DOI: 10.15649/2346030X.3461
Aibi revista de investigación, administración e ingeniería. Volume 12, Number 1, Pág 152-162, January - April 2024, ISSN 2346-030X

Assessment of face image quality for improvement of recognition system with machine learning

Parul Choudhary[1], Pooja Pathak[2], Phalguni Gupta[3]
[1,2,3]GLA University, *Mathura - India*
ORCID: [1]*0009-0006-2490-4110*, [2]0000-0002-8381-6405, [3]0000-0002-4122-4970

Received: August 06, 2023.
Accepted: November 02, 2023.
Published: January 01, 2024.

***Abstract*—** Face recognition systems are crucial in security and surveillance. However, their effectiveness depends on the quality of input face images. This paper introduces a novel Face Image Quality Assessment (FIQA) framework to enhance recognition accuracy. It evaluates image quality using various factors and combines them to classify the face quality index. The technique is evaluated on the Color-Feret dataset, outperforming existing methods with a strong correlation (0.95) between system and human scores.

Keywords: face, neural networks, machine learning.

*Corresponding author.
Email: cparul2605@gmail.com (Parul Choudhary).
Peer reviewing is a responsability of the Universidad de Santander.

How to cite this article: P. Choudhary, P. Pathak y P. Gupta, "Assessment of face image quality for improvement of recognition system with machine learning", *Aibi research, management and engineering journal*, vol. 12, no. 1, pp. 152-162 2024, doi: 10.15649/2346030X.3461

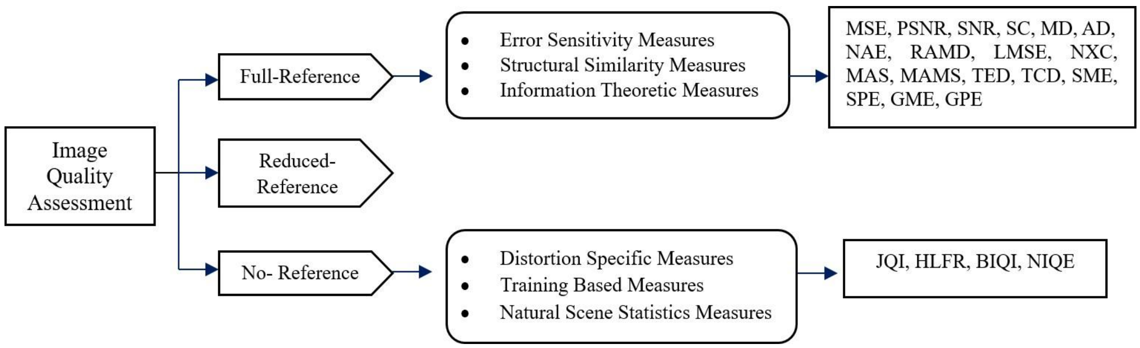
I. INTRODUCTION

Image quality assessment (IQA) is a critical research domain in the field of computer vision and image processing. As the demand for high-quality images continues to grow across various applications, including multimedia, medical imaging and autonomous systems, the need for accurate and reliable methods to assess image quality becomes increasingly important.

There are various image quality factors such as focus, contrast, etc. that gives high influence on performance of biometric systems. By feeding a low-quality image to the system, the system's accuracy will suffer; in contrast to the high quality of images which will improve the recognition rate. Thus, it is significant to determine an image quality before processing to improve system matching performance [1]. Basic three general terminologies related to such topic are Quality parameters, also known as quality factors, include pose, focus, etc.; Quality metrics, often referred to as performance indicators or measurement parameters; they are also known as "quality measures," and they are divided into two types: specific measures that focus on issues related to the face [2-4], iris [5], and fingerprint [6] and generic measures that are general in nature. And quality index which is a solitary number used to represent an overall quality of image.

According to [7], FIQA can be classified in three ways: photographically, scene-related and digitally. Photographically-oriented FIQA methods center their evaluation criteria on the inherent characteristics of the photographic process, emphasizing factors that pertain directly to the acquisition and capture of face images. Key elements encompassed within this category include- sharpness, illumination, etc. Factors related to scene requirements are image background, illumination, glass, Visualization of mouth and eyes, etc. which expand their perspective beyond the technical attributes of the face image itself, incorporating environmental and contextual factors. Meanwhile, digitally formatted methods concentrate on the manipulation and processing of digital face images including factors such as grayscale image contrast, resolution, etc.

Over the past few decades, IQA has drawn much more attention and become popular among technologies. As shown in the Fig.1, there are three kinds of IQA: Full-Reference IQA methods [8-11] compare the quality of an image to a reference (original) image with known quality. These methods aim to quantify the extent of distortion between the reference and the distorted image; Reduced-Reference IQA approaches [12-13] strike a balance between full-reference and no- reference methods by using partial information from the reference image. These methods extract and compare specific features such as edge information, color histograms or local structural patterns, to estimate image quality and No-Reference IQA methods [14-18] assess image quality without access to any reference image. These approaches rely on statistical models, machine learning or perceptual models to predict image quality based on the characteristics of the distorted image itself. NR IQA is especially useful in scenarios where a reference image is unavailable such as assessing image quality in legacy systems or user-generated content on the internet.



Fiura 1: Types of Image Quality Assessment [1].
Source: Elaborated based on contributions from [1].

It's worth mentioning that no single IQA metric is universally applicable to all scenarios as IQA depends on various factors like image content, distortions and the viewing conditions. Researchers continue to develop new metrics and algorithms to improve the accuracy of IQA systems. Additionally, IQA techniques can also be combined with machine learning and deep learning approaches to enhance their performance and adaptability to specific applications.

The well-organized research paper is crucial for effectively conveying ideas and supporting such arguments. Hence, an organization of this research is mentioned in Fig. 2.

Introduction- this section explores IQA and its various types, followed by an overview of the overall paper structure.

Literature review- an essential background information is provided on basis of literature review. Methodology- explicates on the research design, dataset description, the specification of quality factors and metrics, and a comprehensive exploration of multiple fusion schemes. Experiment- briefly contrasts the recognition system with the obtained results. Finally, conclude the paper and provide a list of references.

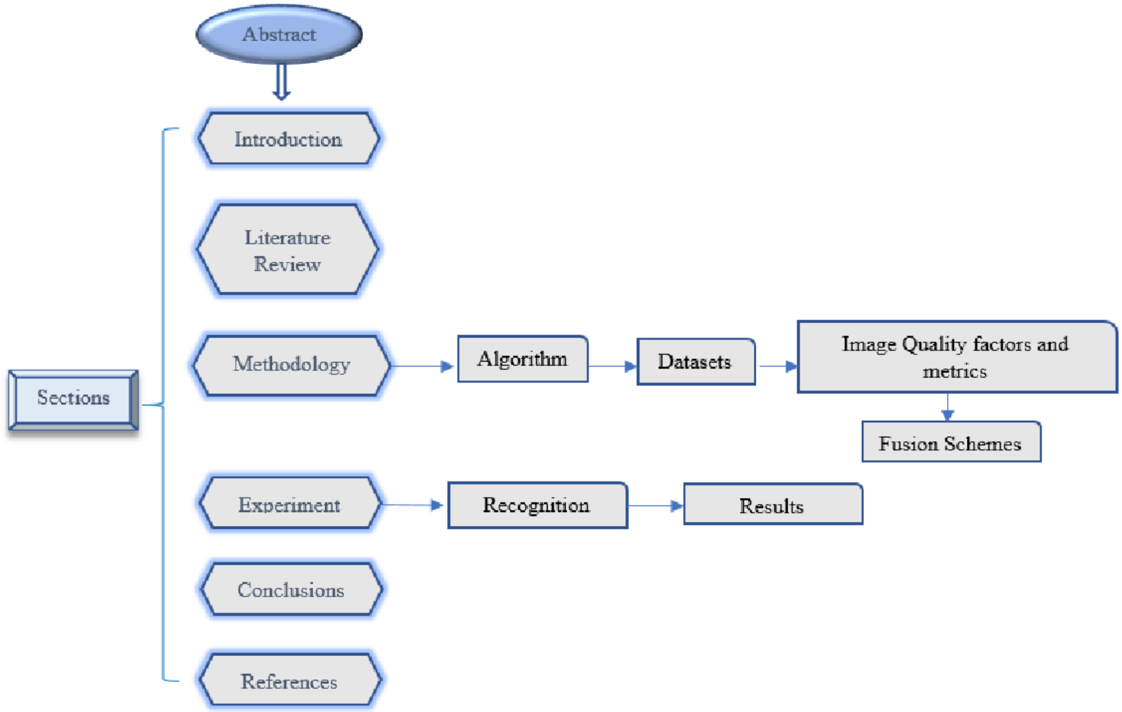


Figure 2: Organization of Paper.
Source: Own elaboration.

II. THEORETICAL FRAMEWORK

During these days, various fusion-based techniques are used to combine several quality factors- blurring, brightness, contrast, focus, etc. to predict quality of image. Some studies are done for FQA mentioned below and evaluated in Table 1.

- In the study of Hsu et al. [19], a comprehensive assessment of quality factors such as exposure, focus, pose, illumination and others is undertaken. This assessment integrates these diverse factors through the utilization of a neural network framework, culminating in the prediction of an authenticity or genuineness score.
- Nasrollahi et al. [20], a focused solely on four discerning factors, namely, brightness, sharpness, out-of-plane rotation and resolution as determinants of video sequence quality enhancement.
- Wong et al. [21] have devised a method designed to address a spectrum of challenges encompassing blurriness, cast shadows, pose variations and alignment errors among others.
- Aggarwal et al. [22] conducted an evaluation encompassing factors such as illumination, pose and expression, among others. The Multi-Dimensional Scaling (MDS) technique is subsequently employed to map these factors, facilitating the derivation of a genuine score, which serves as a decisive indicator of the system's success or failure.
- Phillips et al. [23] conducted an extensive quality assessment by integrating twelve distinct measures, including but not limited to exposure time, eye distance, face saturation and standard deviation. These measures are then employed as features within a Support Vector Machine (SVM) model, enabling the prediction of image quality in binary values, indicative of either successful quality or its absence.
- Dutta et al. [24] conducted training of a Probability Density Function (PDF) model to establish the relationship between image quality and recognition performance. This training process involved the integration of factors such as pose, illumination, noise, sharpness, etc. The outcome of this modelling endeavour was the determination of the False Rejection Rate (FRR) as a critical metric in the context of recognition systems.
- Abaza et al. [25] employed a neural network with a single layer to facilitate the amalgamation of quality measures, incorporating elements such as focus, contrast, brightness, and various others.
- Vignesh et al. [26], their proposed model leverages a Convolutional Neural Network (CNN) architecture to concurrently address both face recognition and quality prediction tasks.
- Kim et al. [27] undertook an investigation that encompasses factors such as blurriness, pose and brightness which are interpreted as constituents of visual quality. Furthermore, the study investigates into the issue of mismatch between training and testing images.
- Ning Zhuang et al. [28] conducted training of a Deep Convolutional Neural Network (DCNN) model, specifically VGG16, which incorporates a comprehensive set of factors comprising pose, blurring, occlusion, contrast and brightness. The primary objective of this model is to predict image quality and categorize it as either "good" or "bad."
- Bharadwaj et al. [29] combined a set of features encompassing pose, illumination, low- resolution, expression, and occlusion. This combination was achieved through the application of a Multi-Support Vector Machine (Multi-SVM) technique. The outcome of this approach was the categorization of image quality into four distinct levels, namely, poor, fair, good, and excellent.
- In references [30, 31, 32, and 33], the application of tattoos or paint is employed as a strategy to conceal specific facial regions with particular attention to mitigating the adverse effects of occlusion on Face Recognition (FR) systems.
- In response to the ongoing COVID-19 pandemic, various public and private establishments have directed the widespread use of facial masks as a preventive measure to mitigate virus transmission, as documented in reference [34]. However, it is noteworthy

that the use of facial masks has introduced significant challenges to the efficacy of face recognition systems, thereby interpreting automated face recognition considerably more complicated.

- A novel Blind Face Image Quality Assessment (FIQA) model, denoted as GFIQA-20k, is introduced, utilizing generative face priors [35]. Furthermore, Jo et al. [36] propose an alternative method for assessing the perceptual authenticity of face images. However, it is noteworthy that there remains a relative scarceness of research attention directed toward the objective and subjective quality evaluation of compressed facial video content.

Table 1: Related Work Summary for Automatic Face Quality Assessment.

Study (Year)	Dataset	Predicted QualityValue	Learning Method
Hsu et al. [19], 2006	1886 (N/A): FRGC 2000(N/A): Passports 1996(N/A): Mugshots	Continuous produced genuine score	27 quality measures (exposure, focus, position, illumination, etc.) will be collected by a neural network to predict real scores.
Nasrollahi et.al. [20], 2008	FRI CVL, Hermes project	Quality (1 to 5)	FQA system obtained values for out-of-plan rotation, sharpness, brightness, resolution.
Y. Wong et al. [21], 2011	Face datasets (FERET and PIE)	In order to get sorted quality of images	Patch-based FQA algorithm obtained pose variations, cast shadows, blurriness, and alignmenterrors.
Aggarwal et al. [22], 2011	FacePix:1830(30) multi-PIE:6740(337)	Continuous (genuine score) or binary (success vs. failure)	MDS learns mapping from illumination, pose, expression, etc. factor to genuine score. This score compares to decide whether it is success or failure.
Phillips et al. [23], 2013	GU+: 4340(437) PaSC:4688(n/a)	Binary (low vs. high)	PCA + LDA determine EdgeDensity, Face saturation, Exposure time, Eye distance, Focus, Standard deviation, Illumination direction, etc.
Dutta et al. [24], 2014	MultiPIE: 3370(337)	Continuous (FRR)	PDF between quality and recognition (pose, illumination, noise, sharpness, etc.)
Abaza et al. [25], 2014	GU+ :4340(437)	Binary (good vs. ugly)	Collecting all value metrics of contrast, brightness, sharpness, focus, and lighting using a single layer neural network
Vignesh et al. [26], 2015	Chokepoint:48 videos	Continuous (genuine score)	CNN
Kim et al. [27], 2015	FRGC: 10,448 (322)	Binary (low vs. high)	All features pose, blur, brightness is fed to AdaBoost binary classifier and relative (colour mismatch between train and test images) facial image quality information.
Ning Zhuang et al. [28], 2019	Color FERET, KinectFace	Binary (Good vs. Bad)	DCNN-VGG16 to combine Blurring, Contrast, Occlusion,Pose, Brightness.
Bharadwaj et.al [29], 2019	CAS-PEAL and SC Face	Quality (poor, fair, good, excellent)	Multi-class SVM combines features (illumination, expression, pose, low-resolution and occlusion)

Source: Own elaboration.

III. METHODOLOGY

a. Proposed work

The Image Quality Index (IQI) is a numerical measure employed to assess the perceptual quality of digital images in comparison to a reference image or a predefined ideal. It encapsulates the visual fidelity of an image by considering various factors including artifacts, noise, sharpness, and color accuracy. The calculation of IQI typically involves a combination of objective quality metrics each quantifying specific aspects of image quality. Commonly used metrics encompass the Structural Similarity Index (SSIM), Peak Signal-to-Noise Ratio (PSNR), etc. Therefore, in this study, author's look at several factors to determine the suggested FQI and demonstrate how it can support utilization through a series of experiments. Fig. 3 shows concept of block diagram of proposed model. Given dataset followed by converting it into grayscale in pre-processing; calculate region of interest to get only essential features of image, also quality factors are evaluated. Next to get a single quality index apply clustering method. Additionally, it's worth mentioning that image quality can be subjective and different individuals may have different preferences and perceptions of what constitutes a "good" image. Therefore, objective metrics are quit useful, but they might not capture all aspects of image quality that a human observer might consider important. Hence collecting the data from three human and ultimately obtain correlation between the fusion scheme and human score.

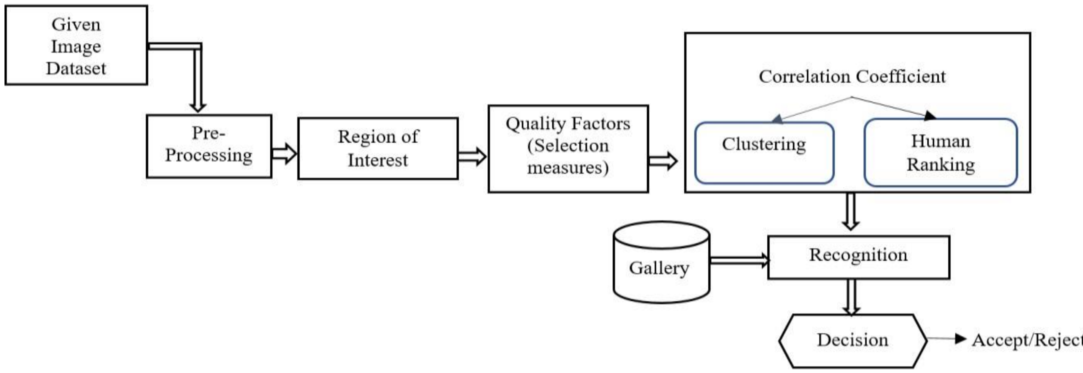


Figure 3: Block Diagram of Proposed system.
Source: Own elaboration.

b. Algorithm

Step 1: Data Acquisition

- Acquire an unlabeled dataset comprising grayscale images for analysis.

Compute seven distinct quality factors.

Step 2: Quality Factor Computation

- Contrast: Measures the difference in intensity between objects in the image.
- Brightness: Quantifies the overall luminance or illumination of the image.
- Sharpness: Assesses the clarity and presence of fine details.
- Illumination: Evaluates the distribution and consistency of lighting.
- Focus: Determines the degree of focus and image sharpness.
- Occlusion: Detects any obstructions in the image.

Step 3: Clustering Techniques

- Apply a variety of clustering techniques to the computed quality factors. These techniques include
- KMeans.
- Bisect Means.
- Agglomerative.
- DBSCAN (Density-Based Spatial Clustering of Applications with Noise).
- Affinity-Propagation.
- Gaussian-Mixture.
- Birch (Balanced Iterative Reducing and Clustering using Hierarchies).

Step 4: Determining the Best Clustering Technique

- Evaluate the clustering techniques by assessing inter-cluster and intra-cluster algorithms to identify the most suitable technique for grouping the images effectively.

Step 5: Human-Eye Models for IQA

- Employ three different human-eye models tailored for Image Quality Assessment. These models mimic human perception and evaluate image quality accordingly.

Step 6: Correlation Analysis

- Establish a correlation between the selected clustering technique and the outcomes derived from the human-eye models. This analysis provides insights into the alignment between machine-based and human-based assessments of image quality.

Step 7: Statistical Testing

- Perform statistical testing to quantify the level of confidence in the image quality assessment results. This step aids in validating the robustness and reliability of the algorithm's findings.

c. database

This section introduces two popular databases such as Color-Feret and Kinect-Face which are used in our work elaborated each given below:

1. Color-Feret database [37]:

The Color-FERET (Face Recognition Technology) database stands as a seminal resource in the field of facial image analysis and recognition. The dataset offers a rich collection of color facial images and serves as a benchmark for assessing the performance of various facial recognition systems. It has total 11,338 images of 994 individuals. These images acquired by several poses encompassing the rotations of ±00, ±22.50, ±67.50 and ± 900 and also considering arbitrary angles over 15 sessions in the mid of year 1993 to 1996. Advantage of this dataset - it gives more feasibility for automatic face recognition.

2. Kinect-Face database [38]:

It represents a notable contribution to the field of computer vision, providing a valuable dataset for research in facial recognition and related applications. It is characterized by several salient features - 3D Facial Data, Expression Variation, Pose Variation, etc. Images are taken by 52 individuals i.e. 38 males and 14 females that are captured in 2 periods and 9 states of various face expressions or by changing light and occlusion conditions such as smile, left/right profile, occlusion of eyes/mouth, etc.

In our work, 150 images are taken of each such as Color-Feret database and Kinect-Face database and convert it into Gray-scale images.

3. In-house database:

Consists of faces from university biometric dataset. Total 380 images: 300 of Men's and 80 of Women's of size 240 x 320 px.

d. Image quality factors and metrics

Several metrics have been reported in this paper obtained for face image recognition. The most used for measurement of image quality factors [39] are contrast, brightness, focus, etc., but in our addition, include some more factors to get automatic focus of image such as entropy- if it is high that means the randomness is also high; also, occlusion plays very important role to improve system performance. However, researchers cannot rely in only three or four factors, it will decrease recognition system performance. Also, Image enhancement can be subjective means quality of image differs on person to person. Hence it is very much efficient to get quantitative metrics for determining image quality. Such quality factors mentioned below-

1. Contrast

It refers to the difference in luminance or color intensity between different elements or regions within an image. It quantifies the visual distinction between objects or features in an image, contributing to the perceptual clarity and saliency of these elements.

$$C_{rms}[40] = \sqrt{\frac{\sum_{x=1}^{M}\sum_{y=1}^{N}[I(x,y)-u]^2}{M*N}}$$

2. Brightness

It pertains to the overall intensity or luminance of an image. It represents the perceived level of lightness or darkness in the image and influences its overall illumination. Brightness [41] can be measure as:

$$B[41] = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}[\mathrm{Max}(R,G,B)]$$

3. Focus

A degree of image blurring is referred by image focus as,

$$\sum_{X=1}^{M}\sum_{y=1}^{N}|Gxx(x,y)| + |Gxx(x,y)| + |\,Gxx(x,y)|$$

4. Sharpness

Sharpness characterizes the degree of detail and fine structure present in an image. A sharp image exhibits well-defined edges, high contrast between adjacent regions and a clear representation of fine textures and patterns. It is expressed as,

$$S\,[43] = \frac{1}{2}\Big[\frac{1}{(N-1)*M}\sum_{x=1}^{M}\sum_{y=1}^{N-1}|I_{x,y}-I_{x,y+1}|\Big] + \frac{1}{(M-1)*N}\Big[\sum_{x=1}^{M-1}\sum_{y=1}^{N}|I_{x,y}-I_{x+1,y}|\Big]$$

5. Illumination

An illumination encompasses the quantity, quality and direction of light incident on the subject. It is measured as weighted sum of mean values intensity which is divided into 4x4 blocks.

$$I\,[44] = \sum_{i=1}^{4}\sum_{j=1}^{4} Wij * \ddot{I}ij$$

6. Entropy

Representing a measure of disorder, randomness or uncertainty within a system or a dataset. It is determined by histogram which shows different grey level probabilities in given image.

$$E\,[45] = -\sum^{n} P_i\ \mathrm{Log}_2\,(P_i)$$

7. Occlusion

It is a fundamental concept that pertains to the partial or complete concealment of objects within an image by other objects. Author written formulae as,

$$O = X * \frac{(M * M)}{(N * N)} * 100$$

Where, X= number of blocks whose maximum value = minimum value.
M= block of image and N = Size of image.

To check the result of above quality factors- author used the Color-Feret DB for contrast, brightness, focus, illumination, and sharpness. For Contrast, images are saturated at low and very high intensities in step of 10%; i.e. mapping of low[in]=0, high[out]=1 to low[in]=0.05, high[out]=0.95. Next gamma parameter is used to change brightness of image such as γ=0.5, 0.6…..1.4. If gamma <1 indicates bright and gamma > 1 indicates dark images. Towards verifying the degree of blurriness in image circular average filter is used by modify disk values like d=3, 5…19; Hence various values of focus and sharpness are obtained. The effect of illumination can be seen considering different sets of Yale DB; all results are mentioned in table 2. For entropy and occlusion; there are alternate options to modify it; hence, considered as it is. Take 10 images of this database and determine the values as shown in table 3.

Table 2: Contrast (C), brightness (B), focus (F), sharpness (S),and illumination (I) measures by applying above techniques.

Color-Feret DB	C	B	F	S	I
Normal	0.32	0.49	0.03	0.84	0.87
10%	0.48	0.46	0.02	0.56	0.82
20%	0.52	0.44	0.006	0.48	0.84
30%	0.56	0.39	0.003	0.37	0.76
40%	0.62	0.37	0.002	0.24	0.62
50%	0.66	0.32	0.003	0.23	0.59
60%	0.73	0.23	0.002	0.20	0.25
70%	0.78	0.25	0.0013	0.19	0.91
80%	0.85	0.22	0.0012	0.15	0.82
90%	0.92	0.15	0.0014	0.13	0.72

Source: Own elaboration.

Table 3: Entropy (E) and Occlusion (O) measures by considering random 5 images of DB.

Color-Feret DB	E	O
I1	2.4	0.95
I2	3.45	0.85
I3	2.96	0.76
I4	2.74	0.72
I5	2.45	0.56

Source: Own elaboration.

e. Fusion schemes

Machine learning methodologies [46] designed to enable computers to learn from data and improve their performance on specific tasks. Broadly, this can be categorized into three main types, each distinguished by its underlying learning paradigm and application domain such as supervised, unsupervised, and semi-supervised [47]. Whereas author's go with unsupervised learning because there are no explicit target labels provided [48]. This algorithm autonomously identifies hidden patterns or groupings within the data. It capable to perform the task like clustering, anomaly detection, dimension reduction, feature based learning, etc. Clustering and dimensionality reduction are two key subdomains of unsupervised learning. Clustering algorithms aiding in tasks in form of groups while dimensionality reduction techniques helps to reduce the complexity of high-dimensional data for visualization and analysis. Unsupervised learning is critical in applications such as anomaly detection, recommendation systems and exploratory data analysis. In this work, author prefer to choose clustering because it offers several distinct advantages - by grouping similar data points into clusters, it helps analysts gain a deeper understanding of the relationships, dependencies and natural divisions present in the dataset. It also identifies similarities and dissimilarities among data points, highlighting commonalities and anomalies. In order to determine efficient clustering technique, the author review few literatures and obtain the drawbacks. A centroid-based clustering algorithm K-Means serves as a foundational method for partitioning a dataset into distinct clusters, each characterized by a central representative point known as a centroid. There are so many variations of K-Means clustering that differ each other for obtaining centroids such as Bisect K-Means, Fuzzy C-Means [49], affinity-propagation [50], etc. The centroid-based clustering has biggest drawback that it unable to handle noise, outliers, high dimension feature space from objects. Unlike centroid based clustering has high computational cost because of its uniqueness result shows in dendrogram which helps to visualize the relationship between clusters in hierarchical format [51].

Next popular techniques are hierarchical clustering consists of agglomerative and divisive follows bottom-to-up and top-to-bottom approach; it is a valuable technique for uncovering hierarchical structures within data. However, it is not without its limitations and drawbacks. It has a higher computational complexity compared to centroid-based algorithms like K-Means. The time and memory requirements increase rapidly with the number of data points, making them less suitable for large datasets. Agglomerative hierarchical clustering, in particular, can become computationally expensive as it repeatedly merges data points or clusters.

It is important to note that DBSCAN [52] is a robust and versatile clustering algorithm that offers several advantages over traditional clustering techniques, particularly in scenarios where the data distribution is irregular or contains noise. One of its notable strengths is an ability to automatically detect the number of clusters in the data, without requiring the user to specify the number of clusters beforehand. This is in contrast to K-Means and other centroid-based methods, which rely on a predefined K value.

In order to verify clustering results, author opt inter-cluster and intra-cluster which are most frequently used in the context of clustering analysis for data analysis. They refer to the relationships and distinctions between clusters formed by grouping data points based on similarity. Understanding these terms is essential for assessing the quality and characteristics of clustering results where Intra-cluster refers to the properties, relationships or characteristics that exist within individual clusters formed by a clustering algorithm and Inter-cluster pertains to the relationships and distinctions between different clusters within a dataset. It focuses on how well-separated or distinct clusters are from each other. Both aspects are essential for evaluating the quality of a clustering solution and understanding the structure of the data. Also, outlined that if inter-cluster is more and intra-cluster is less, this means clustering algorithm is pertained to be good result found with DBSCAN as mentioned in table 4.

Table 4: Inter and Intra Class Clustering with different datasets.

Algorithm/Dataset	**Color-Feret**		**Kinect-Face**		**GLA**	
	Inter	**Intra**	**Inter**	**Intra**	**Inter**	**Intra**
K-Means	75.69	80.95	59.08	88.65	47.11	75.92
Agglomerative	89.85	93.23	75.07	94.22	56.63	91.24
DBSCAN	**91.49**	**25.02**	**72.38**	**54.48**	**67.58**	**13.43**
Birch	89.83	96.23	60.0	94.22	56.60	91.20
Gaussian Mixture	16.05	57.31	19.22	56.41	42.07	75.52
Affinity Propagation	59.78	70.57	70.52	98.12	41.88	94.63
Bisect-Means	75.69	90.95	59.08	88.65	47.11	75.92

Source: Own elaboration.

f. Correlation between system and human score

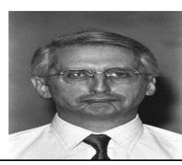
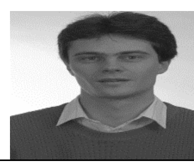
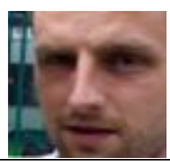
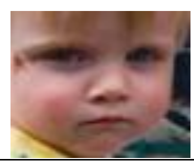
After system evaluation, there is need for Human quality assessment (HQA), it involves evaluating the quality and perceptual aspects of images through human judgment. This assessment is crucial in image processing, etc. HQA involves human evaluators rating the perceived quality of images. Typically, author has created the group of three individuals that assesses images and provides subjective scores based on their visual perception.

To obtain correlation between system score and human score, Spearman correlation coefficient (SROCC) is used. SROCC, often referred to as Spearman's rho (ρ), is a statistical measure used to assess the strength and direction of the monotonic relationship between two quality scores. It often denoted by the symbol ρ (rho) which is a degree of monotonic relationship between the two different scores. It assesses both strength and direction of association between the ranks of data scores rather than their actual values. SROCC is measured by,

$$\rho[52] = r_R = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Note that the SROCC ranges from -1 to 1. This means 1 identifies a perfect increasing monotonic relationship, -1 identifies a perfect decreasing monotonic relationship and 0 specifies the no monotonic relationship between the variables. Closer absolute value of ρ is to 1, is stronger monotonic relationship between such variables. Hence result is 0.88, 0.90 and 0.95 from which human score 3 is quite better to show relation as shown in table 5.

Table 5: Quality based rankings.

Score/Data				
System Score	0	1	3	2
Human Score-1	2	1	2	2
Human Score-2	1	1	3	2
Human Score-3	0	1	3	4

Source: Own elaboration.

Further, Z-test helps to make informed decisions about whether observed differences are statistically significant or not. This statistical hypothesis test is used to assess whether the means of two scores are significantly different when the quality standard deviation is known. Hence, computing the z-score is-

$$Z\ [28] = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1}} + \sqrt{\frac{\sigma_2^2}{n_2}}}$$

g. Experiment

During experiment, various face recognition algorithms analyse and identify face images enabling automated recognition, verification and analysis of individuals based on their facial features.

The techniques are classified like, intensity-based like principal component analysis (PCA), independent component analysis (ICA) [54]; distribution-based like LBP [55], and local ternary patterns, etc. The commercial software developed by Pittsburgh Pattern Recognition (PittPatt) for face detection gives best results of 98.6% shown in table 6.

Table 6: Various techniques for face recognition; where Rank1 represents the performance.

Algorithms	Rank %
PCA	88.9
ICA	84.6
LBP	92
LTP	91.2
PittPatt	98.6
Logistic regression [15]	75.60%
Support vector regression [15]	79.38%
Neural network[1]	81.02%

Source: Own elaboration.

The Face detection algorithm gives best result obtaining by comparing several existing techniques from which cascade classifier gives best face detection rate and PittPatt predicts good recognition mentioned in the table 7.

Table 7: Comparison of various face detection, fusion schemes to classify the input images.

Database	Description	Face Detection Rate	Fusion Rule	Correct Matching (%)	Correlation	Recognition
FRI CVL [56]	S, R, P, B	94.3%	$S\ [18] = \sum_{i=1}^{4} W_i * S_i$	92.1%	--	
Hermes project [57]		90.5% (Cascade Classifier)		87.1%	--	
FOCS	C, B, F, S, I	91.3% (LBP)	Neural Network [1]	---	--	
Color-Feret [37]	C, B, S, P, O	92.7% (DCNN)	DCNN [15]	---	0.85	
Kinect Face [38]					0.76	
Clustering	C, B, F, S, I, O, E	93.2 %	DBSCAN	---	0.88	PittPatt (98.6%)
					0.90	
					0.95	

Source: Own elaboration.

IV. CONCLUSIONES

In this study, the author explored various methodologies and metrics in Face Image Quality Assessment (FIQA), which aim to evaluate face image quality accurately. These methods have evolved to address new challenges, including machine learning and deep learning techniques. The use of unsupervised learning methods, like clustering, is crucial for understanding the relationship between quality measures and score prediction in practical applications. This research has led to the development of a more efficient Face Quality Index (FQI) that effectively reflects changes in input quality factors and their correlation with face recognition performance.

V. ACKNOWLEDGMENTS

I sincerely thank all who contributed to this research. Special gratitude to Prof. Phalguni Gupta, my PhD supervisor, and Dr. Pooja Pathak, my co-supervisor. I am also grateful to Prof. Dakshina Ranjan Kisku from NIT Durgapur for their invaluable insights. This work benefited greatly from the support of many individuals and organizations.

VI. STATEMENT AND DECLARATIONS

Funding: Authors declare that no funds, grants, or other support were received during the preparation of manuscript.

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

Author Contributions: Conceptualization, first and second authors'; Formal analysis, first and third authors'; Expert System and Software, first author; Validation, all three authors'. All authors' read and approved the final manuscript.

VII. REFERENCES

[1] D. Agarwal and A. Bansal, "Assessment of latent fingerprint image quality based on level 1, level 2, and texture features," in International Conference on Innovative Computing and Communications: Proceedings of ICICC 2020, Volume 1, Singapore, 2021.
[2] A. Abaza et al., "Design and evaluation of photometric image quality measures for effective face recognition," IET Biometrics, 3(4), pp. 314-324, 2014.
[3] R. L. Hsu et al., "Quality assessment of facial images," in Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference, IEEE, pp.1-6, 2006.
[4] D. Bhattacharjee et al., "No-reference image quality assessment for facial images," in Advanced Intelligent Computing Theories and Applications, with Aspects of Artificial Intelligence: 7th International Conference, ICIC 2011, Zhengzhou, China, August 11-14, pp. 594-601, 2011.
[5] Y. Wong et al., "Patch-based probabilistic image quality assessment for face selection and improved video-based face recognition," in CVPR 2011 WORKSHOPS, IEEE, pp. 74-81, 2011.

[6] J. Zuo and N.A. Schmid, “Adaptive quality-based performance prediction and boosting for iris authentication: methodology and its illustration,” IEEE transactions on information forensics and security, 8(6), pp. 1051-1060, 2013.

[7] J. Merkle et al., “Towards improving the NIST fingerprint image quality (NFIQ) algorithm,” 2010.

[8] J. Sang et al., “Face image quality evaluation for ISO/IEC standards 19794-5 and 29794-5,” in Advances in Biometrics: Third International Conference, ICB 2009, Alghero, Italy, Proceedings 3, pp. 229-238, June 2-5, 2009.

[9] H.R. Sheikh et al., “An information fidelity criterion for image quality assessment using natural scene statistics,” IEEE Transactions on image processing, 14(12), pp. 2117-2128, 2005.

[10] Z. Wang et al., “Image quality assessment: from error visibility to structural similarity,” IEEE transactions on image processing, 13(4), pp. 600-612, 2004.

[11] L. Zhang et al., “Fsim: A feature similarity index for image quality assessment,” IEEE Trans. Image Process., 20 (8), pp. 2378–2386, 2011.

[12] K. Gu et al., “Analysis of distortion distribution for pooling in image quality prediction,” IEEE Trans. Broadcasting, 62 (2), pp. 446–456, 2016.

[13] X. Lv and Z.J. Wang, “Reduced-reference image quality assessment based on perceptual image hashing reduced-reference image quality assessment based on perceptual image hashing,” in: Proceedings of the 16th IEEE International Conference on Image Processing (ICIP), pp. 4361–4364, 2009.

[14] Q. Li and Z. Wang, “Reduced-reference image quality assessment using divisive normalization- based image representation,” IEEE J. Sel. Topics Signal Process. 3 (2), pp. 202–211, 2009.

[15] M. Saad et al., “Blind image quality assessment: a natural scene statistics approach in the DCT domain,” IEEE Trans. Image Process. 21 (8), pp. 3339–3352, 2012.

[16] A. K. Moorthy et al., “Blind image quality assessment: from natural scene statistics to perceptual quality,” IEEE Trans. Image Process. 20 (12), pp. 3350–3364, 2011.

[17] P. Ye et al., “Unsupervised feature learning framework for no- reference image quality assessment,” in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1098–1105, 2012.

[18] Mittal et al., “No-reference image quality assessment in the spatial domain,” IEEE Trans. Image Process. 21 (12), pp. 4695–4708, 2012.

[19] L. Kang and P. Ye, “Convolutional neutral networks for no-reference image quality assessment,” in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1733–1740, 2014.

[20] L. V. Hsu et al., “Quality assessment of facial images,” in Proc. Biometrics Symp., Special Issue Res. Biometric Consortium Conf., pp. 1–6, Sep./Aug. 2006.

[21] Nasrollahi et al., “Face quality assessment system in video sequences,” in: Proceedings of the Biometrics and Identity Management, First European Work-shop (BIOID), pp. 10–18, 2008.

[22] Y. Wong et al., “Patch-based probabilistic image quality assessment for face selection and improved video-based face recognition,” Proc. SPIE - Int. Soc. Opt. Eng. 7255: pp. 74–81, 2011.

[23] G. Aggarwal et al., “Predicting performance of face recognition systems: An image characterization approach,” in Proc. IEEE CVPRW, pp. 52–59, Jun. 2011.

[24] P. J. Phillips et al., “On the existence of face quality measures,” in Proc. IEEE Conf. Biometrics, Theory, Appl. Syst. (BTAS), pp. 1–8, Sep./Oct. 2013.

[25] R. Veldhuis Dutta and L. Spreeuwers, “A Bayesian model for predicting face recognition performance using image quality,” in Proc. IEEE Int. Joint Conf. Biometrics (IJCB), pp. 1–8, Sep./Oct. 2014.

[26] Vignesh, S. et al., “Face image quality assessment for face selection in surveillance video using convolutional neural networks,” in IEEE Global Conference on Signal and Information Processing (Global SIP), pp. 577-581, 2015.

[27] P. Choudhary and P. Pathak, “A review of convolution neural network used in various applications,” in 5th international conference on information systems and computer networks (ISCON), IEEE, pp. 1-5, 2021.

[28] H.-I. Kim et al., “Face image assessment learned with objective and relative face image qualities for improved face recognition,” in Proc. IEEE Int. Conf. Image Process. (ICIP), pp. 4027–4031, Sep. 2015.

[29] N. Zhuang et al., “Recognition oriented facial image quality assessment via deep convolutional neural network,” Neurocomputing, 358, pp. 109-118, 2019.

[30] S. Bharadwaj et al., “Can holistic representations be used for face biometric quality assessment?” in Proc. IEEE Int. Conf. Image Process. (ICIP), pp. 2792–2796, Sep. 2013.

[31] C. Chen Dantcheva and A. Ross, “Can facial cosmetics affect the matching accuracy of face recognition systems?” in Proc. IEEE Int. Conf. Biometrics Theory Appl. Syst. (BTAS), pp. 391–398, 2012.

[32] A. Dantcheva Chen and A. Ross, “Impact of facial cosmetics on automatic gender and age estimation algorithms,” in Proc. Int. Conf. Comput. Vis. Theory Appl. (VISAPP), vol. 2, pp. 182–190, 2014.

[33] Rathgeb et al., “Impact and detection of facial beautification in face recognition: An overview,” IEEE Access, vol.7, pp. 152667–152678, 2019.

[34] M. Ibsen et al., “Impact of facial tattoos and paintings on face recognition systems,” IET Biometrics, vol. 10, pp. 706–719, Apr. 2021.

[35] Masks4All., “What countries require masks in public or recommend masks?” 2020.

[36] S. Su et al., “Going the extra mile in face image quality assessment: A novel database and model,” arXiv preprint arXiv:2207.04904, 2022.

[37] D. Cho Jo et al., “Ifqa: Interpretable face quality assessment,” in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 3444–3453, 2023.

[38] P. Phillips et al., “The feret evaluation methodology for face algorithms,” IEEE Trans. Pattern Anal. Mach. Intell. 22, pp. 1090–1104, 2000.

[39] R. Min et al., “Kinectfacedb: a Kinect database for face recognition,” IEEE Trans. Syst. Man Cybern.: Syst. 44 (11), pp. 1534–1548, 2014.

[40] J. Sang et al., “Face image quality evaluation for ISO/IEC standards 19794–5 and 29794–5,” in Int. Conf. on Biometrics (ICB), Sassari, Italy, 2009.

[41] X. Gao et al., “Standardization face image sample quality,” in Int. Conf. on Biometrics (ICB), Seoul, Korea, 2007.

[42] A. Mishra et al., “A robust approach for palmprint biometric recognition,” International Journal of Biometrics, 11(4), pp. 389-408, 2019.

[43] S. Bezryadin et al., “Brightness calculation in digital image processing,” in International symposium on technologies for digital photo fulfillment, Vol. 1, pp. 10-15, 2007.

[44] P. T. Yap and P. Raveendran., “Image focus measure based on Chebyshev moments,” IEEE Proceedings-Vision, Image, and Signal Processing, 151(2), pp. 128-136, 2004.
[45] K. Kryszczuk and A. Drygajlo, “On combining evidence for reliability estimation in face verification,” in 14th European Signal Processing Conference, IEEE, pp. 1-5, 2006.
[46] M. Abdel-Mottaleb and M.H. Mahoor., “Application notes-algorithms for assessing the quality of facial images,” Computational Intelligence Magazine, IEEE, 2(2), pp. 10-17, 2007.
[47] X. Chen et al., “No-reference color image quality assessment: From entropy to perceptual quality,” EURASIP Journal on Image and Video Processing, pp. 1-14, 2019.
[48] P. Pathak and P. Choudhary, “A Comprehensive Review of Various Machine Learning Techniques,” Explainable Machine Learning Models and Architectures, pp. 1-10, 2023.
[49] L. Cheng and T. Yu., “A new generation of AI: A review and perspective on machine learning technologies applied to smart energy and electric power systems,” International Journal of Energy Research, 43(6), pp. 1928-1973, 2019.
[50] J. Yang et al., “Information extraction from electronic medical records using multitask recurrent neural network with contextual word embedding,” Applied Sciences, 9(18), 3658, 2019.
[51] J. C. Dunn, “A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters,” 1973.
[52] B. J. Frey and D. Dueck, “Clustering by passing messages between data points,” Science, vol. 315(5814), pp. 972-976, 2007.
[53] K. Nezamabadi et al., “Unsupervised ECG analysis: A review,” Reviews in Biomedical Engineering, IEEE, 16, pp. 208-224, 2022.
[54] M. Hahsler et al., “dbscan: Fast density-based clustering with R,” Journal of Statistical Software, 91, pp. 1-30, 2019.
[55] P.C. Yuen and J. H. Lai, “Face representation using independent component analysis,” Pattern recognition, 35(6), pp. 1247-1257, 2002.
[56] Y. Zhao, “Theories and applications of LBP: a survey,” in Advanced Intelligent Computing Theories and Applications,” With Aspects of Artificial Intelligence: 7th International Conference, ICIC 2011, Zhengzhou, China, Revised Selected Papers 7, pp. 112-120, August 11-14, 2011.
[57] F. Solina et al., Color-based face detection in the "15 seconds of fame" art installation, 2003.