 2 and 3 show the multiple-constrained faces of some subjects in the test image databases 1 and 2 respectively.

These multiple-constrained face images (Figs. 2 and 3) are subsequently reconstructed using the MICE, MissForest and RegEM imputation techniques and captured into different test image databases. Test image database 3 contains the MICE, MissForest and RegEM reconstructed images from the 10% occluded images while test image database 4 contains the MICE, MissForest and RegEM reconstructed images from the 20% occluded faces.

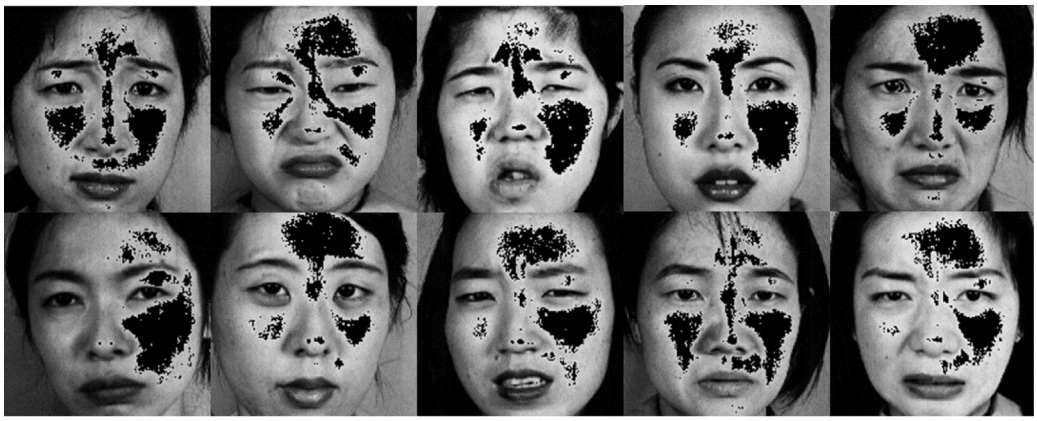
Reconstruction via multiple imputation methods

Multiple imputation by chain equations (MICE)

The MICE algorithm [19] is a parametric (regression-based) multiple imputation algorithm. As with all multiple imputation algorithms, MICE creates m copies (via Gibbs Sampling) of each missing pixel intensity value in an occluded face image data matrix and obtains single imputations by pooling. MICE uses the distribution of observed pixels in the data matrix corresponding to a multiple-constrained face to obtain plausible values for the missing pixels. This is achieved by fitting conditional (regression) models for each variable (column) based on all other variables. This is iterated until a stopping criterion is met.

MissForest

The MissForest imputation method is a non-parametric imputation algorithm based on random forests [20]. Random forest is an extension of classification trees, which are predictive models that recursively subdivide the data based on values of the predictor variables [21]. It is a supervised learning algorithm that grows and combines multiple decision trees to create a forest. Random forest



(a) Subjects in the JAFFE database (10% random occlusions with sad expressions)

(b) Subjects in the CKFE database (10% random occlusions with surprise expressions)

Fig. 2. Sample faces of subjects with 10% synthetic occlusions, showing sad and surprise expressions.

models are able to handle well-mixed types of data (categorical and continuous) and address complex interactions and nonlinearities. In missForest, the imputation of missing values is carried out by regressing each variable (column) in turn on all other variables, just as in MICE, but predictions for the dependent variable are obtained from the fitted forest as the average of the outputs of all trees. This is iterated until a stopping criterion is met.

Regularized Expectation Maximization (RegEM)

The Expectation Maximization (EM) algorithm is based on the concept of maximum likelihood for estimating unknown population parameters. By modeling it as a prediction (regression) task, the EM technique is used to compute parameter estimates in the presence of missing data [22]. The EM algorithm is well-known for having desired qualities such as minimal cost per iteration, simplicity of implementation, and reliable global convergence [23]. The Regularized EM technique is a penalized regression approach proven to work well, especially when the design matrix is ill-conditioned. L2 Regularization or ridge regression was used in this study, and the regularization parameter was calculated using the generalized cross-validation technique [24]. The missing pixels in the occluded facial photographs were imputed as follows:

Starting with an initial mean and covariance matrix estimate,

- Regress each column with missing pixel intensity values on all other columns to get the (ridge multiple) regression parameters for each row of the image matrix of an occluded face with missing pixel intensity values.
- Then obtain the conditional expectation values as a product of observed pixels and the estimated regression coefficients and use it to fill in the missing pixel intensity values.
- Re-estimate the mean and covariance matrix, then continue steps 1–3 until convergence is achieved.

Figs. 4 and 5 show the de-occluded face images using MICE, RegEM and the Missforest algorithms at 10% and 20% occlusion rates respectively.

(a) Subjects in the JAFFE database (20% random occlusions with sad expressions)

(b) Sample of subjects in the CKFE database (20% random occlusions with surprise expressions)

Fig. 3. Sample faces of subjects with 20% synthetic occlusions, showing sad and surprise expressions.

Research design

First, frontal face images (having neutral expressions with no occlusions as shown in Fig. 1) in the train-image database are passed to the recognition module, denoised in the DCT-based transform domain or DWT-based Transform domain following contrast enhancement via Histogram Equalization. The enhanced images are then sent to the extraction chamber and PCA-based feature extraction is carried out. The resultant features are stored in memory as a created knowledge for recognition. When a test image (form either database 3 or 4) reaches the recognition module, it is also enhanced, as in the case of the train images, and its discriminative PCA-based features are extracted. The extracted features are passed on to the recognition unit and matched with the stored knowledge created from the train images for recognition. The City Block distance (L1) is used for classification and a closer match is obtained when the corresponding distance is minimum. We note that only one multiple imputation method is employed for facial reconstruction at a time and that only one test face image is passed to the recognition module along with the train images at a time. The design of the recognition module is presented in Fig. 6.

Preprocessing

Most image preprocessing procedures aim to enhance the quality of an image to obtain more details for further analysis while minimizing possible distortions or degradation. The main preprocessing mechanisms adopted in this study are Discrete Cosine Transform (DCT) denoising, Discrete Wavelength Transform (DWT), Histogram Equalization (HE) and Mean Centering (MC).

Discrete Cosine Transform (DCT)

The Discrete Cosine Transform is a transform-based image enhancement technique. As such, it converts an image from the spatial domain to the frequency domain for the purposes of image enhancement, compression or information hiding/embedding. The DCT is an invertible linear transform whose kernel is defined by a set of complete orthogonal discrete cosine functions [25]. The DCT

(a)

(b)

Fig. 4. Sample (10%) de-occluded faces acquired with varying expressions (Test-image database 3).

decomposes an image into a series of cosine harmonic functions [26]. Thus, the input image is expressed as a linear combination of weighted sinusoidal basis functions that relate to its frequency components [27]. A remarkable feature of DCT is that after transformation, the resultant image has good time–frequency property and the visual quality is greatly enhanced [28]. DCT aim to separate the gross features key to visualizing an image from the fine detail, which are less essential and sometimes imperceptible by the eye. Due to its high energy compaction, the DCT has found wide applicability in pattern recognition and data compression. The DCT is closely related to the Discrete Fourier Transform. For a sequence of length N, a corresponding $2N$-point sequence is obtained by mirroring the former. The DCT is then obtained as the first N points of the $2N$-point Discrete Fourier Transform [29]. As with the Discrete Fourier Transform, the computational efficiency of the DCT is achieved via fast algorithms which could be based on other discrete orthogonal transforms (eg. Fast fourier transform (FFT)), thus rendering it practically feasible [30].

Given a $2D$ reconstructed face image f, the DCT computes the $(i,j)th$ entry of the $2D$ DCT-transformed image using the following equations:

$$D(i,j) = \omega(i)\omega(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos\left[\frac{\pi(2x+1)i}{2N}\right] \cos\left[\frac{\pi(2y+1)j}{2N}\right], \tag{1}$$

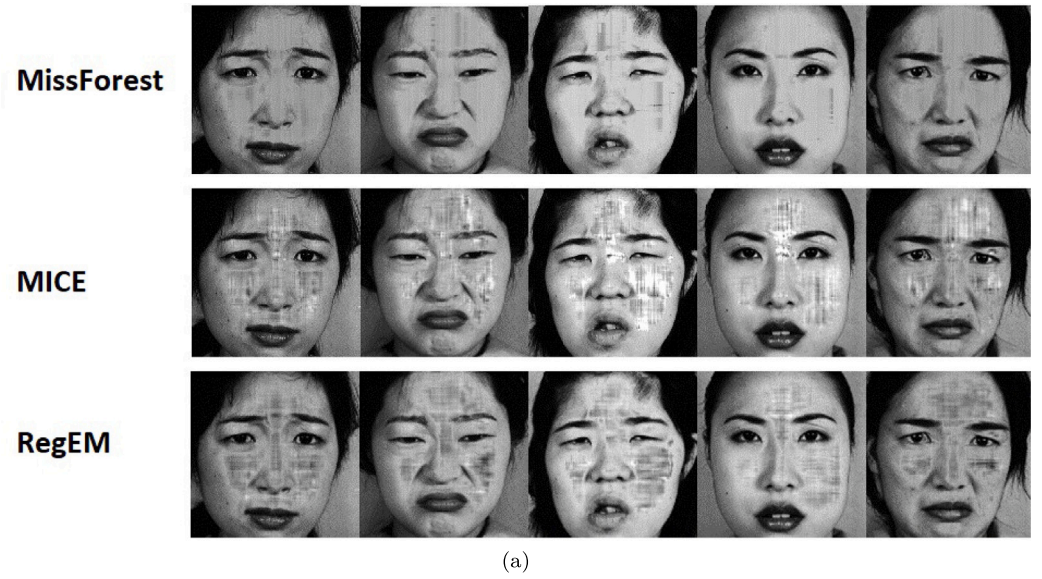


Fig. 5. Sample (20%) de-occluded faces acquired with varying expressions (Test-image database 4).

where $f(x, y)$ is the $(x, y)th$ element (pixel intensity) of the image matrix f, N is the size of the image matrix on which DCT is performed, $D(i, j)$ is the $(i, j)th$ DCT coefficients for $i, j = 0, 1, 2, \ldots, N-1$ and

$$\omega(i) = \omega(j) = \begin{cases} \frac{1}{\sqrt{N}} & i, j = 0 \\ \sqrt{\frac{2}{N}} & i, j \neq 0 \end{cases}. \tag{2}$$

A zero-mean Gaussian filter is then applied to denoise and normalize illuminations, and then the image is reconstructed using Inverse Discrete Cosine Transform (IDCT).

The inverse transform is

$$f(x, y) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \omega(i)\omega(j) D(i, j) \cos\left[\frac{\pi(2x+1)i}{2N}\right] \cos\left[\frac{\pi(2y+1)j}{2N}\right] \tag{3}$$

$x, y = 0, 1, 2, \ldots, N-1.$

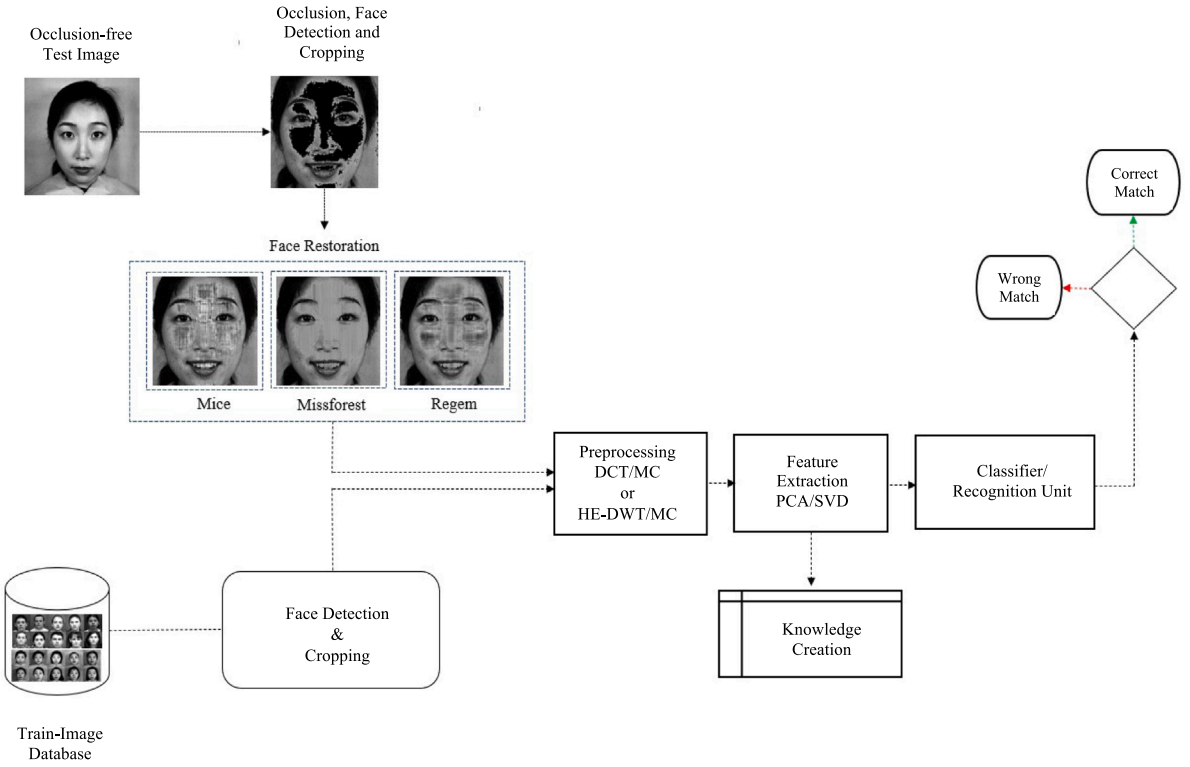


Fig. 6. Research design.

Histogram equalization

Histogram equalization (HE) is a well-known contrast enhancement technique in image processing. Global Histogram equalization modifies the distribution of pixel intensities in an image to approximate uniform distribution for maximum entropy [31]. Though global histogram equalization contrast enhancements are known to cause unnatural effects in images, it could potentially enhance the recognition accuracy of automatic face recognition modules when used with other enhancement mechanisms.

Let $I = I(x, y)$ denote a discrete gray scale face image of size $M \times N$ (number of pixels) with L intensity levels, where $I(x, y) \in \{0, 1, 2, \ldots, L-1\}$ represents the intensity level of the pixel at location (x, y).

- The histogram of the image intensities is $\{n_k\}$, where n_k is the number of pixels whose intensity level is k.
- The associated probability density function (pdf) (normalized histogram) and cumulative density function (cdf) are $p(k) = \frac{n_k}{M} \times N$ and $F(k) = \sum_{k=0}^{L-1} p(k)$ respectively, with $F(L-1) = 1$.
- The equalized intensity level in the output image is obtained by

$$I_{EQ} = I_{EQ}(k) = (L-1) \times F(k), \tag{4}$$

where $k = 0, 1, 2, \ldots, L-1$.

Discrete Wavelet Transform (DWT)

Denoising the face images via a 1-level Discrete Wavelet Transform (DWT) comprised decomposing the image into four (4) sub-bands (LL, LH, HL, HH) of various frequency components, applying a gaussian filter to the respective sub-bands and reconstructing the image via inverse Discrete Wavelet Transform. The LL sub-band gives the low-resolution form of the image and contains global information whilst the remaining sub-bands give the high-resolution forms of the image and contain local information such as eyes, mouth and nose.

According to [32], DWT is the most stable invertible transform in transforming signals in diverse domains. Its efficacy in denoising signals can be attributed to its multiresolution property which allows the analysis of a signal at different resolutions or scales, making it easier to identify patterns and anomalies in large datasets. The Haar wavelet, among other well-known wavelets (Coiflet, Daubechies), was chosen for the purposes of this study due to its simplicity and orthogonal property.

Although the different enhancement mechanisms could give satisfactory results when applied independently, we posit that a unified enhancement framework that combines Histogram equalization and Discrete Wavelet Transform can jointly leverage the advantages of each method to improve recognition accuracy in multiple constrained environments where the MICE, Missforest and RegEM multiple imputation methods are used for image reconstruction.

Feature extraction: Principal Component Analysis (PCA)

The human face image is a high-dimensional structure. Therefore, processes involving their use and analysis suffer from the problems associated with of high-dimensional data. As such, dimensionality reduction is of utmost importance in developing efficient face recognition systems. Given a data matrix $\mathcal{Y}$, dimensionality reduction techniques seek a lower dimensional representation Γ that captures the essential or discriminative features of the matrix without or with minimal loss of information.

According to [33], the need for such techniques is due to the fact that not all image features of an image are important for understanding an underlining phenomenon. Hence, these methods come in handy to offset the high computational cost associated with methods that deals with high dimensional data, even though such methods could yield high level of accuracy in predictions.

The principal component analysis is a statistical dimensional reduction technique that has achieved much success in many facets of high dimensional data analysis, with minimal loss of information in terms of the mean square error [34]. Given a feature matrix associated with a set of face images, PCA finds an orthogonal set of vectors of the feature matrix with the largest variance. For face images, varying facial expressions cause larger intra-personal variation in appearance relative to inter-personal variation in appearance. Such intra-personal variations in facial appearance are known to be handled well by discriminatory-based methods such as PCA [35]. Therefore, PCA is adopted for feature extraction in this study.

For a given set of n face images, the feature matrix associated with each of the images is vectorized and stored as a column of the face matrix $\mathcal{Y} = (\mathcal{Y}_1, \mathcal{Y}_2, \dots, \mathcal{Y}_n)$, where $\mathcal{Y}_1, \mathcal{Y}_2, \dots, \mathcal{Y}_n$ are the vectorized face images. Next, $\mathcal{Y}$ is mean-centered by subtracting the mean image from the individual images (columns) in $\mathcal{Y}$.

The mean is given by

$$\bar{\mathcal{Y}}_j = E(\mathcal{Y}_j), \quad j = 1, 2, \dots, n.$$

The variance–covariance matrix is calculated, following mean centering, as;

$$C = \frac{1}{n} \mathcal{W} \mathcal{W}^T, \tag{5}$$

where the mean centered matrix is $\mathcal{W} = (w_1, w_2, \dots, w_n)$.

The eigenvalues and their corresponding eigenvectors of the variance–covariance matrix are then calculated via singular value decomposition (SVD), $C = \mathcal{U} \boldsymbol{\Sigma} \mathcal{V}^T$.

This decomposes the covariance matrix C into two orthogonal matrices $\mathcal{U}$ and $\mathcal{V}$ and a diagonal matrix $\boldsymbol{\Sigma}$.

The eigenfaces can be obtained by

$$z_j = w_j u_j^T, \tag{6}$$

where u_j is the jth column vector of $\mathcal{U}$.

From the training set, the principal components are extracted (by multiplying them with the eigenvectors which act as weighting coefficients) as;

$$\boldsymbol{\gamma}_j = z_j^T (\mathcal{Y}_j - \bar{\mathcal{Y}}), \tag{7}$$

and $\Gamma^T = [\gamma_1, \gamma_2, \dots, \gamma_n]$.

When a new face (test image) is passed through the recognition module, its unique features are extracted as;

$$\gamma_j^* = z_j^T (\mathcal{Y}_r - \bar{\mathcal{Y}}),$$

and $\boldsymbol{\Gamma}_r^{*T} = [\gamma_1^*, \gamma_2^*, \dots, \gamma_n^*]$.

Then, recognition is done using the city block distance as a classifier. The distances are computed as;

$$\boldsymbol{\psi} = \|\Gamma - \Gamma_r^*\|. \tag{8}$$

The minimum city block distance $d_{ji} = min[\psi], \quad j = 1, 2, \dots, n$ and $i = 1, 2$ are used for classification.

Results and discussion

Numerical evaluation: Recognition performance with DCT and HE-DWT as enhancement mechanisms

The average recognition rate: measures the number of correct recognitions relative to the total number of face images in the test image database passed to the recognition module for recognition. This is a numerical measure used to assess the accuracy of recognition algorithm in recognizing face image captured under the study constraints.

First, we present the results (decisions and associated recognition distances), for a sample of subjects, when the DCT was used as an enhancement mechanism following facial reconstruction for all three reconstruction algorithms.

Fig. 7 shows the results of using the 10% reconstructed expression test faces, following DCT enhancement, for twelve (12) subjects [six (6) each from JAFFE (Fig. 7a) and CKFE (Fig. 7b)] databases for recognition. It can be observed that there were no mismatches when the DCT-enhanced MICE and RegEM test faces were used for recognition. However, there was one (1) mismatch (wrong match) when the DCT-enhanced MissForest reconstructed expression test faces were used for recognition. Fig. 8 shows the results (but with 20% reconstructed expression faces as test images) for the same sample of 12 subjects. The number of wrong

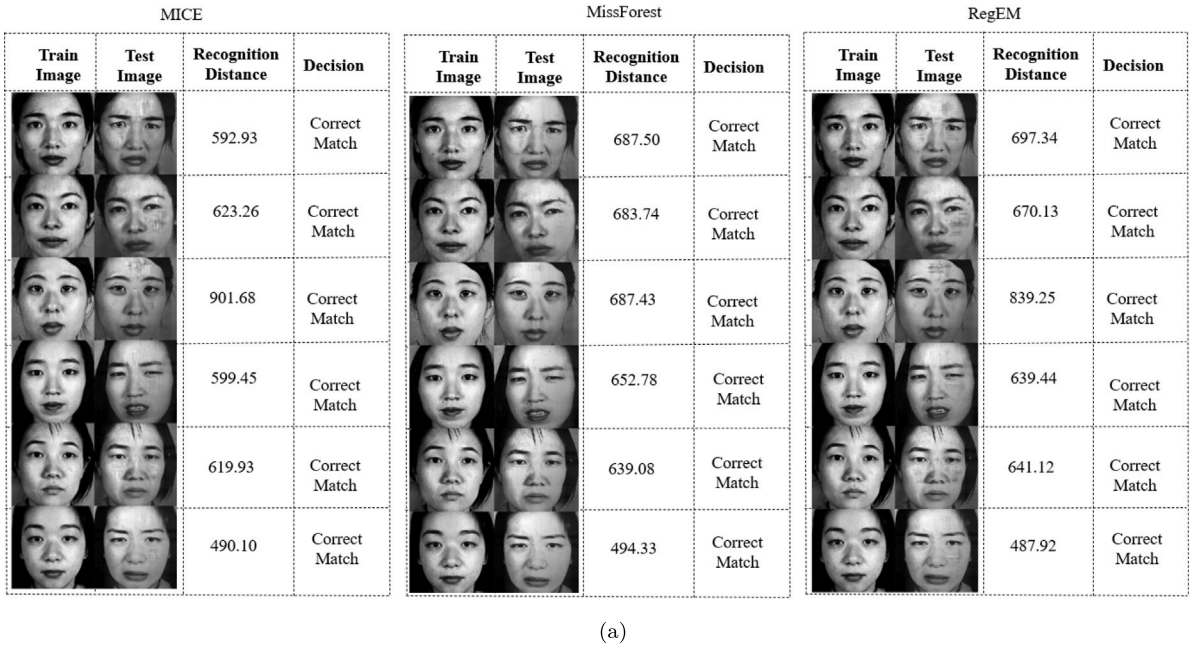


Fig. 7. Sample recognition results per reconstruction method for the JAFFE and CKFE database (10% occlusions).

matches corresponding to the use of the DCT-enhanced MICE, RegEM and MissForest test faces were three (3), three (3) and two (2) respectively. It is, therefore, obvious that there is an increase in the number of mismatches, relative to the results obtained under a 10% reconstruction rate.

Next, we present in Table 1 the recognition rate of the study algorithm (PCA/SVD-L1) obtained with DCT-based enhancement or DWT-based enhancement following Histogram Equalization (HE-DWT) for all three (3) reconstruction/imputation algorithms at 10% and 20% occlusion rates.

It is evident from Table 1 that, the DCT and HE-DWT enhancement mechanisms augment well the MissForest imputation method for reconstruction and recognition of face images acquired under multiple constrained environments. This is because, the use of test faces reconstructed using Missforest imputation mechanism resulted in the highest recognition rates (90.63% & 91.15% using DCT and HE-DWT preprocessing mechanisms respectively at 10% Occlusion rate). Also, at 20% occlusion rate the Missforest gave the

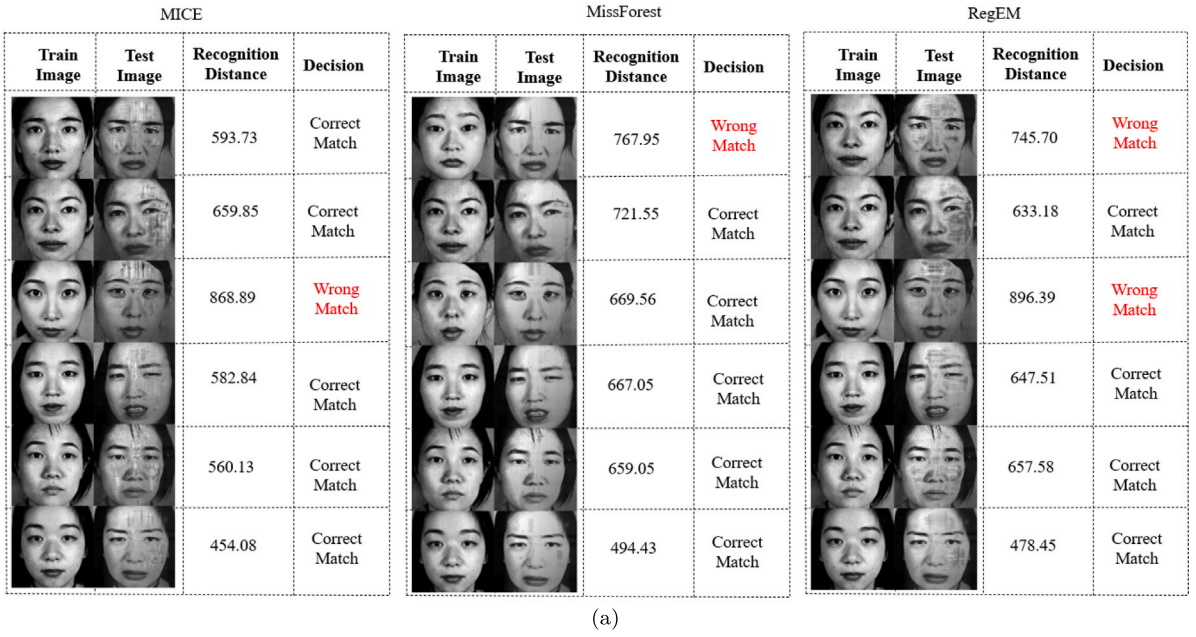


Fig. 8. Sample recognition results per reconstruction method for the JAFFE and CKFE database (20% occlusions).

highest recognition rates (86.98% & 85.94% using DCT and HE-DWT preprocessing mechanisms respectively). Also, the DCT and HE-DWT enhancement mechanisms gave appreciable recognition rates, for MICE and RegEM reconstruction algorithms. The average runtime for the recognition of a subject in the test-image database was less than one second for each algorithm.

Statistical evaluation

A statistical evaluation was performed to establish the equivalence or otherwise of the DCT and HE-DWT enhancement mechanisms. A statistical evaluation of the observed difference in average recognition distances (pairwise comparison) obtained with the DCT and HE-DWT was conducted at $\alpha = 0.05$ significance level. Specifically, the recognition distances obtained by using

Table 1
Recognition rates using the proposed augmented enhancement mechanisms with MICE, MissForest and RegEM for reconstruction.

		Recognition rates		
Occlusion rate	Method	DCT	DWT	HE-DWT
10%	MICE	87.50%	85.94%	85.94%
	MissForest	90.63%	68.75%	91.15%
	RegEM	84.47%	84.44%	84.49%
20%	MICE	80.73%	78.65%	83.85%
	MissForest	86.98%	54.69%	85.94%
	RegEM	82.81%	76.54%	81.25%

Table 2
Normality of difference: Doornik-Hansen test.

Occlusion rate	Test stat	df	*p*-value	MVN
10%	9.855105	12	0.6287	YES
20%	12.92368	12	0.374612	YES

the de-occluded images following DCT enhancement for recognition was compared with the recognition distances obtained for the de-occluded images following HE-DWT enhancement.

Generally, a relatively lower average recognition distance signifies a closer match. The multivariate pairwise comparison test was employed to first ascertain whether statistical significant difference exist between the two algorithms (DCT enhanced PCA-based algorithm and HE-DWT enhanced PCA-based algorithm) across the principal facial expressions (Angry, Disgust, Fear, Happy, Sad, and Surprise). If there exist statistically significant difference in average recognition distance between the two algorithms (enhancement mechanisms), then the algorithm with relatively lower average recognition distance is preferred since it gives closer matches. This multivariate test of difference is a more data-driven evaluation mechanism adopted to established whether statistically significant difference exist between the average recognition distance obtained with the distinct study algorithms. This was done separately for the MICE, MissForest and RegEM de-occlusion methods. For brevity of presentation we would present only the results for MissForest de-occlusion mechanism since it outperformed the MICE and RegM de-occlusion mechanisms in the numerical evaluation shown in Table 1.

For any two treatments r and s, let d_{jer} and $d_{jes}, j = 1, 2, \ldots, n, e = 1, 2, \ldots, p$ denote the recognition distances for the jth subject having p facial expressions in the treatment databases s and r respectively. Then the observed difference is

$$D_{je} = d_{jer} - d_{jes}. \tag{9}$$

We test the assumption that the paired difference $D = (D_{j1}, D_{j2}, \ldots, D_{jp})$ is multivariate normal. That is $D \sim N_p(u, \Sigma_D)$. The hypothesis to be tested under the Hotelling's T^2 test is $H_0 : u = 0$. The Hotelling's T^2 statistic is

$$T^2 = n\bar{D}' \boldsymbol{\Sigma}_D \bar{D}. \tag{10}$$

$\bar{D}$ is the mean of the observed difference D. Under H_0,

$$T_*^2 = \beta(n, p)T^2 \sim F_{p,n-p}(\alpha). \tag{11}$$

where $\beta(n, p) = \dfrac{n-p}{p(n-1)}$, $F_{p,n-p}(\alpha)$ is the upper (100α) percentile of a F-distribution with $(n, n-p)$ degrees of freedom. Therefore, H_0 is rejected if $T_*^2 > F_{p,n-p}(\alpha)$.

The Bonferroni $100(1-\alpha)\%$ simultaneous confidence interval for the individual mean differences is

$$u_i = \bar{d}_i + t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{\tau_{d_i}^2}{n}}, \tag{12}$$

where $\tau_{d_i}^2$ is the ith diagonal element of Σ_D.

Table 2 shows the results of the test of normality of observed difference in recognition distances. The observed differences in recognition distances were found to be multivariate normal with the use of both (10% and 20%) de-occluded set of images.

Table 3 gives the average recognition distances and their respective standard errors (under DCT and HE-DWT) corresponding to the use of the various (10% and 20%) MissForest de-occluded expression test images for recognition. The observed mean difference in recognition distance between DCT and HE-DWT enhanced images are also shown in Table 3. The result gives some indication that the recognition distances are lower under the DCT enhancement mechanism. Table 4 shows the results of testing for a significant difference in the observed average recognition distances using the multivariate paired test. It is seen that there exists statistically significant difference in recognition distance ($p < 0.05$) under DCT and HE-DWT enhancement with the use of each of the 10% and 20% MissForest de-occluded test images for recognition. It can be inferred from Tables 3 and 4 that, at 5% level of significance, the average recognition distances are lower under DCT enhancement relative to HE-DWT enhancement, making the DCT enhancement preferable since it gives a closer match.

Table 3
Mean and Standard Deviation of the recognition distance per expression.

Occlusion rate	Expressions	DCT		HE-DWT		Mean difference
		Mean	S.e	Mean	S.e	(DCT) - (HE-DWT)
10%	Angry	744.1767	49.7777	965.2719	60.9281	−221.0952
	Disgust	762.1061	40.4855	1023.1750	64.4541	−261.0689
	Fear	689.7295	39.9497	937.4438	67.3649	−247.7142
	Happy	638.1117	47.6348	914.8531	74.5677	−276.7414
	Sad	664.5213	54.6439	957.9094	85.5900	−293.3881
	Surprise	941.1787	65.8231	1303.0438	101.7442	−361.8651
20%	Angry	881.0605	60.8198	1201.2438	90.7440	−320.1833
	Disgust	925.9782	55.8913	1291.378	92.1291	−365.4000
	Fear	839.2603	52.8494	1201.8812	97.1722	−362.6210
	Happy	784.1014	58.1735	1182.291	108.5025	−398.1893
	Sad	772.4425	57.0739	1188.9250	114.2386	−416.4825
	Surprise	982.0889	67.8732	1470.375	118.4195	−488.2861

Table 4
Multivariate test of difference (Hotelling's T^2).

Occlusion rate	Test statistic	F	df	*p*-value
10%	75.8181	10.59823	(6, 26)	5.809×10^{-6}
20%	59.35826	8.297392	(6, 26)	4.4528×10^{-5}

Table 5
Confidence Interval.

Number	Expressions	10%		20%	
		Lower limit CI	Upper limit CI	Lower limit CI	Upper limit CI
1	Angry	100.5440	341.6464	152.6187	487.7479
2	Disgust	122.8378	399.3000	184.5800	546.2199
3	Fear	121.3165	374.1119	181.7394	543.5026
4	Happy	163.6841	389.7987	215.7959	580.5826
5	Sad	153.5196	433.2566	197.7809	635.1842
6	Surprise	212.5932	511.1370	299.4959	677.0763

The results of constructing a 95% Bonferroni Simultaneous Confidence Interval to assess the direction of the achieved significance is presented in Table 5. It can be seen from Table 5 that there exist statistically significant difference in average recognition distances between DCT and HE-DWT enhanced MissForest de-occluded test images across all the principal face expressions (Angry, Disgust, Fear, Happy, Sad and Surprise).

Conclusion and recommendation

This study sought to propose enhancement mechanisms that augment the use of three (3) multiple imputation methods (MICE, MissForest and RegEM) for image reconstruction in PCA-based face recognition characterized by multiple constraints. It was demonstrated that the proposed enhancement mechanisms (DCT and HE-DWT) improved the recognition rate in recognizing frontal face images acquired under moderately low occlusion rates and varying facial expressions compared with using DWT as an enhancement mechanism for all three reconstruction methods employed in this study. The highest recognition rates (90.63% and 86.98%) were attained when MissForest was used for reconstruction at both 10% and 20% reconstruction rates respectively. Subsequent statistical evaluation revealed that average recognition distance with DCT augmentation, following MissForest reconstruction, was found to be relatively lower at a 5% level of significance, making the DCT enhancement mechanism preferable when compared to the HE-DWT enhancement mechanism.

The achieved average recognition rates are improvements over those of [17], who attained average recognition rates of 77.31% and 76.85% under similar multiple constraints. The above findings are also consistent with those of [36] who conducted a comprehensive survey on recognition techniques under occlusion.

The study also revealed that there exists a statistically significant difference in average recognition distances between DCT and HE-DWT enhanced MissForest reconstructed test images across all the principal face expressions (Angry, Disgust, Fear, Happy, Sad and Surprise). The DCT enhanced algorithm gave relatively lower average recognition distances (closer matches) than the HE-DWT enhanced algorithm.

Overall, the study recommends augmenting DCT enhancement with MissForest reconstruction for PCA-based face recognition modules when moderately low levels of occlusion (10% and 20%) and varying facial expressions are the underlying constraints. Future work will focus on assessing the relative merits of using these multiple imputation mechanisms in highly constrained environments as well as exploring the suitability of the proposed enhancement mechanisms on other face recognition algorithms when used for recognition under multiple constraints.

CRediT authorship contribution statement

Joseph Agyapong Mensah: Conceptualization, Methodology, Writing – original draft. **Eric Ocran:** Data curation, Formal analysis. **Louis Asiedu:** Supervision, Validation, Reviewing and editing.

Declaration of competing interest

The authors declare that there is no conflict of interest.

References

[1] H. Oufdou, L. Bellanger, A. Bergam, A. El Ghaziri, K. Khomsi, E.M. Qannari, et al., Comparison of different regularized and shrinkage regression methods to predict daily tropospheric ozone concentration in the grand Casablanca area, Adv. Pure Math. 8 (10) (2018) 793.
[2] L. Asiedu, F.O. Mettle, J.A. Mensah, Recognition of reconstructed frontal face images using FFT-PCA/SVD algorithm, J. Appl. Math. 2020 (2020).
[3] O.N. Oyelade, A.E. Ezugwu, Characterization of abnormalities in breast cancer images using nature-inspired metaheuristic optimized convolutional neural networks model, Concurr. Comput.: Pract. Exper. 34 (4) (2022) e6629.
[4] R. Josphineleela, P. Raja Rao, A. Shaikh, K. Sudhakar, A multi-stage faster RCNN-based iSPLInception for skin disease classification using novel optimization, J. Digit. Imaging (2023) 1–17.
[5] A.M. Bronstein, M.M. Bronstein, R. Kimmel, Robust expression-invariant face recognition from partially missing data, in: European Conference on Computer Vision, Springer, 2006, pp. 396–408.
[6] H. Drira, B.B. Amor, A. Srivastava, M. Daoudi, R. Slama, 3D face recognition under expressions, occlusions, and pose variations, IEEE Trans. Pattern Anal. Mach. Intell. 35 (9) (2013) 2270–2283.
[7] P.-L. Loh, M.J. Wainwright, High-dimensional regression with noisy and missing data: Provable guarantees with non-convexity, Adv. Neural Inf. Process. Syst. 24 (2011).
[8] T.Y. Kim, K.M. Lee, S.U. Lee, C.-H. Yim, Occlusion invariant face recognition using two-dimensional PCA, in: Advances in Computer Graphics and Computer Vision, Springer, 2007, pp. 305–315.
[9] W. Zhang, S. Shan, X. Chen, W. Gao, Local gabor binary patterns based on Kullback–Leibler divergence for partially occluded face recognition, IEEE Signal Process. Lett. 14 (11) (2007) 875–878.
[10] Z. Zhang, X. Ji, X. Cui, J. Ma, A survey on occluded face recognition, in: 2020 the 9th International Conference on Networks, Communication and Computing, 2020, pp. 40–49.
[11] R. Weng, J. Lu, Y.-P. Tan, Robust point set matching for partial face recognition, IEEE Trans. Image Process. 25 (3) (2016) 1163–1176.
[12] M. Zou, M. You, T. Akashi, Reconstruction of partially occluded facial image for classification, IEEJ Trans. Electr. Electron. Eng. 16 (4) (2021) 600–608.
[13] L. Asiedu, J.A. Mensah, F. Ayiah-Mensah, F.O. Mettle, Assessing the effect of data augmentation on occluded frontal faces using DWT-PCA/SVD recognition algorithm, Adv. Multimed. 2021 (2021).
[14] A. Kumar, M. Kumar, A. Kaur, Face detection in still images under occlusion and non-uniform illumination, Multimedia Tools Appl. 80 (2021) 14565–14590.
[15] X. Yin, D. Huang, Z. Fu, Y. Wang, L. Chen, Segmentation-reconstruction-guided facial image de-occlusion, in: 2023 IEEE 17th International Conference on Automatic Face and Gesture Recognition, FG, IEEE, 2023, pp. 1–8.
[16] J.K. Appati, K.S. Adu-Manu, E. Owusu, Implementation of missing data imputation schemes in face recognition algorithm under partial occlusion, Adv. Multimed. 2022 (2022).
[17] J.A. Mensah, L. Asiedu, F.O. Mettle, S. Iddi, Assessing the performance of DWT-PCA/SVD face recognition algorithm under multiple constraints, J. Appl. Math. 2021 (2021) 1–12.
[18] M.E. Rana, A.A. Zadeh, A.M.M. Alqurneh, Use of image enhancement techniques for improving real time face recognition efficiency on wearable gadgets, J. Eng. Sci. Technol. 12 (1) (2017) 155–167.
[19] S. Van Buuren, K. Oudshoorn, Flexible Multivariate Imputation by MICE, TNO, Leiden, 1999.
[20] D.J. Stekhoven, P. Bühlmann, MissForest—Non-parametric missing value imputation for mixed-type data, Bioinformatics 28 (1) (2012) 112–118.
[21] A.D. Shah, J.W. Bartlett, J. Carpenter, O. Nicholas, H. Hemingway, Comparison of random forest and parametric imputation models for imputing missing data using MICE: A CALIBER study, Am. J. Epidemiol. 179 (6) (2014) 764–774.
[22] G.E. Hinton, S. Sabour, N. Frosst, Matrix capsules with EM routing, in: International Conference on Learning Representations, 2018.
[23] H. Li, K. Zhang, T. Jiang, The regularized EM algorithm, in: AAAI, 2005, pp. 807–812.
[24] G.H. Golub, M. Heath, G. Wahba, Generalized cross-validation as a method for choosing a good ridge parameter, Technometrics 21 (2) (1979) 215–223.
[25] V. Britanak, P.C. Yip, K.R. Rao, Discrete Cosine and Sine Transforms: General Properties, Fast Algorithms and Integer Approximations, Elsevier, 2010.
[26] S.P. Singh, G. Bhatnagar, Perceptual hashing-based novel security framework for medical images, in: Intelligent Data Security Solutions for E-Health Applications, Elsevier, 2020, pp. 1–20.
[27] G.A. Licciardi, Hyperspectral compression, in: Data Handling in Science and Technology, Vol. 32, Elsevier, 2020, pp. 55–67.
[28] Z.-M. Lu, S.-Z. Guo, Lossless Information Hiding in Images, Syngress, 2016.
[29] K. Sayood, Introduction to Data Compression, Morgan Kaufmann, 2017.
[30] M. Nixon, A. Aguado, Feature Extraction and Image Processing for Computer Vision, Academic Press, 2019.
[31] S.E. Umbaugh, Computer Vision and Image Processing: A Practical Approach using Cviptools with Cdrom, Prentice Hall PTR, 1997.
[32] D. Devi, S. Sophia, S.B. Prabhu, Deep learning-based cognitive state prediction analysis using brain wave signal, in: Cognitive Computing for Human-Robot Interaction, Elsevier, 2021, pp. 69–84.
[33] I.K. Fodor, A Survey of Dimension Reduction Techniques, Tech. Rep., Lawrence Livermore National Lab., CA (US), 2002.
[34] I.T. Jolliffe, Principal component analysis, Technometrics 45 (3) (2003) 276.
[35] H.V. Nguyen, L. Bai, L. Shen, Local gabor binary pattern whitened PCA: A novel approach for face recognition from single image per person, in: International Conference on Biometrics, Springer, 2009, pp. 269–278.
[36] D. Zeng, R. Veldhuis, L. Spreeuwers, A survey of face recognition techniques under occlusion, IET Biometrics 10 (6) (2021) 581–606.