SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
http://saucis.sakarya.edu.tr/
e-ISSN: 2636-8129
Publisher: Sakarya University
Vol. 7, No. 3, 2024, 361-377
DOI: 10.35377/saucis...1495856

RESEARCH ARTICLE

A Hybrid Approach for Color Face Recognition Based on Image Quality Using Multiple Color Spaces

Mohammad Mehdi Pazouki[1], Önsen Toygar[1*], Mahdi Hosseinzadeh[2]

[1] Eastern Mediterranean University, Faculty of Engineering, Department of Computer Engineering, Gazimağusa, Turkish Republic of Northern Cyprus
[2] Tarbiat Modares University, Faculty of Electrical and Computer Engineering, Department of Computer Engineering, Tehran, Iran

Corresponding author:
Önsen Toygar, Eastern Mediterranean University, Faculty of Engineering, Department of Computer Engineering, Gazimağusa, Turkish Republic of Northern Cyprus
onsen.toygar@emu.edu.tr

Article History:
Received: 04.06.2024
Accepted: 18.09.2024
Published Online: 29.11.2024

ABSTRACT
In this paper, the color face recognition problem is investigated using image quality assessment techniques and multiple color spaces. Image quality is measured using No-Reference Image Quality Assessment (NRIQA) techniques. Color face images are categorized into low, medium, and high-quality face images through the High Low Frequency Index (HLFI) measure. Based on the categorized face images, three feature extraction and classification methods as Scale Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), and Convolutional Neural Networks (CNN) are applied to face images using RGB, YCbCr, and HSV color spaces to extract the features and then classify the images for face recognition. To enhance color face recognition systems' robustness, a hybrid approach that integrates the aforementioned methods is proposed. Additionally, the proposed system is designed to serve as a secure anti-spoofing mechanism, tested against different attack scenarios, including print attacks, mobile attacks, and high-definition attacks. A comparative analysis that assesses the proposed approach with the state-of-the-art systems using Faces94, ColorFERET, and Replay Attack datasets is presented. The proposed method achieves 96.26%, 100%, and 100% accuracies on ColorFERET, Replay Attack, and Faces94 datasets, respectively. The results of this analysis show that the proposed method outperforms existing methods. The proposed method showcases the potential for more reliable and secure recognition systems.

Keywords: Face recognition, Image quality assessment measures, Color spaces, Feature extraction, Deep learning

1. Introduction

Exploring the recent technology for face recognition has shown that it is an important application in the world, especially for passport control, entrance to secure systems, and authorization for mobile phones and other technical devices. A digital image or a single frame extracted from a video recording can be used for the identification or verification of faces [1]. The technology analyzes specific facial features in an image and compares them with those in a database of faces stored in a central repository. In particular, applications based on color face images perform significantly better than those using grayscale facial images, largely because grayscale conversion separates luminance from chrominance, and luminance is more efficient for detecting visual features [2]. Successful recognition of face images requires the inclusion of multiple attributes of the face, namely orientation, position, expression, color, and scale [3]. However, color alone does not provide a substantial advantage in face recognition beyond luminance information [4]. Many face recognition methods rely solely on luminance data, converting color face images into grayscale images to streamline processing, as color adds complexity and slows down computation [2].

In the literature, Olayede et al., in 2020, presented the recent challenges and approaches for face recognition systems [3]. The authors indicated that under the illumination challenge, the best result was 90.38% on the ORL database using the Convolutional Neural Networks (CNN) approach. In the presence of pose variations, the highest performance was obtained as 93.5% on the PIE dataset using a 3D reconstruction procedure based on facial landmarks and sparse regression. Face recognition under various expressions was also discussed and the best recognition rate was achieved as 89.76% on the JAFFE dataset using the CNN approach. On the other hand, in the presence of occlusions, the highest recognition rate was reported as 92.5% on the AR dataset using a fuzzy max-pooling approach based on CNN [3].

Recently, in 2023, Rusia and Singh [4] presented a comprehensive survey on face recognition approaches and challenges. The authors reviewed the fusion of feature-based and texture-based methods that represent 91% accuracy under plastic

Cite as: M. M. Pazouki, Ö. Toygar, M. Hosseinzadeh, "A Hybrid Approach for Color Face Recognition Based on Image Quality Using Multiple Color Spaces," *Sakarya University Journal of Computer and Information Sciences*, vol.7, no. 3, pp. 361-377, 2024. Doi: 10.35377/saucis…1495856

surgery challenges. Besides, for detecting facial makeup, a 93.5% detection rate was mentioned which was achieved using shape, texture, and skin color analysis-based methods with RGB, Lab, and HSV color spaces. Additionally, face-based indirect spoofing was discussed with photo attack, video attack, and mask attack categories. The best error rate was reported as 0.024 Half Total Error Rate (HTER) on the Replay Attack dataset using an image quality assessment based fast non-intrusive method with Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM) methodologies [4].

Moreover, several color spaces have been studied for performing color face recognition. For instance, Yang et al. [5] introduced a global Eigen scheme that handles color components independently, showing potential improvements in face recognition with color information. Yip and Sinha [6] developed a non-negative matrix factorization (NMF) approach that outperforms grayscale methods in recognizing color face images. While some important investigations of the researchers emphasize the role of color cues in recognition [7], the rest of the researchers indicate that color does not significantly alter face recognition processes since it does not affect shape-from-shading mechanisms [8]. Nevertheless, it can be stated that color is important in face recognition technology, especially in high-resolution images.

The most widely used color spaces, namely YCbCr, RGB, and HSV are utilized in this study to assess image quality under various conditions. The aforementioned color spaces are analyzed for each channel individually, afterwards, the concatenation of outputs is employed. Several feature extraction and/or classification techniques such as Principal Component Analysis (PCA), Local Binary Patterns (LBP), Color Local Binary Patterns (ColorLBP), Scale Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), Oriented FAST and rotated BRIEF (ORB), and Convolutional Neural Networks (CNN), are employed to extract or classify the facial features in face recognition experiments. Additionally, Image Quality Measures (IQM) and No-Reference Image Quality Measure (NR-IQM) metrics are used to measure the image quality and distinguish between real and fake face images [9], [10]. All the aforementioned techniques for feature extraction and classification are integrated in this study and a hybrid color face recognition algorithm for color face images with different image qualities is proposed. Consequently, the experimental results using multiple feature extraction methods and color spaces on publicly available datasets are presented.

The novelty of the paper can be summarized with the following contributions:

- This study combines the usage of image quality by utilizing No-Reference Image Quality Assessment techniques and employing various hand-crafted feature extraction methods, including ColorLBP, LBP, SIFT, SURF, ORB, and PCA to extract features from both color and grayscale images.
- Hand-crafted feature extraction methods and deep learning-based Convolutional Neural Networks are used together for color face recognition in a hybrid approach.
- Three different color spaces (YCbCr, HSV, and RGB) across three face databases (ColorFERET, Faces94, and Replay Attack) are applied.
- High-Low Frequency Index (HLFI) quality measure is employed in determining image quality, surpassing other no-reference quality measures.
- For medium-quality images, the texture-based SIFT feature extraction method combined with RGB color space is found to be the most suitable. Deep-learning-based CNN demonstrates high performance on low-quality images, particularly in the HSV color space. The YCbCr color space, in conjunction with the SURF feature extraction method, proves to be the triumphant approach for high-quality images.
- The proposed approach is a hybrid algorithm in which the outcomes are promising, outperforming state-of-the-art face recognition approaches on the same datasets.

The remaining sections of the paper cover fundamental concepts, literature review, the proposed method, experimental results, and concluding remarks, all contributing to the understanding and advancement of face recognition technology.

2. Materials and Methods

The databases used in the experiments; materials and methodologies employed in this study such as image quality assessment techniques, color spaces, feature extractors, and classification methods; and the details of the proposed method are explained in the following subsections.

2.1. Databases

Three color face databases are used to measure the quality level of face images and evaluate the performance of color face recognition methods. The databases used are Replay-Attack (Kumar et al. [11]), Faces94 (Wang et al. [12]), and ColorFeret (Rowley et al. [13]).

The Replay-Attack database, sourced from Idiap, is a valuable resource for studying face spoofing. It comprises 1300 video clips capturing attempts at image and movie attacks on 50 subjects, conducted under varying illumination conditions. These color videos, captured by a webcam at 320x240 pixels resolution, provide rich data for analysis. In our study, we carefully selected 100 real images, with 50 designated for testing and the remaining 50 for training purposes. These images encompass three distinct attack types, namely High-Def, Mobile, and Print. The High-Def attack involves displaying high-quality images on a tablet, while the Mobile attack showcases images on a mobile screen and presents them to the photo lens. The Print attack involves printing the image and presenting it to the photo lens. For each attack type, we curated 100 images from the Replay-Attack database, allocating 50 for training and 50 for testing.

Conversely, the Faces94 database boasts a collection of 153 images, each rendered at a resolution of 180x200 pixels. This database thoughtfully segregates its images into distinct categories, delineating between male and female individuals. Specifically, it comprises 20 images of females, 113 of males, and 20 of male staff members. The majority of subjects represented in this database are first-year undergraduate students, predominantly falling within the age range of 18 to 20 years. The imagery within Faces94 exhibits a deliberate artificial illumination setting, with some subjects captured wearing glasses. Furthermore, the photos exhibit a blend of both tungsten and fluorescent lighting conditions, adding diversity to the dataset for comprehensive analysis.

The expansive ColorFERET database boasts a staggering collection of 11,338 color images, each generously sized at 512 x 768 pixels. These images were meticulously captured against a semi-controlled backdrop, featuring subjects adopting 13 distinct poses. Impressively, the database encompasses imagery from a diverse pool of 994 subjects. The training stage of the machine relies on the standard frontal image set (denoted as Fa), where one frontal image per subject is employed. Subsequently, the system's prowess is assessed on a different frontal pose drawn from the complementary set (referred to as Fb). In the context of the experiments conducted in this study, a subset comprising 268 training images and 268 test images is judiciously utilized to ensure robust evaluation.

Table 1 provides a concise overview of the image distribution utilized for both training and testing across the three pivotal databases featured in our experimental analysis. Furthermore, a visual glimpse into these datasets is given in Figure 1. A compelling showcase of color face images sourced from the Replay Attack, Faces94, and ColorFERET datasets are presented in Figure 1 (a), (b), and (c), respectively. In particular, Figure 1a thoughtfully presents an assortment of sample images from the Replay Attack database, illustrating high-definition attacks, mobile attacks, and print attacks.

A series of comprehensive experiments have been conducted in the realm of face recognition utilizing color-rich images. These experiments leverage the power of No-Reference Image Quality Measurement (NR-IQM) methods to ascertain the quality levels of the images under scrutiny. Four distinct IQM methods were meticulously implemented to accomplish this. Through rigorous evaluation, the most suitable IQM method was judiciously chosen for integration into the image quality determination process.

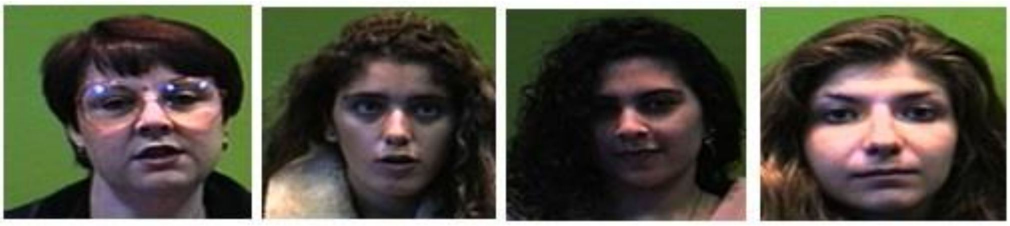
Furthermore, these experiments ventured into the exploration of three distinctive color spaces: RGB, HSV, and YCbCr, each offering a unique perspective on image representation. Additionally, a robust array of seven feature extraction methods, including PCA, LBP, ColorLBP, SIFT, SURF, ORB, and CNN, were meticulously implemented to capture the diverse facets of image characteristics.

The experiments are conducted on a Windows 10 Professional OS, Skylake 6700k Core i7 CPU operating at 4.3 GHz and 16 GB of 2400 Hz dual channel RAM. The version of Python is 3.6 and the OpenCV is 3.4. The following subsections provide an in-depth presentation of these experiments.

Table 1. Summary of Datasets and Experimental Setup Used in the Experiments

<table>
<tr><th>Dataset</th><th colspan="3">Test images</th><th colspan="3">Train images</th></tr>
<tr><td>Faces94</td><td colspan="3">1520 real</td><td colspan="3">1520 real</td></tr>
<tr><td>ColorFERET</td><td colspan="3">268 real</td><td colspan="3">268 real</td></tr>
<tr><td rowspan="3">Replay Attack</td><td colspan="3">150 fake</td><td colspan="3">150 fake</td></tr>
<tr><td colspan="3">50 real</td><td colspan="3">50 real</td></tr>
<tr><td>50 Print Attack</td><td>50 Mobile Attack</td><td>50 High-def Attack</td><td>50 Print Attack</td><td>50 Mobile Attack</td><td>50 High-def Attack</td></tr>
</table>

(a)



(b)

(c)

Figure 1. Sample Color Face Images from (a) Replay Attack (b) Faces94 and (c) ColorFERET

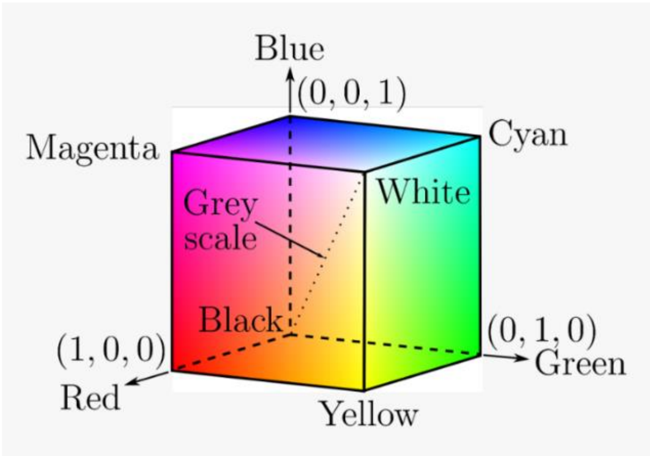
2.2. Materials

Image processing investigates several color channels in which the most widely used is RGB color space. Investigation of color face images in the RGB domain showed some limitations due to the close relationship between the channels of red (R), green (G), and blue (B). In this research, our primary goal is to use a color space with reduced inter-element correlation to enhance the performance of face classification. To achieve this objective, a combination of diverse color spaces is employed in the proposed color face recognition method, including RGB color space, HSV, and YCbCr color spaces. Color space models are shown in Figure 2 [14]. HSV is renowned for its alignment with human perception and encompasses hue, saturation, and value components. In contrast, the YCbCr color domain which was initially designed for image compression, separates the luminance (Y) channel from chrominance channels (Cb, Cr). On the other hand, as another advantage, YCbCr also serves as a functional space for the segmentation of skin color applications [4].

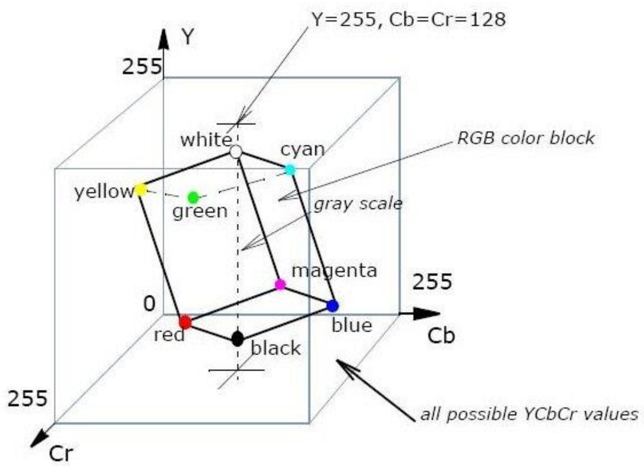
Image quality assessment is a significant field that is employed for many image processing applications. Within this field, Full Reference Image Quality Measurement (FR-IQM) and Reduced Reference Image Quality Measurement (RR-IQM) methods can be used on complete or partial images to evaluate visual quality effectively [9]. On the other hand, it is also possible not to use a reference image and measure the quality. Specifically, No-Reference Image Quality Measures provide a quality measurement of the images without using reference images [10]. Image quality can be predicted accurately by the majority of NR-IQM metrics for a single type of distortion. There are four notable NR-IQMs employed in this study: Blind Image Quality Index (BIQI), JPEG Quality Index (JQI), Naturalness Image Quality Estimator (NIQE) and High-Low Frequency Index (HLFI).

The first quality metric employed in this study is JQI which is successful in estimating image quality when block-based algorithms, which are common in compression methods like Joint Photographic Experts Group (JPEG) [2], introduce distortion at low bit rates. Meanwhile, the second quality metric is HLFI which demonstrates sensitivity to image sharpness by calculating the difference in power between lower frequency values and upper frequency values in the Fourier Spectrum. Blind Image Quality Index is the third metric employed which capitalizes on natural scene statistics (NSS) and prior knowledge of unaltered natural scene images to train its primary model. Lastly, the fourth metric is NIQE, which is a comprehensive blind image quality analyzer that constructs an informed quality profile through a Gaussian multivariate model of natural scenes. The assessment of all image quality measures is performed collectively focusing on improving the accuracy and reliability. Methods that fail for a specific application to deliver appropriate and detailed evaluations are excluded from consideration.

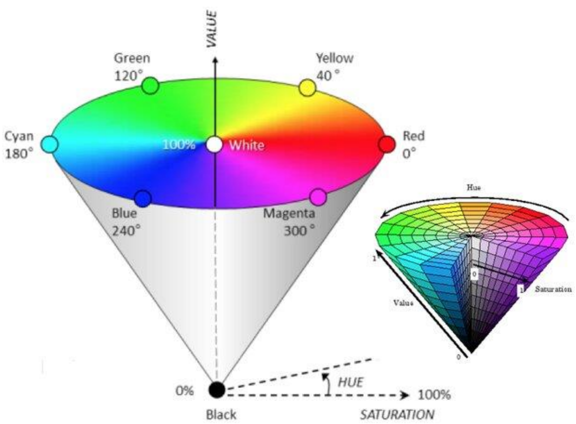
Conversely, the most prominent step for face recognition is the feature extraction step. Facial feature extraction specifically involves the isolation of essential facial components including the nose, mouth, and eyes. This process stands as a fundamental prerequisite for initiating various sophisticated techniques, such as facial tracking, face recognition, and facial expression recognition. Feature extraction can be classified into several categories. One of the most widely used categories is appearance-based feature extraction methods. As an example of this category, the Principal Component Analysis (PCA) method performs data summarization by condensing a comprehensive set of variables into a reduced set through specific transformations. These transformations aim to convert linearly correlated variables into uncorrelated ones, effectively simplifying the data.



(a)



(b)



(c)

Figure 2. Color Models (a) RGB Color Space, (b) YCbCr Color Space, (c) HSV Color Space [14]

On the contrary, texture-based feature extraction methods involve several methods such as Local Binary Patterns (LBP) and Scale Invariant Feature Transform (SIFT) that are based on the texture of images and known as local descriptors. The fundamental concept behind LBP revolves around the replacement of pixel values within an image with decimal characters to produce a code that extracts the local structural information surrounding each pixel [15]. In this process, a central pixel undergoes analysis in conjunction with its eight neighboring pixels. Pixels with values lower than the center pixel are assigned bit 0, while the others that are equal to or greater than the center pixel receive bit 1. Consequently, for every center pixel, a binary number is created by concatenating these binary digits in a clockwise fashion, starting from one of its top-left neighbors.

Color Local Binary Patterns (ColorLBP) method is another texture-based feature extraction method that is a variant of the LBP method. ColorLBP method is tailored for the analysis of color images, which comprise three distinct channels in RGB color space, representing the blue, green, and red color components. In the realm of color space analysis, various methods offer distinct advantages and applications. RGB images, characterized by their red, green, and blue channels, are known for their clarity and adaptability under varying conditions. This is achieved by normalizing these color elements within the color space of RGB. Contrarily, different color spaces can also be considered for extracting features from images within the ColorLBP method. For example, the color space of HSV is designed to emulate the visual system's perception of color with three channels as hue, which defines color; saturation, representing chrominance; and value, specifying luminance. Similarly, the other color space, namely YCbCr, divides the image into three channels. The first channel is the luminance (Y) while the remaining channels are chrominance blue (Cb) and chrominance red (Cr). In general, the processing steps for these color spaces follow a similar pattern. Condensation is applied to each channel's histogram independently and then concatenation of histograms is conducted to extract the whole feature vector.

Another powerful texture-based feature extraction method is the Scale-Invariant Feature Transform (SIFT) [16]. SIFT is renowned for its important ability to identify and characterize features in different scales, making it invaluable for tasks like image recognition, classification, image registration, and 3D reconstruction. Similarly, Speeded-Up Robust Features (SURF) [17] is a significant variant of SIFT which emerges as an efficient alternative to SIFT. It represents a faster iteration of SIFT with high speed and robustness against image transformations. The SURF method is a powerful and fast feature detector and descriptor that is adaptable for several tasks such as image classification, object recognition, and 3D reconstruction.

Oriented FAST and rotated BRIEF (ORB) method was proposed in 2011 [18] as another alternative for local descriptors. ORB offers a rapid and robust solution for tasks like object recognition and 3D reconstruction. The method builds upon the FAST keypoint detector and employs a modified version of the visual descriptor known as Binary Robust Independent Elementary Features (BRIEF). ORB method is aimed at providing a reliable and efficient alternative to the SIFT algorithm.

Recently, the development of deep learning techniques attracted many researchers in many fields such as face recognition. In computer vision and biometrics fields, deep learning techniques such as Convolutional Neural Networks (CNN) are universalized [3, 4]. Object detection, tracking, surveillance, and object recognition tasks have all harnessed the power of CNN. In the realm of surveillance applications, CNN finds widespread use, particularly in face recognition. In this respect, numerous university automation systems and intelligent entry management solutions have successfully incorporated face recognition technology [19] into their operations.

Deep learning approaches such as Convolutional Neural Networks have shown superior results for detecting and recognizing faces across diverse conditions [3, 4]. These conditions may include scenarios where intra-class and inter-class variations exist in the facial images. These variations can be listed as facial expressions, partial occlusions, illumination variations, the similarity between identical twins' faces [20], and so on.

2.3. Proposed Method

This study proposes a novel and hybrid approach for the recognition of facial images across various quality levels, encompassing low, medium, and high-quality color images. To achieve this, an extensive series of preliminary experiments were conducted, exploring different color spaces and diverse techniques for extracting features in face recognition. The outcome of these experiments led to the development of a versatile color-based face recognition system employing a hybrid methodology. To determine the most suitable quality assessment measure for this study, various image quality assessment metrics were applied, ultimately selecting HLFI for this purpose. Furthermore, numerous methods to extract facial features including LBP, SIFT, SURF, ORB, and CNN, were independently implemented for both grayscale and color images within the RGB, HSV, and YCbCr color spaces.

This study adopts a tailored approach based on image quality: low-quality images are processed using the CNN approach within the HSV color space, medium-quality images employ the SIFT feature extractor using RGB color space, and high-quality images utilize the SURF method in the YCbCr color space. Each of these scenarios was thoroughly explored in preliminary experiments to determine the most effective feature extractor and color space for the respective image quality category. In terms of matching and classification, Manhattan Distance measurement is employed, and the face classification is carried out using the Nearest Neighbor classifier or a CNN model. This matching and classification process is conducted using a comparison with the test image features with the corresponding using the same color space to train image feature sets, ultimately yielding the system's decision (ID). A detailed and step-by-step algorithm of the proposed method is presented in Algorithm 1. The flowchart of the proposed method is shown in Figure 3. Every phase of the suggested approach shown in the flowchart is explained in detail in the following subsections.

Algorithm 1. Proposed Method for Color Face Recognition

```
1   Start
2   Read the training images and calculate HLFI value for
    each of them and apply preprocessing.
3   Check the average HLFI value and using the threshold
    values, determine whether an image is low, high, or
    medium-quality image. Store the quality type of
    training images.
4   Use CNN approach to extract features from low-quality
    images in HSV color space. Go to step 7.
5   If the image is medium quality image, then apply
    feature extraction on RGB color space using SIFT
    technique. Go to step 7.
6   If the image is high quality image, then apply feature
    extraction on YCbCr color space using SURF technique.
    Go to step 7.
7   Perform matching and classification after applying
    preprocessing, HLFI calculation and image quality type
    decision on test images according to the same threshold
    values as in training.
8   Report the decision of the classifier.
9   Stop
```

2.3.1. Reading Train Images, Applying HLFI Quality Metrics and Preprocessing

First of all, the initial stage of the proposed method helps to read the database images and convert them into the selected image format. Afterward, the images are saved in specific directories so that it is easier to read them. Since the size of each image is different, resizing the images to the same size is performed. Image quality is then studied by calculating the HLFI image quality measure [21] on each training image as follows:

$$HLFI(I) = \frac{\sum_{i=1}^{i_l}\sum_{j=1}^{j_l}|F_{i,j}| - \sum_{i=i_h+1}^{N}\sum_{j=j_h+1}^{M}|F_{i,j}|}{\sum_{i=1}^{N}\sum_{j=1}^{M}|F_{i,j}|} \tag{1}$$

The High-Low Frequency Index (HLFI) feature is sensitive to image sharpness and this is performed by computing the difference between the power in the lower and upper frequencies of the Fourier Spectrum. In HLFI calculation as shown in (1), jl, jh, il, ih are the indices corresponding to the lower and upper-frequency thresholds, M and N demonstrate the number of rows and columns of the images, Fij is the Fourier transform of image matrix I. In this study, the values of the parameters used are jl = jh = 0.15M and il = ih = 0.15N.

2.3.2. Classifying Image Quality by Specifying Threshold

Using the HLFI image quality measure, the individual images are examined and the results are saved. Using the average HLFI method, a threshold is found to recognize the quality of each image and then the quality type is stored as high-quality, medium-quality, and low-quality. Threshold values for each quality type are specified as described in the next section using the average HLFI values based on preliminary experiments on RGB, HSV, and YCbCr color spaces.

2.3.3. Feature Extraction on Different Color Spaces

SIFT and SURF methods are used as feature extractors in this study on different channels of color spaces. Additionally, deep learning-based CNN is used as a feature extractor and classifier in the proposed method.

Scale Invariant Feature Transform (SIFT) approach [16] detects the points of interest which are called keypoints and produce features using four steps as follows: (1) Scale-space Extrema Detection which identifies keypoints at different scales; (2) Keypoint Localization which refines the location and scale of keypoints; (3) Orientation Assignment that assigns orientations to keypoints; and (4) Keypoint Descriptor that generates descriptors for keypoints. Detecting the keypoints in the SIFT framework starts with convolving the image with Gaussian filters at different scales, and then the difference of successive Gaussian-blurred images is calculated. Keypoints are then obtained as maxima/minima of the Difference of Gaussians (DoG) that occur at multiple scales.

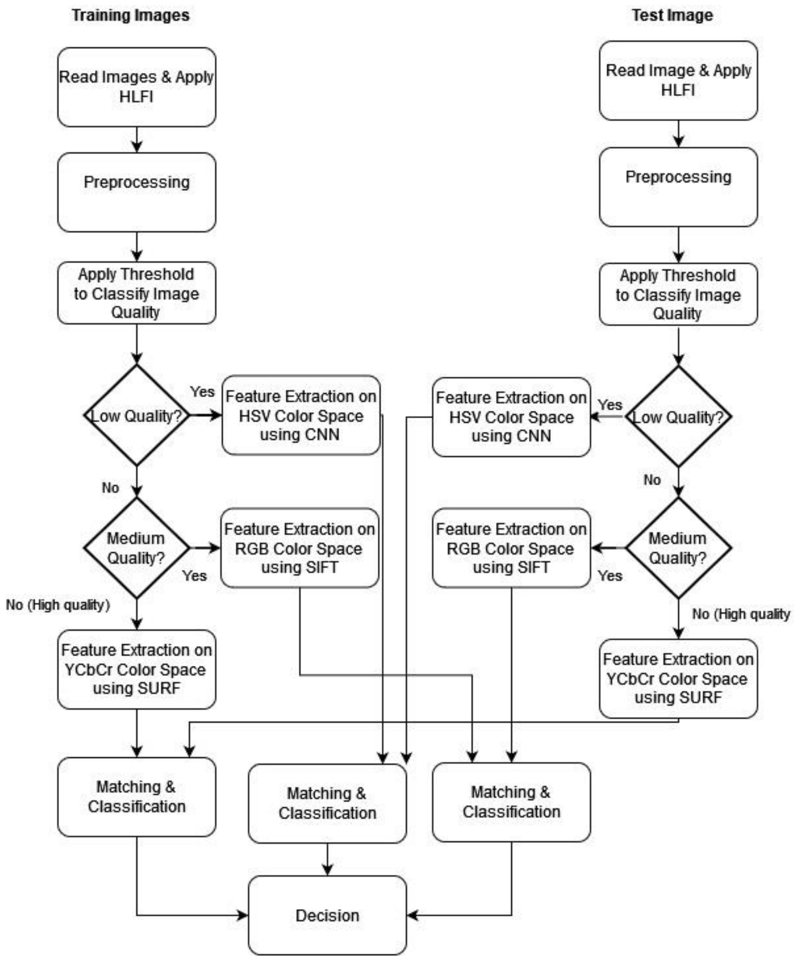


Figure 3. Proposed method flowchart

Gaussian function is the only possible scale-space kernel under a variety of reasonable assumptions. Consequently, the scale space of an image is defined as a function, $L(x, y, \sigma)$, that is produced from the convolution of a variable-scale Gaussian, $G(x, y, \sigma)$, with an input image, $I(x, y)$ as follows:

$$L(x, y, \sigma) = G(x. y, \sigma) * I(x, y), \quad (2)$$

where * is the convolution operation in x and y, and

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}. \quad (3)$$

To detect stable keypoint locations in scale space, scale-space extrema in the difference-of-Gaussian function convolved with the image, $D(x, y, \sigma)$, is proposed [16] that can be computed from the difference of two nearby scales separated by a constant multiplicative factor k as follows:

$$\begin{aligned} D(x, y, \sigma) &= \big(G(x, y, k\sigma) - G(x, y, \sigma)\big) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma). \end{aligned} \quad (4)$$

The function in (4) is a particularly efficient function to compute since a simple image subtraction is performed and the smoothed image, L, can be computed in any case for scale space feature description.

Speeded Up Robust Features (SURF) [17] is a local feature detector and descriptor which is a faster variant of the SIFT approach. The algorithm has three main parts: (1) Interest point detection; (2) Local neighborhood description; (3) Matching. SURF approach detects the interest points using an integer approximation of the determinant of the Hessian blob detector that can be computed with 3 integer operations using a precomputed integral image. The feature descriptor of SIFT is based on the sum of the Haar wavelet response around the point of interest which can also be computed with the integral image. Using the multi-resolution pyramid technique, the image is transformed into coordinates to copy the original image with a Pyramidal Gaussian or Laplacian Pyramid shape. This operation aids in obtaining an image of the same size but with reduced bandwidth. Consequently, a special blurring effect on the original image is achieved which is called Scale-Space. The process helps to produce points of interest that are scale invariant.

Specifically, SURF uses square-shaped filters as an approximation of Gaussian smoothing [17]. Therefore, filtering the image with a square is much faster if the integral image is used as follows:

$$S(x, y) = \sum_{i=0}^{x} \sum_{j=0}^{y} I(i, j) \tag{5}$$

The integral image in (5) helps to quickly evaluate the sum of the original image within a rectangle that requires evaluations at the rectangle's four corners.

To find points of interest with SURF, a blob detector based on the Hessian matrix is used. The determinant of the Hessian matrix is calculated. Afterward, it is used as a measure of local change around the point. The points are selected where the determinant is maximal. Moreover, SURF uses the determinant of the Hessian for choosing the scale. Given a point $p = (x, y)$ in an image I, the Hessian matrix $H(p, \sigma)$ at point p and scale σ, is as follows:

$$H(p, \sigma) = \begin{pmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{xy}(p,\sigma) & L_{yy}(p,\sigma) \end{pmatrix} \tag{6}$$

where $L_{xx}(p, \sigma)$ is the convolution of the second-order derivative of Gaussian with the image $I(x, y)$ at the point p.

Specifically, the box filter of size 9×9 is an approximate of a Gaussian with $\sigma = 1.2$ which represents the lowest level with the highest spatial resolution for blob-response maps.

On the other hand, Convolutional Neural Network (CNN) [22] is a deep learning-based method and an improved version of feedforward neural networks that extracts features from input data using convolution structures. The inspiration for CNN architecture comes from visual perception. Four specific components are needed to build a CNN model as follows: (1) Padding, which enlarges the input with zero values in order not to lose information on the borders of the input image; (2) Convolution, which is a pivotal step for feature extraction using activation functions or convolution kernel; (3) Pooling (or downsampling), that is used to prevent redundancy with max pooling and average pooling; (4) Classification with Fully Connected (FC) Layers, that classifies the input and produces the output. An example of a CNN procedure is shown in Figure 4.

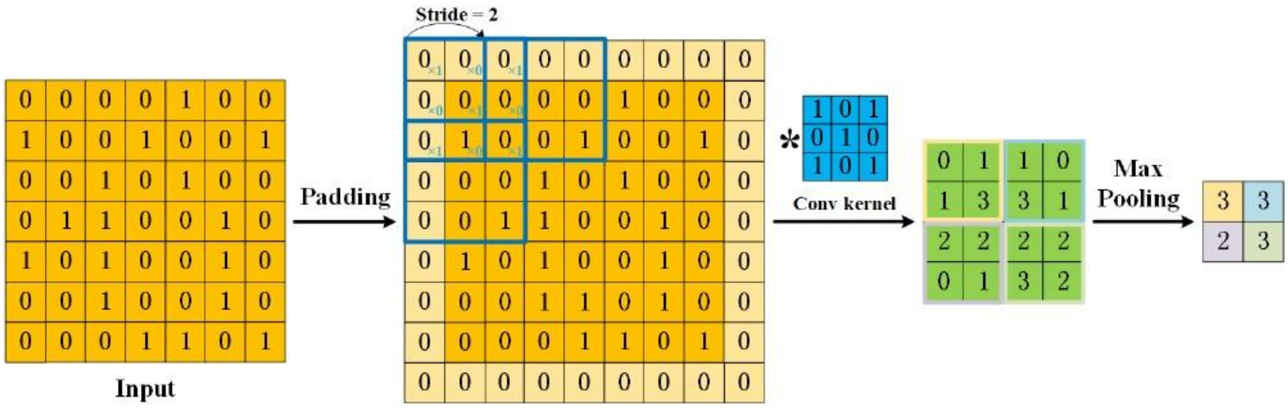


Figure 4. A 2D Convolutional Neural Network Procedure [22]

In the proposed method, SIFT, SURF, and CNN feature extraction and classification approaches are used as follows. The preprocessed images that are classified as low, medium, or high-quality images in the previous steps go through the feature extraction process with different feature extraction methods on different color spaces. According to the preliminary experiments which tried all possibilities based on the implemented feature extractors in this study, the most appropriate feature extractor on a specific color space is selected to be used in the feature extraction stage. Consequently, each color channel from the selected color space is processed by the feature extractor and the average of all channels is considered as the final feature set. In the proposed method, for low-quality images, a deep-learning-based CNN approach is applied to HSV color space, and the resulting feature vectors are used as the output of the feature extraction stage for low-quality images. On the other hand, medium-quality images are processed with a texture-based SIFT method to extract the features on RGB color space. High-quality images are handled using the fast texture-based SURF feature extractor on YcbCr color space. The extracted feature vectors for all train images are stored to be used in the matching and classification stages.

2.3.4. Reading and Processing Test Images

Images in the test phase are read, HLFI is applied to them, and they are preprocessed similarly to training images. HLFI quality measure values are used to specify the condition level of the test images based on the specified thresholds. Consequently, the condition level of the test images is determined as high, moderated, or low condition. Afterward, the same feature extraction methods and color spaces are used for each test image according to the quality level in the same way as applied for train images. Specifically, features of low-quality test images are extracted using the CNN approach on HSV color space; for medium-quality test images, the SIFT method is applied on RGB color space; for high-quality test images, SURF features on YCbCr color space are extracted. Afterward, the extracted feature vectors for each test image are used for the matching and classification processes.

2.3.5. Matching & Decision

In the context of this study, face image classification is facilitated through the application of the Nearest Neighbor classifier [23]. The classification process hinges on a voting mechanism where a face image's classification is determined by the cumulative votes of its neighboring images. Consequently, the face image is allocated to the nearest neighbor that exhibits the highest resemblance.

As a result, the system employs a criterion of selecting the most akin face identification from among the pool of training images as its final decision. The accuracy of this decision is contingent on the concordance between the identification of the test image and the identification made by the system. The decision is considered accurate when the test image's ID aligns with the ID arrived at by the system. The Nearest Neighbor classifier excels in achieving robust performance through the utilization of multiple statistical patterns. Typically, this classifier necessitates comprehensive training with both genuine and counterfeit cases. When classifying a new sample, it calculates the distance to the nearest training case, and the label of the sample is determined by the class of that nearest training case. Expanding on this principle, the Nearest Neighbor classifier goes further by considering the k closest points and assigning the class label that predominates among them. The validation process, involving the input test set, enables the computation of an n-dimensional pattern vector based on the training samples, ultimately leading to classification based on minimum distance.

In the training phase, the examples consist of vectors annotated with class labels. These vectors are stored in the training stage of the nearest neighbor classification, and class labels are assigned accordingly. Notably, in this study, the Manhattan distance measure is employed for matching purposes since it is effective for face matching [20]. Subsequently, the Nearest Neighbor classifier is applied to execute the classification of the test images.

3. Results

In this section, we delve into a series of experiments conducted with the newly proposed method and offer a comparison analysis of its outcomes alongside state-of-the-art methods. The subsequent subsections delve into the finer details of our experimental setup, providing insights into database particulars. Furthermore, we explore the preliminary results of initial experiments that leverage several no-reference image quality assessment techniques to ascertain the image quality. Finally, we present the results of color face recognition experiments using our proposed method.

3.1. Experimental Results

This study places significant importance on the color aspect as a crucial criterion for face image recognition. Color-rich images, consisting of three channels, are stored in diverse formats, each with its unique advantages and drawbacks. The investigation delves into the merits of three distinct color spaces: RGB, HSV, and YcbCr, exploring their potential contributions to the recognition process.

Additionally, the study encompasses three different tiers of image quality levels. This diversity in quality levels allows for the consideration of an index that strikes a balance between ensuring a minimum image quality threshold while optimizing recognition performance. To enhance processing speed and address the challenges posed by growing image datasets, the study tackles the critical issues of program runtime and available storage space. These considerations become pivotal as the number of images escalates, and the study provides a comprehensive discussion on strategies for optimization in this context. The subsequent sections offer a detailed exploration of these aspects.

The objective of this study is to determine the similarity in the face patterns observed after extracting the features using PCA, LBP, SIFT, SURF, and ORB feature extraction methods. Measurement of Manhattan distance using two images' extracted features for face recognition is not efficient on single-channel images or low-quality images. However, it works well with multi-channel images and extracts more useful information. They must also be tested in different color spaces.

3.1.1. Preliminary Experiments using NR-IQM Results

To classify the image quality type as high, low, or medium quality, four No-Reference Image Quality Measurement (NR-IQM) methods are employed on each dataset. The results obtained through these methods, including the (JQI), (HLFI), (BIQI), and (NIQE), are presented in Tables 2 and 3.

Table 2 showcases the experimental outcomes on the Replay Attack datasets, specifically addressing subsets such as Print Attack, Mobile Attack, and High-Def Attack. These subsets contain medium-quality images. Notably, the results reveal that RGB color-based images yield superior outcomes when utilizing the HLFI method. The HLFI method exhibits a wider range of effectiveness across different image types, making it a suitable choice for estimating image quality in this context. Similarly, the HSV color space is investigated for image quality assessment on the Replay Attack dataset, and the results indicate that the HLFI method excels in this color space as well. The extensive range of HLFI's effectiveness in the HSV color space further solidifies its appropriateness for making image quality determinations.

Additionally, the YCbCr color space is explored in the experiments with the Replay Attack dataset, and once again, the HLFI method outperforms other methods across various image types. This robust performance across image types underscores the suitability of the HLFI method for assessing image quality in the YCbCr color space within this specific scenario.

In conclusion, the HLFI method is chosen as the primary measure for assessing image quality due to its consistently low probability of misclassifying images. Table 3 provides an overview of the average No-Reference Image Quality Measurement (NR-IQM) results for images from the Faces94 and ColorFERET datasets, each analyzed in different color spaces. The range of results presented in Table 3 aids in identifying the most suitable color space for each method. Notably, HLFI emerges as the most appropriate measure for evaluating image quality across the considered datasets. This determination is based on the comprehensive evaluation of results encompassing JQI, HLFI, BIQI, and NIQE metrics.

Table 4 provides valuable insights into the no-reference image quality assessment metric HLFI, highlighting minimum, average, and maximum values. These statistics help establish thresholds for categorizing images into different quality types. Here are the key findings from Table 4. High-Quality Images: In high-quality images, the YCbCr color space exhibits a more consistent alternation pattern in HLFI values. Low-Quality Images: For low-quality images, the HSV color space demonstrates greater stability and consistency in HLFI values compared to other color spaces. Medium-Quality Images: When it comes to medium-quality images, the RGB color space outperforms other color spaces based on HLFI values. These observations underscore the importance of selecting the appropriate color space for assessing image quality, depending on the quality type (high, low, or medium) under consideration.

Table 2. Average of NR-IQM Results for Real and Fake Images on Replay Attack Dataset

Dataset	Color space	JQI	BIQI	NIQE	HLFI
Real	RGB	7.70	27.23	2.44	8271132.84
	HSV	17.57	27.35	3.94	35398.87
	YCbCr	7.74	36.77	4.21	5993760.03
Highdef Attack	RGB	7.64	26.74	2.61	9191549.95
	HSV	17.36	26.92	4.29	32722.45
	YCbCr	8.20	37.01	4.52	6101433.80
Mobile Attack	RGB	7.69	38.14	3.00	10565951.74
	HSV	17.82	30.11	4.17	36521.13
	YCbCr	8.29	41.83	4.32	6660354.29
Print Attack	RGB	8.33	32.62	2.25	8966469.80
	HSV	18.50	28.80	3.44	30806.30
	YCbCr	8.33	43.32	4.08	6087265.60
Average of All Images	RGB	7.84	31.18	2.57	9248776.08
	HSV	17.81	28.29	3.96	33862.18
	YCbCr	8.14	39.73	4.28	6210703.43

Table 3. Average of NR-IQM Results for Real and Fake Images for Faces94 and ColorFERET Databases

Dataset	Color space	JQI	BIQI	NIQE	HLFI
Faces94	YCbCr	8.87	47.97	4.73	2483901.57
	HSV	19.15	36.76	4.25	13629.71
	RGB	8.76	32.54	3.27	2982913.14
ColorFERET	YCbCr	12.61	49.35	3.71	33056379.03
	HSV	11.37	30.12	3.71	17199617.13
	RGB	10.97	29.69	3.36	49998026.51

Table 4. HLFI Values for Specifying Thresholds

Color space	High quality			Medium quality			Low quality		
	Max	Average	Min	Max	Average	Min	Max	Average	Min
YCbCr	65M	33M	2M	9-11M	5-6M	3-4M	3M	2M	1M
HSV	26M	17M	3M	45-58K	30-36K	7-14K	21K	13K	4K
RGB	73M	49M	16M	11-13M	8-10M	2-6M	73M	2M	765k

3.1.2. Face Recognition Results

The face recognition experiments conducted on three datasets using seven feature extraction methods have yielded results summarized in Table 5. Here is an overview of the experimental outcomes for each feature extraction method on each dataset:

Replay Attack Dataset: This dataset utilized 2 training images (1 real and 1 fake) and 2 testing images (1 real and 1 fake) for the experiments. Faces94 Dataset: Experiments on the Faces94 dataset involved 10 training and 10 testing images. ColorFERET Dataset: The ColorFERET dataset experiments were performed with 1 training and 1 testing scenario. The results are presented in terms of recognition rate, calculated as specified in the document. These recognition rates provide insights into the performance of various feature extraction methods on different datasets, facilitating the evaluation of their effectiveness in face recognition tasks:

$$Recognition\ rate(\%) = \frac{Number\ of\ correctly\ recognized\ test\ images}{Total\ number\ of\ test\ images} * 100 \quad (7)$$

According to the presented results in Table 5, SIFT achieves the best recognition rate (93%) on the Replay Attack dataset while SURF performed well (96.26%) on the ColorFERET dataset. CNN works very well on the Faces94 dataset with a 100% recognition rate. The proposed method combines the advantages of all these superior achievements and produces the best results on all datasets. Therefore, the recognition rates of the proposed method are superior in general on all datasets. The graphical representation of the results is also demonstrated in Figure 5.

The results of face recognition experiments using a deep-learning-based CNN model are shown in Table 6. The identification rate is low on the Replay Attack dataset because the number of images used in that dataset is low, it includes both real and fake images, and it needs to perform data augmentation to increase the number of images to achieve better results with this model. However, with Faces94 and ColorFERET datasets, the accuracies (recognition rates) are very high and the model can be used for face recognition.

3.2. Comparison with the State-of-the-art

The performance of the suggested face recognition method is compared to state-of-the-art studies on three different datasets, as presented in Tables 7, 8, and 9. On Replay Attack Dataset, as shown in Table 7, the proposed method achieves a remarkable accuracy of 100%, making it comparable to the top-performing studies by Yu et al. [24] and Benlamoudi et al. [25], which also achieved 100% accuracy. This demonstrates the excellence of the proposed approach in recognizing faces under spoofing attacks.

On Faces94 Dataset which is presented in Table 8, the proposed method excels with an accuracy of 100%, outperforming other studies. The closest result to the proposed method is from the study by Karanwal and Diwakar [26], which achieved an accuracy of 98.89%. This highlights the effectiveness of the proposed method for face recognition in the Faces94 dataset. The proposed method exhibits superior performance on the ColorFERET dataset, as shown in Table 9, with an accuracy of 96.26%. It outperforms state-of-the-art studies, with the closest performance achieved by Terhörst et al. [27] at 95.9%. This underscores the capability of the proposed approach in recognizing faces in color images. Overall, the proposed method consistently achieves high accuracy levels and demonstrates competitiveness with state-of-the-art studies across different datasets, showcasing its effectiveness in face recognition tasks. In general, three different databases are used in this study and the results show that the proposed method is superior to the other methods across all datasets with 100%, 100%, and 96.26% on Replay Attack, Faces94, and ColorFERET datasets, respectively.

Table 5. Face Recognition Results

Dataset	Number of Face Images	Recognition Rate (%)							
		CNN	LBP	Color LBP	SIFT	SURF	ORB	PCA	Proposed Method
Faces94	3040 real	100	99.86	99.86	43.10	42.03	40.72	99.53	100
Color Feret	536 real	92.50	57.46	92.91	89.55	96.26	88.80	61.19	96.26
Replay Attack	300 fake 100 real	12.90	30.00	63.00	93.00	85.66	77.33	19.00	100

Table 6. CNN Accuracy and Loss Results

Dataset	Train		Test	
	Loss (%)	Accuracy (%)	Loss (%)	Accuracy (%)
Replay Attack	0.48	84.5	9.15	12.99
Faces94	~0	100	~0	100
Color Feret	0.04	100	0.42	92.53

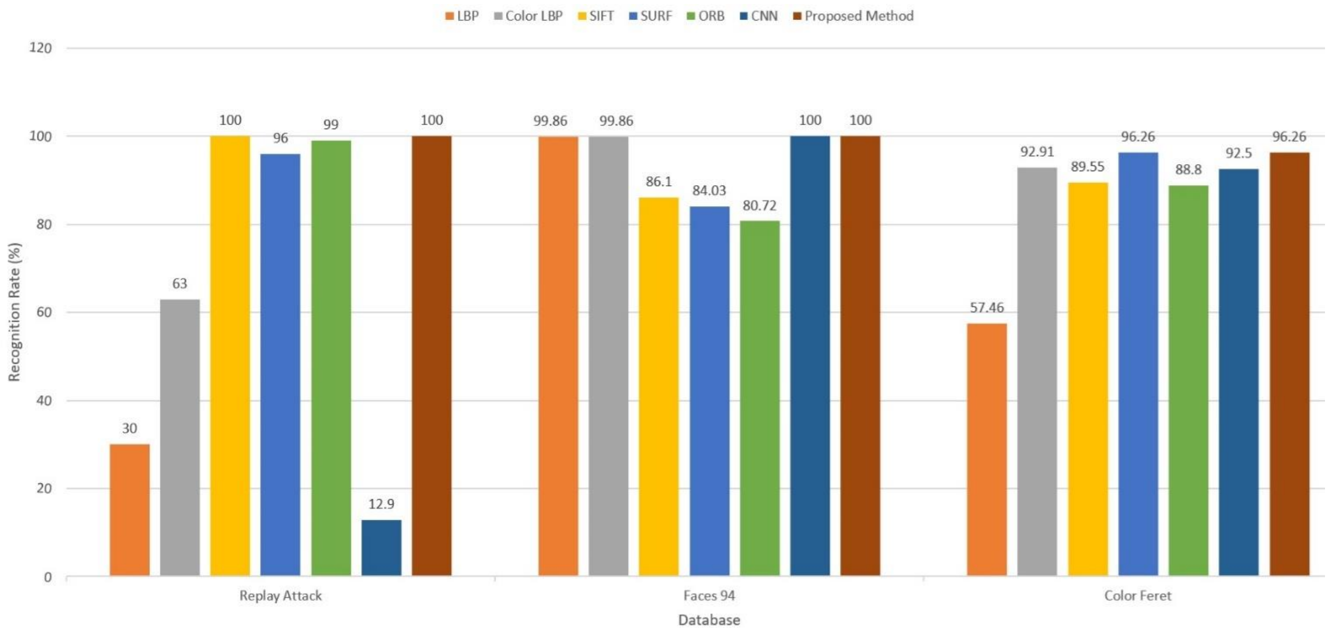


Figure 5. Face Recognition Rates on Replay Attack, Faces94, and ColorFERET Datasets

Table 7. Comparison with the State-of-the-art Face Recognition on Replay Attack Dataset

Reference	Year	Experimental Setup	Methods	Accuracy (%)
Zhang et al. [28]	2020	Train:800, Test:800	CNN	96.6%
Boulkenafet et al. [29]	2017	Train:240, Test:240	SVM+LBP	98.17%
Yu et al. [24]	2020	Train:1300, Test:500	CDCN++	99.89%
Yu et al. [24]	2020	Train:1300, Test:500	CDCN	100%
Benlamoudi et al. [25]	2022	Train:650, Test:650	BS-CNN+MV	100%
Proposed Method	2023	Train:200, Test:200	SIFT	100%

Table 8. Comparison with the State-of-the-art Face Recognition on the Faces94 Dataset

Reference	Year	Experimental Setup	Methods	Accuracy (%)
Chen et al. [31]	1998	Train:768, Test:768	LDA	79.39%
Karanwal [33]	2021	Train:1530, Test:510	CLBP	89.14%
Vinay et al. [30]	2018	Train:2295, Test:765	LBP	93.6%
Sikder et al. [32]	2021	Train:572, Test:52	PCA	94.4%
Tran et al. [34]	2017	Train:608, Test:2432	LIAD	98.16%
Karanwal and Diwakar [26]	2021	Train:1360, Test:680	CMBZZBP	98.89%
Proposed Method	2023	Train:1520, Test:1520	CNN	100%

Table 9. Comparison with the State-of-the-art Face Recognition on the ColorFERET Dataset

Reference	Year	Experimental Setup	Methods	Accuracy (%)
Chou et al. [35]	2019	Train:675, Test:675	Deep CNN	83.79%
Zhao et al. [36]	1998	Train:1316, Test:298	PCA	95%
Geng and Jiang [37]	2009	Train:1340, Test:1340	LBP	95%
Du et al. [38]	2009	Train:804, Test:268	SURF	95%
Terhörst et al. [27]	2020	Train:9125, Test:5000	kNN	95.9%
Proposed Method	2023	Train:268, Test:268	SURF	96.26%

4. Conclusion

Securing sensitive data has become a challenging task leading to significant research in biometrics. Within the realm of biometrics, robust and secure face recognition plays a pivotal role. According to this study, image quality is assessed by utilizing No-Reference Image Quality Assessment Measures (NR-IQMs). Four NR-IQMs are employed in this study, namely the Blind Image Quality Index (BIQI), JPEG Quality Index (JQI), Naturalness Image Quality Estimator (NIQE) and High-Low Frequency Index (HLFI). Besides, various feature extraction methods are employed including ColorLBP, LBP, SIFT, SURF, ORB, and PCA to extract features from both color and grayscale images. Additionally, a deep learning approach using CNN architecture is applied when a substantial volume of image data is available. The experiments encompass three different color spaces (YCbCr, HSV, and RGB) across three face databases (ColorFERET, Faces94, and Replay Attack).

The key findings from the study are as follows: The High-Low Frequency Index (HLFI) quality measure excels in determining image quality, surpassing other no-reference measures. For medium-quality images, the texture-based SIFT feature extraction method combined with RGB color space is found to be the most suitable. Deep-learning-based CNN demonstrates high performance on low-quality images, particularly in the HSV color space. The YCbCr color space, in conjunction with the SURF feature extraction method, proves to be the triumphant approach for high-quality images. As a result, the proposed approach leverages these insights and combines them into a hybrid algorithm that capitalizes on these findings.

Specifically, the proposed hybrid face recognition algorithm starts with reading the color face images and calculating the HLFI quality measure to decide on the quality of the face image as low-quality, medium-quality, or high-quality image. For low-quality images, the proposed method employs the CNN approach on the images in HSV color space. Mediaum-quality images are processed in RGB color space with the SIFT feature extractor. High-quality images are used in YcbCr color space and their features are extracted with the SURF method.

The proposed hybrid color face recognition method achieves 100% accuracy on the Replay Attack dataset with the SIFT feature extractor employed for medium-quality images of the dataset. Additionally, the accuracy of the proposed method on low-quality images of the Faces94 dataset is 100% by employing a deep learning-based CNN approach. Moreover, high-quality images in the ColorFERET dataset are recognized by the proposed method which employs a SURF feature extractor with 96.26% accuracy. The outcomes of the proposed method are promising, outperforming extant state-of-the-art face recognition approaches on the same datasets. Future research could explore further enhancements in face recognition algorithms by incorporating diverse and robust techniques for extracting features in other color spaces and investigating the potential of texture-based and deep learning-based techniques for an even more advanced face recognition system.

References

[1] A. K. Jain, A. A. Ross, and K. Nandakumar, “Introduction to Biometrics”, Springer Publishing Company, Incorporated, 2011.

[2] R. Szeliski, “Computer Vision: Algorithms and Applications”, 1st. ed., Springer-Verlag, Berlin, Heidelberg, 2010.

[3] M. O. Oloyede, G. P. Hancke, H. C. Myburgh, “A review on face recognition systems: recent approaches and challenges”, Multimed Tools Appl, 79, pp. 27891–27922, 2020. https://doi.org/10.1007/s11042-020-09261-2

[4] M. K. Rusia, D. K. Singh, “A comprehensive survey on techniques to handle face identity threats: challenges and opportunities”, Multimed Tools Appl, 82, pp. 1669–1748, 2023. https://doi.org/10.1007/s11042-022-13248-6

[5] J. Yang, D. Zhang, Y. Xu, and J. Y. Yang, “Recognize color face images using complex Eigenfaces”, Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 3832 LNCS, pp. 64–68, 2006, doi: 10.1007/11608288_9.

[6] A. W. Yip and P. Sinha, “Contribution of color to face recognition”, Perception, vol. 31, no. 8, pp. 995–1003, 2002, doi: 10.1068/p3376.

[7] Q. B. Sun, W. M. Huang, and J. K. Wu, “Face detection based on color and local symmetry information”, Proc. - 3rd IEEE Int. Conf. Autom. Face Gesture Recognition (FG1998), pp. 130–135, 1998, doi: 10.1109/AFGR.1998.670937.

[8] D. H. Brainard and B. A. Wandell, “Asymmetric color matching: how color appearance depends on the illuminant”, J. Opt. Soc. Am. A, vol. 9, no. 9, p. 1433, 1992, doi: 10.1364/josaa.9.001433.

[9] M. Pedersen and J. Y. Hardeberg, “Full-Reference Image Quality Metrics: Classification and Evaluation”, Foundations and Trends in Computer Graphics and Vision, vol. 7, no. 1, pp. 1–80, 2012, doi: 10.1561/0600000037.

[10] L. Liu, B. Liu, H. Huang and A. C. Bovik, “No-reference image quality assessment based on spatial and spectral entropies”, Signal Process. Image Commun., vol. 29, no. 8, pp. 856–863, 2014, doi: 10.1016/j.image.2014.06.006.

[11] I. Chingovska, A. Anjos, S. Marcel, "On the Effectiveness of Local Binary Patterns in Face Anti-spoofing"; IEEE BIOSIG, 2012, https://ieeexplore.ieee.org/document/6313548

[12] Facial Images: Faces94, Computer Vision Science Research Projects website, Designed and maintained by Dr Libor Spacek on 13th June 2009, https://cmp.felk.cvut.cz/~spacelib/faces/faces94.html

[13] Color FERET Database, National Institute of Standards and Technology (NIST) website, Designed by P. Jonathon Phillips, Created January 31, 2011, Updated December 3, 2019, https://www.nist.gov/itl/products-and-services/color-feret-database

[14] N. Khediri, M. Ammar and M. Kherallah, “Comparison of Image Segmentation using Different Color Spaces”. 2021 IEEE 21st International Conference on Communication Technology (ICCT), pp. 1188-1192, 2021, doi: 10.1109/ICCT52962.2021.9658094.

[15] S. Banerji, A. Verma, and C. Liu, “Novel color LBP descriptors for scene and image texture classification”, Proc. 2011 Int. Conf. Image Process. Comput. Vision, Pattern Recognition, IPCV 2011, vol. 2, pp. 537–543, 2011.

[16] D.G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”. International Journal of Computer Vision, vol. 60, pp. 91–110, 2004, https://doi.org/10.1023/B:VISI.0000029664.99615.94

[17] H. Bay, A. Ess, T. Tuytelaars, L.V. Gool, "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), vol. 110, no. 3, pp. 346--359, 2008.

[18] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “ORB: An efficient alternative to SIFT or SURF”, Proc. IEEE Int. Conf. Comput. Vis., pp. 2564–2571, 2011, doi: 10.1109/ICCV.2011.6126544.

[19] S. Kakarla, P. Gangula, M. S. Rahul, C. S. C. Singh, and T. H. Sarma, "Smart Attendance Management System Based on Face Recognition Using CNN", 2020 IEEE-HYDCON, Hyderabad, India, 2020, pp. 1-5, doi: 10.1109/HYDCON48903.2020.9242847.

[20] A. Afaneh, F. Noroozi and Ö. Toygar, “Recognition of Identical Twins Using Fusion of Various Facial Feature Extractors”, EURASIP Journal on Image and Video Processing, vol. 2017:81, pp.1-14, Dec. 2017.

[21] J. Galbally, S. Marcel, and J. Fierrez, "Image Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint and Face Recognition", IEEE Transactions on Image Processing, vol. 23, no. 2, pp. 710.724, Feb. 2014, doi: 10.1109/TIP.2013.2292332.

[22] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," in IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 12, pp. 6999-7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.

[23] T.K. Ho, "Nearest Neighbors in Random Subspaces", In: A. Amin, D. Dori, P. Pudil, H. Freeman (eds), Lecture Notes in Computer Science, Springer, Germany, pp 640–648, 1998.

[24] Z. Yu et al., "Searching central difference convolutional networks for face anti-spoofing", Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 5294–5304, 2020, doi: 10.1109/CVPR42600.2020.00534.

[25] A. Benlamoudi et al., "Face Presentation Attack Detection Using Deep Background Subtraction", Sensors, vol. 22, no. 10, 2022, doi: 10.3390/s22103760.

[26] S. Karanwal and M. Diwakar, "Two novel color local descriptors for face recognition", Optik (Stuttg)., vol. 226, 2021, doi: 10.1016/j.ijleo.2020.166007.

[27] P. Terhörst, M. Huber, N. Damer, F. Kirchbuchner, and A. Kuijper, "Unsupervised Enhancement of Soft-biometric Privacy with Negative Face Recognition", arXiv:2002.09181v1 [cs.CV], 2020, [Online]. Available: http://arxiv.org/abs/2002.09181.

[28] B. Zhang, B. Tondi, and M. Barni, "Adversarial examples for replay attacks against CNN-based face recognition with anti-spoofing capability", Comput. Vis. Image Underst., vol. 197–198, 2020, doi: 10.1016/j.cviu.2020.102988.

[29] Z. Boulkenafet, J. Komulainen, L. Li, X. Feng, and A. Hadid, "OULU-NPU: A Mobile Face Presentation Attack Database with Real-World Variations", 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), Washington, DC, USA, 2017, pp. 612-618, doi: 10.1109/FG.2017.77.

[30] A. Vinay, A. Gupta, A. Bharadwaj, A. Srinivasan, K. N. B. Murthy, and S. Natarajan, "Deep Learning on Binary Patterns for Face Recognition", Procedia Comput. Sci., vol. 132, pp. 76–83, 2018, doi: 10.1016/j.procs.2018.05.164.

[31] L. F. Chen, H. Y. M. Liao, M. T. Ko, J. C. Lin, and G. J. Yu, "A New LDA-based Face Recognition System Which Can Solve the Small Sample Size Problem", Proc. Jt. Conf. Inf. Sci., vol. 4, pp. 282–286, 1998.

[32] J. Sikder, R. Chakma, R. J. Chakma, and U. K. Das, "Intelligent Face Detection and Recognition System", 2021 Int. Conf. Intell. Technol. CONIT 2021, 2021, doi: 10.1109/CONIT51480.2021.9498291.

[33] S. Karanwal, "A comparative study of 14 state of art descriptors for face recognition", Multimed. Tools Appl., vol. 80, no. 8, pp. 12195–12234, 2021, doi: 10.1007/s11042-020-09833-2.

[34] C.-K. Tran et al., "Local intensity area descriptor for facial recognition in ideal and noise conditions", J. Electron. Imaging, vol. 26, no. 2, p. 023011, 2017, doi: 10.1117/1.jei.26.2.023011.

[35] H. R. Chou, J. H. Lee, Y. M. Chan, and C. S. Chen, "Data-Specific Adaptive Threshold for Face Recognition and Authentication", Proc. - 2nd Int. Conf. Multimed. Inf. Process. Retrieval, MIPR 2019, pp. 153–156, 2019, doi: 10.1109/MIPR.2019.00034.

[36] W. Zhao, A. Krishnaswamy, R. Chellappa, D. L. Swets, and J. Weng, "Discriminant Analysis of Principal Components for Face Recognition", In: Wechsler, H., Phillips, P.J., Bruce, V., Soulié, F.F., Huang, T.S. (eds) Face Recognition. NATO ASI Series, vol 163. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-72201-1_4

[37] C. Geng and X. Jiang, "Face recognition using SIFT features", Proc. - Int. Conf. Image Process. ICIP, pp. 3313–3316, 2009, doi: 10.1109/ICIP.2009.5413956.

[38] G. Du, F. Su, and A. Cai, "Face recognition using SURF features", MIPPR 2009 Pattern Recognit. Comput. Vis., vol. 7496, p. 749628, 2009, doi: 10.1117/12.832636.

Acknowledgments

Authors would like to thank Idiap Research Institute in Switzerland for providing the Replay Attack database; Dr. Libor Spacek for providing the Faces94 database; and Dr. P. Jonathon Phillips from the National Institute of Standards and Technology for providing the ColorFERET database.

Authors Contributions

Mohammad Mehdi PAZOUKI: Performed experimental analysis and wrote the paper.
Önsen TOYGAR: Developed and designed the analysis and reviewed the paper.
Mahdi HOSSEINZADEH: Collected and prepared data and wrote the paper.

Conflicts of Interest

Authors declare no conflict of interest.

Ethical Approval

It is declared that during the preparation process of this study, scientific and ethical principles were followed, and all the studies benefited from are stated in the bibliography.

Availability of data and material
Not applicable.

Plagiarism Statement
This article has been scanned by iThenticate ™.