Journal of Advanced Zoology
ISSN: 0253-7214
Volume 45 Issue -3 Year 2024 Page 465-482

Enhancing Face Recognition Accuracy On Low-Resolution Databases Using Interpolation Techniques And Feature Extraction Techniques

Bhavna Bhadkare[1*], Dr. Varsha Jotwani[2]

[1]Ph.D Scholar, Department of Computer Science, RNTU Bhopal*
[2]Professor and HOD, Department of Computer Science & IT, RNTU Bhopal

***[*]Corresponding Author:** Bhavna Bhadkare*
[*]Ph.D Scholar, Department of Computer Science, RNTU Bhopal

Abstract –

The images in high resolution contain more useful information than the images in low resolution. Thus, high-resolution digital images are preferred over low-resolution images. Image super-resolution is one of the principal techniques for generating high-resolution images. This research investigates the impact of image resolution on the performance of face recognition systems and proposes methods to enhance recognition accuracy on low-resolution face databases. In the first phase, several holistic face recognition algorithms, including Support Vector Machine (SVM), Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), and RESNET50, are evaluated for their performance on low-resolution face images. Subsequently, three interpolation techniques - nearest neighbor, bilinear, and bicubic interpolant - are applied as preprocessing steps to increase the resolution of the input images. The study aims to determine the effectiveness of these techniques in improving recognition accuracy. Various evaluation metrics, including accuracy, precision, sensitivity, specificity, Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), and Structural Similarity Index (SSIM), are employed to assess the performance of the recognition systems. The results demonstrate the efficacy of the proposed approach in enhancing recognition accuracy on low-resolution face datasets, thereby contributing to the advancement of face recognition technology in practical applications.

Keyword -face recognition Low-resolution images Interpolation techniques Feature extraction

I INTRODUCTION

The Single Image Super-Resolution (SISR), a technique for restoring a visually pleasing high-resolution (HR) image from its low-resolution (LR) version, is still a challenging task within the computer vision research community [1]. Since multiple solutions exist for the mapping from LR to HR space, SISR is highly ill-posed and a variety of algorithms, especially the current leading learning-based methods are proposed to address this problem. Understanding what the SISR problem represents is crucial in order to develop a method that is capable of solving it. Having a low-resolution image at inference time means that there is no ground truth answer on how the high-resolution counterpart image is generated. That being said, in order to recover a higher-resolution image, assumptions need to be made that do not violate the visible artifacts taken from the low-resolution image.[2-3] The fine details added to the higher-resolution image are subjective, since they only need to follow certain already visible artifacts from the low-resolution image. The task in SISR is to find a model that learns how to make these assumptions and generate high-resolution images as plausible as possible according to the specific task that is being undertaken like, Face SISR. To this day, all current solutions for the SISR problem attempt to reconstruct a single high-resolution image based on a given low-resolution input image. In other words, the process of generating a high-resolution image is deterministic and given the same low-resolution image multiple times as input will yield the same high-resolution image.

Digital images captured by poor resolution cameras have three primary limitations, namely aliasing, blurring, and noise. Aliasing can occur due to inadequate image sensor elements that lead to an under-sampled spatial resolution, which results in a significant loss of high-frequency (HF) information, such as edges and textures. Image blur occurs due to camera motion, jitter, out-of-focus, etc. In addition to blur, various noises can also be added to the image during the imaging process and can degrade the image quality. Degradation may also occur because of the sensor element's point spread function (PSF). Super-resolution (SR) refers to those techniques designed to build high-resolution (HR) images from single or more observed low-resolution (LR) images by increasing the HF components, replicating larger dimensional multipliers, and removing the degradation caused by the low-resolution camera imaging process. In essence, the super-resolution process should reconstruct lost HF details while minimizing aliasing and blurring.

As stated before, HR images are obtained by increasing the number of image sensor elements and reducing the pixel size. This increases the pixel density. However, a reduction in pixel size causes shot noise and degrades the quality of the image captured. In addition, it may result in additional costs due to an increase in the number of sensor elements. Therefore, the employment of novel signal processing approaches is required to post-process the captured LR images. A simple approach is by interpolating the LR image to the size of the desired HR image. However, traditional interpolation approaches, such as bilinear, bi-cubic, and nearest neighbor algorithms, result in blurry images, as the missing pixel is found by averaging it from a neighboring pixel. The blurry effect introduced by interpolation techniques contributes to the loss of HF details, and hence, the fundamental problem in SR reconstruction, i.e., aliasing effect (loss of HF details), remains unsolved. Typically, an image that holds fine details is said to be an HR image and that with fewer details is referred to as an LR image. Image resolution provides the least measure of detail with which an image can be resolved into a more intricate and clearly defined pixel. As the resolution of an image is increased, it conveys a more complex structure. Therefore, image resolution is vital in all wings of digital image processing, and the performance of an image-processing algorithm depends on image resolution. It is therefore one of the key aspects of digital image processing.[4] The resolution of an image depends primarily on the sensor elements used in the imaging device. For obtaining an HR image, a sophisticated, complex sensor is therefore needed. This can be very expensive and, in many cases, not affordable. Resolution is an important term for the quality assessment of image acquisition and processing devices in digital image processing. Image resolution is defined as the smallest measurable visual data in an image. The resolution of an optical device can be quantified by measuring its OTF, which is a measure of the system response to different spatial frequencies. Digital image processing can generally classify the image resolution into the following four types:[5]

Pixel or Spatial Resolution: An image consists of several distinguishable pixel image elements. The spatial distance between pixels in an image is called pixel or spatial resolution. The first number is the number of pixel columns (width), while the second is the number of pixel lines (high), named m by n. It is represented by a set of two positive integers. High spatial resolution improves the image quality by allowing a clear insight into fine details and vivid color transitions. Instead, an image with fine details not shown with enough pixels suffers from aliasing and introduces undesired artifacts, such as the blocking effect.

Intensity Resolution: The number of grey levels used to represent an individual pixel is referred to as intensity resolution. It is represented by the number of bits used to represent each intensity level. A small, discernible change in grey level can be perceived with a large number of bits used to represent a pixel. However, increasing the number of bits increases the image size. A monochromatic image's typical intensity resolution is 256 grey levels, implying 8 bits are required to represent a pixel.

Temporal Resolution: Temporal resolution refers to the frame rate captured by a camera in a motion picture. It carries the motion information between two subsequent frames. Movements in a scene can be viewed without smearing using a higher frame rate. A typical frame rate to view motion pictures is above 25 frames per second.

Spectral Resolution: The ability to resolve an image into its respective frequency or spectral components is known as a spectral resolution. However, spectral resolution is not discernible to human eyes as much as spatial resolution. Hence, the spectral analysis generally allows a higher tolerance range since small changes in the spectral resolution often go undetected.[6-9]

II RELATED WORK

Low resolution of light field images is the process of reconstructing a high-resolution light field image from a given low-resolution light field image. This section will mainly introduce the traditional super-resolution methods of light field images.

Daniel Schulz et al. (2022)[10] proposed a novel No-Reference method for assessing the quality of ID card images by combining Face Image Quality Assessment and Text Quality Assessment. Their approach leveraged a private dataset of 12,960 Chilean ID cards to evaluate the performance of the proposed method, which showed simultaneous improvements in face and text verification metrics as low-quality images were discarded.

Ang Li Jian Hu et al. (2022)[11] introduced AFSNet, a novel Attribute-Conditioned Face Swapping Network, designed to generate high-quality face-swapped images from low-resolution inputs. They utilized an Image Enhancement Network (IEN) and a Face Exchange Module (FEM) to enhance image quality and preserve attributes, demonstrating superior performance compared to existing methods.

Qiye Lian et al. (2022)[12] presented VLC-FIQA, an unsupervised Face Image Quality Assessment method that quantifies the importance of pixels in a face image and computes the variation of importance as a measure of image quality. Their approach outperformed state-of-the-art methods on LFW, offering a promising solution for face recognition systems.

Sebastián González et al. (2022)[13] addressed the challenge of identity verification using ID cards in remote systems by employing MagFace, a quality-aware face recognition method, in conjunction with a Chilean ID card PAD system. Their approach analyzed the influence of photo ID quality on system effectiveness, enhancing the overall security of verification systems.

Weisong Zhao et al. (2022)[14] tackled the performance degradation of facial recognition systems caused by mask-wearing during the COVID-19 pandemic. Their proposed method utilized a consistent sub-decision network and knowledge distillation to improve masked face recognition performance, outperforming baseline methods on public datasets.

Biying Fu et al. (2022)[15] introduced explainability tools for unsupervised face image quality assessment methods, enabling the derivation of reasoning for different quality decisions and their implications on face recognition performance. Their tools provided insights into the behavior of face recognition models across various quality levels.

Qiyu Wei et al. (2022)[16] proposed an approach for high-quality face image generation with predefined attributes, leveraging StyleGAN and attribute classifiers to control attribute synthesis. Their method demonstrated effectiveness in generating hyper-realistic face images with desired attributes.

Peng Zheng et al. (2022)[17] presented MDFR, a Multi-Degradation Face Restoration model, designed to restore high-quality faces from low-quality inputs under various challenging conditions. Their approach outperformed state-of-the-art methods in both face formalization and restoration tasks.

Ying Tai Feida Zhu et al. (2022)[18] proposed SGPN, a Shape and Generative Prior integrated Network, for blind face restoration from low-quality inputs. By integrating shape and generative priors and employing hierarchical spatial features, their method achieved superior performance compared to existing blind face restoration methods.

M. Benedict Tephila et al. (2022)[19] recommended a Bi-interval contrast enhancement and color correction method for improving the quality of underwater images. Their approach involved subinterval linear transformation, Gaussian low-pass filtering, and Bi-interval histogram equalization to enhance image quality effectively, yielding high-quality underwater images compared to traditional methods.

Wang, Z et.al.(2023)[24] In the realm of remote sensing, practical applications of single-image super-resolution (SISR) methods based on deep convolutional neural networks (CNNs) are hindered by memory consumption and computational burden. To overcome this challenge, we introduce a lightweight feature enhancement network (FeNet) tailored for accurate remote-sensing image super-resolution (SR). Recognizing the constraints of hardware facilities, we devise a lighter FeNet-baseline with approximately 158K parameters. Our approach draws inspiration from lattice structures, leading to the creation of a lightweight lattice block (LLB) as a nonlinear feature extraction function aimed at enhancing expression ability. Leveraging channel separation operations, the upper and lower branches of the LLB focus on distinct sets of features, while weight coefficients calculated via attention mechanisms facilitate efficient communication between these branches. Building upon LLB, we design a feature enhancement block (FEB) in a nested manner to obtain expressive features, with different layers responsible for features with varying texture richness, followed by sequential fusion of features from different layers from deep to shallow. Model parameters and multi-adds operations serve as metrics to assess network complexity, and extensive experiments conducted on two remote-sensing and four SR benchmark test datasets demonstrate that our methods achieve a favorable balance between complexity and performance.

Lijing Bu et.al.(2023)[25] Hyperspectral images (HSI) are renowned for their rich spectral information, but often suffer from insufficient spatial resolution due to limitations in satellite sensors. To address this, a novel hybrid convolution and spectral symmetry preservation network has been proposed. This network leverages the correlation between spectral bands to enhance spatial resolution without compromising spectral fidelity. By integrating information from neighboring spectral bands and employing spatial-spectral symmetric 3D convolution, low-resolution and neighboring band features are effectively extracted. Furthermore, deformable convolution and attention mechanisms aid in information extraction from low-resolution bands. Finally, multiple bands are fused in the reconstruction module, and high-resolution hyperspectral images are obtained through Fourier transform upsampling. Experimental results on various datasets demonstrate superior performance over existing algorithms, achieving high PSNR values while preserving the spectral characteristics of hyperspectral images.

Bhavna Bhadkare, Dr. Varsha Jotwani (2024) [26] provides an overview of deep learning-based methods for enhancing low-resolution images. We explore the evolution of deep learning techniques in this context, from early approaches like CNN to state-of-the-art architectures. Deep learning models, particularly convolutional neural networks (CNNs), have demonstrated remarkable success in up scaling low-resolution images, effectively capturing intricate patterns and textures. We discuss the underlying principles of these DL-based approaches, highlighting their ability to leverage contextual information and learn complex image representations. Additionally, we delve into the various loss functions and training strategies used to optimize network performance. Applications of deep learning-based super-resolution extend to diverse domains, such as improving the quality of medical scans, enhancing surveillance footage for forensic analysis, and refining satellite imagery for better environmental monitoring. We showcase the potential impact of DL-based super-resolution in these real-world scenarios.

The review paper on low-quality images offers a comprehensive and insightful exploration of the challenges, advancements, and applications in the realm of image enhancement using deep learning. Through meticulous research, the authors have delved into the diverse facets of this field, illuminating the critical role played by deep learning architectures such as Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) in elevating the quality of low-resolution images. With a keen focus on inclusion and exclusion criteria, the paper synthesizes a wealth of literature, providing a holistic view of the subject. It skillfully navigates through data analysis, presenting quantitative and qualitative assessments of image improvements while illuminating the study's purpose in revolutionizing image perception across industries. Overall, this review paper serves as an invaluable resource, shedding light on the transformative potential of deep learning in addressing the challenges posed by low-quality images in the digital age.

III MATERIAL AND METHOD

Database To evaluate our system, ORL database (also known as AT&T)is used. It contains a set of face images taken between April 1992 and April 1994.This database, as shown in figure 1 and table 5.1, consists of ten different images of each of forty distinct subjects. Images for some subjects were taken at different times with variation in the lighting, pose, and face expression (open / closed eyes, smiling / not smiling, gloomy, normal). All images are upright frontal position and were taken against a dark background. The images are in JPG format, and the size of each image is 92×112 pixels with 256 grey levels for representing the pixel value.

Figure 1: Samples from ORL database.

Image Interpolation Techniques

Image interpolation can be expressed as the calculation of unknown pixels with the help of the known pixels. When the $M \times N$-sized I_L (low resolution) image is interpolated to the 2 × 2 size as in **Figure 2** $2M \times 2N$-sized I_H (high resolution) image will be obtained. While creating the I_H image, the image in **Figure 2**b is first created by the process shown in Equation (1).

$(2i-1,2j-1)=IL(i,j)$

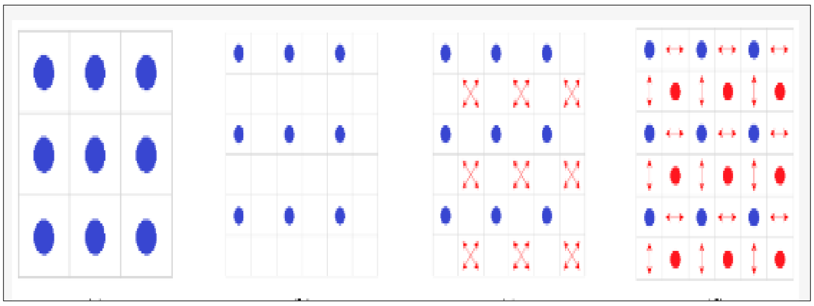


Figure 2. Example of edge-based interpolation. (a) 3 × 3 low-resolution image IL; (b) 6 × 6 high-resolution image IH with known pixels of IL; (c) pixels with diagonal neighborhoods in IH; (d) pixels with linear neighborhoods in IH.

In general, in edge-based techniques [**12**,**14**], after detecting the edges and their directions, the values of the pixels located in the $I_H(2i, 2j)$ position indicated by the diagonal arrows in **Figure 2**c are calculated first. This calculation generally depends on whether the $I_L(i, j)$ pixel is an edge pixel with a diagonal angle. The value of the $I_H(2i, 2j)$ pixel (if $I_L(i, j)$ is an edge pixel with a diagonal angle) is calculated with the help of the known blue-colored pixels adjacent to it. After calculating the diagonally angled pixels, in the second step, calculations are performed for the pixels indicated by the horizontal and vertical arrows in **Figure 2**d. If a pixel is in the position $I_H(2i, 2j + 1)$ and $I_L(i, j)$ has a horizontally oriented edge, this pixel is assigned a value using its neighbors on the horizontal plane. Similarly, if a pixel is at the position $I_H(2i + 1, 2j)$ and $I_L(i, j)$ has a vertically oriented edge, the value of the pixel $I_H(2i + 1, 2j)$ is calculated using its neighbors on the vertical plane. Finally, pixels that are not marked as edges are generally assigned using bicubic interpolation. Thus, the values of all pixels in the I_H image are determined.

In almost all interpolation studies examined, it is seen that the interpolation process has been carried out to increase the image to a 2 × 2 size. In these studies, experimental results have been obtained by comparing I_{org} (original image) to the I_H image constructed by first down sampling the I_{org} by ½ × ½ and then up sampling it to a 2 × 2 size again. To compare the success of the proposed method in this study, the images used in recent studies are first made 1/2 × 1/2 by down sampling. Although the nearest, bilinear, and bicubic techniques can be used for down sampling, in this study, the direct extraction method is used because it is known to both preserve the original pixels of the image and increase success [**14**]. When shrinking an image by direct extraction, double-index rows and double-index columns in the image are deleted. Thus, the $M \times N$-sized I_L image is obtained directly from the $2M \times 2N$-sized I_{org} image. Then, the methods to be tested are applied to the I_L image, and an I_H image of a $2M \times 2N$ size is obtained. Finally, the success of the tested method is measured by comparing the I_{org} and I_H images.[20]

A-Nearest neighbor interpolation

Nearest neighbor interpolation is a simple technique used to resize images. When scaling down an image, each pixel in the original image contributes to exactly one pixel in the resized image. The value of the pixel in the resized image is determined by selecting the value of the nearest pixel in the original image. Mathematically, let Iin denote the original image with dimensions M×N, and let I_{out} denote the resized image with dimensions P×Q, where P<M and Q<N. The nearest neighbor interpolation process for determining the value of a pixel $I_{out}(i,j)$ in the resized image is given by:

$I_{out}(i,j)=I_{in}(round(i.r),round(j.s))$

where:

$r=M/P$ is the scaling factor along the rows,

$s=N/Q$ is the scaling factor along the columns,

Round(x) rounds the value of x to the nearest integer.In this expression, i and j are the row and column indices of the pixel in the resized image, respectively. The expression calculates the closest integer indices in the original image to determine the value of the pixel in the resized image.

B-Bilinear interpolation

Bilinear interpolation is a common method used to interpolate values in a two-dimensional grid. It's often employed in image processing to resize or rescale images while maintaining smoothness and minimizing artifacts. Bilinear interpolation computes new pixel values by considering the weighted average of the four nearest neighboring pixels. Let's denote the four nearest pixels surrounding the desired point ((x,y) in the original image as Q_{11}, Q_{12}, Q_{21}, and Q_{22}, where the subscripts denote their relative positions. The bilinearly interpolated value $I(x,y)$ at point (x,y) can be expressed as:

$$\mathbf{I(x,y)=(1-\alpha)(1-\beta)Q11+\alpha(1-\beta)Q21+(1-\alpha)\beta Q12+\alpha\beta Q22}$$

Where:

α and β are the fractional parts of the coordinates x and y respectively, indicating how far away x and y are from the integer pixel coordinates.$\leq 10 \leq \alpha,\beta \leq 1$

The weights $((1-\alpha)(1-\beta)$, $\alpha(1-\beta)$, $((1-\alpha)\beta$, and $\alpha\beta$ correspond to the contributions of each neighboring pixel to the interpolated value, with higher weights assigned to closer pixels.

C-Bicubic interpolation

Bicubic interpolation is a more sophisticated method compared to bilinear interpolation, commonly used for image resizing and scaling. It offers smoother results and better preserves fine details and sharp edges.Mathematically, bicubic interpolation computes the interpolated pixel value based on a weighted average of sixteen neighboring pixels in a 4x4 grid. Let's denote these sixteen neighboring pixels as Qij, where i and j vary from -1 to 2. The interpolated value $I(x,y)$ at a point (x,y) can be expressed as:

$$\mathrm{I(x,y)}=\sum\sum_{I=-1}^{2}\sum_{J=-1}^{2} h(i-\alpha)\cdot h(\beta-j).Qij$$

Where:

α and β are the fractional parts of the coordinates x and y respectively, indicating how far away x and y are from the integer pixel coordinates.

$h(t)$ is the cubic interpolation kernel function, which typically has a bell-shaped curve to weight nearby pixels more heavily and distant pixels less. One commonly used kernel is the cubic B-spline function.[21]

Feature Extraction : Feature extraction is a process used in computer vision and image processing to identify and extract important information or features from an image or a set of images. The extracted features can then be used for various applications such as image classification, object detection, and image segmentation.

SIFT-SIFT stands for Scale-Invariant Feature Transform, which is a computer vision algorithm used for feature detection and extraction. The SIFT algorithm is designed to identify and describe local features in images that are invariant to scale, rotation, and affine distortion.

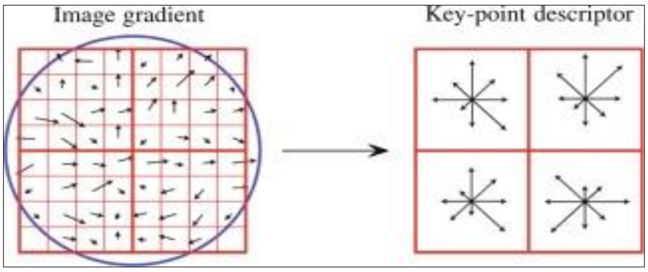


Figure.3 Scale-Invariant Feature Transform,

The SIFT algorithm works by first detecting keypoints in an image using a Difference of Gaussian (DoG) scale space. Then, for each keypoint, a local orientation is calculated based on the gradient directions in the surrounding area. This orientation is used to create a descriptor that captures the appearance of the local region around the keypoint. The descriptor is formed by taking the gradient magnitude and orientation in a number of sub-regions, and then representing this information as a histogram of gradient orientations.

SURF-SURF stands for Speeded up Robust Features, which is a computer vision algorithm used for feature detection and extraction, similar to SIFT. The SURF algorithm is designed to be computationally efficient while maintaining robustness to scale, rotation, and affine distortion. The SURF algorithm works by first detecting interest points in an image using a scale-space Laplacian of Gaussian (LoG) filter. Then, a set of Haar wavelet responses are calculated in the area around each interest point, which are used to create a descriptor that captures the local features. The descriptor is formed by taking the wavelet responses in a number of sub-regions and representing this information as a histogram of oriented gradients.[22]

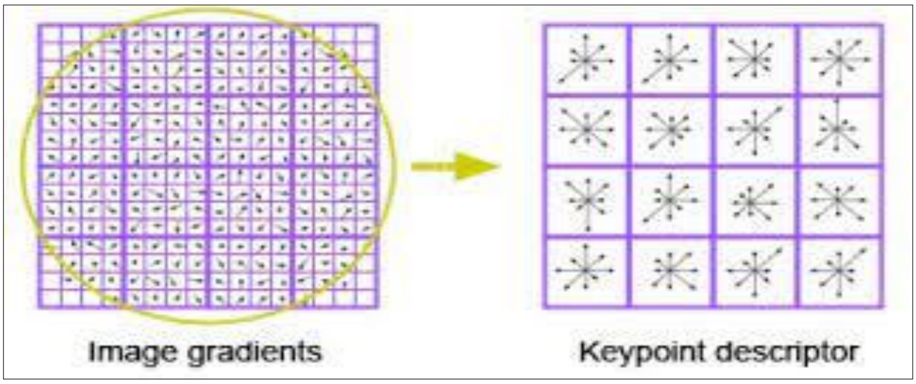


Fig. 4 Speeded up Robust Features

LBP -LBP stands for Local Binary Patterns, which is a computer vision algorithm used for texture classification and feature extraction. The LBP algorithm works by analyzing the local patterns of pixel intensities in an image. It does this by comparing the value of a central pixel in a circular neighborhood with the values of the surrounding pixels. If a neighbor pixel has a higher intensity value than the central pixel, it is assigned a value of 1, and if it has a lower intensity value, it is assigned a value of 0. This process is repeated for all pixels in the neighborhood, creating a binary code that represents the local texture pattern around the central pixel.[23]

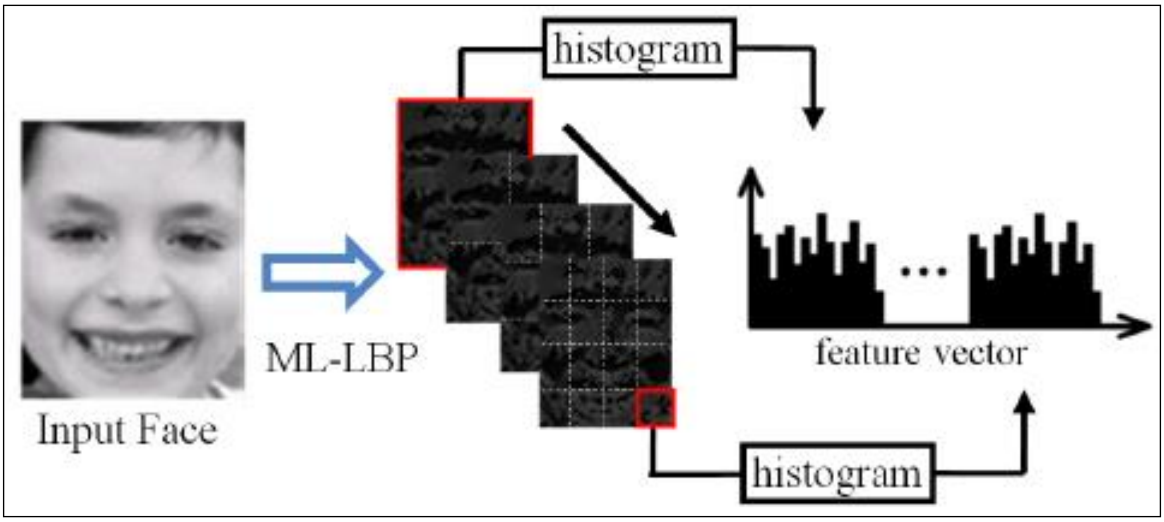


Fig.5 Local Binary Patterns

Local Binary Patterns

The binary codes for all the pixels in an image are then combined into a histogram of frequency counts, which represents the distribution of different texture patterns in the image. This histogram can be used as a feature vector for texture classification and recognition.LBP is a simple and computationally efficient algorithm that is widely used in applications such as face recognition, texture analysis, and object recognition. It has several advantages, including its ability to capture local texture information, its robustness to illumination changes, and its computational efficiency. However, LBP may not be effective for capturing fine-grained details in textures or for distinguishing between highly similar textures.

Block Based Discrete Cosine Transform (BBDCT) -Block Based Discrete Cosine Transform (BBDCT) is a signal processing technique used for data compression in image and video processing. It is based on the Discrete Cosine Transform (DCT), which is a widely used method for signal processing and compression. The BBDCT algorithm divides an image into small blocks and applies the DCT to each block. This transforms the pixel values from the spatial domain to the frequency domain, where they can be compressed more efficiently. The DCT coefficients for each block are then quantized, which means that they are rounded to a limited set of values. The quantized coefficients are then encoded using a lossless or lossy compression algorithm, depending on the application.[24]

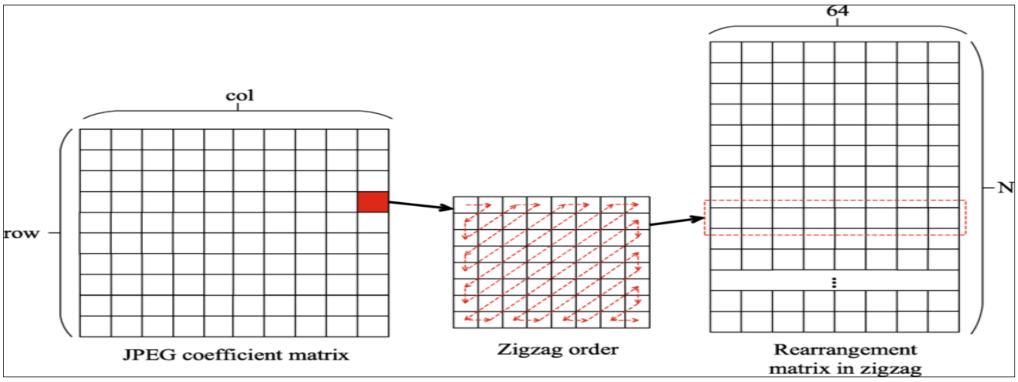


Fig.6 Block Based Discrete Cosine Transform (BBDCT)

CLASSIFICATION TECHNIQUES

Classification techniques are algorithms or methods used to categorize input data into different classes or categories.

Support Vector Machine (SVM): SVM is a supervised machine learning algorithm used for classification and regression tasks. a training dataset with *m* samples and *n* features, SVM aims to find the hyperplane that separates the data into different classes with the maximum margin. The decision function of SVM can be represented as:

$f(x)=\text{sign}(\mathbf{w}\cdot\mathbf{x}+b)$

Where **w** is the weight vector, **x** is the input feature vector, and *b* is the bias term. SVM optimizes the margin by solving the following optimization problem:

$$\min_{\mathbf{w},b} \frac{1}{2}\|\mathbf{w}\|^2$$

Subject to $(w\cdot xi+b)\geq 1$ for all training samples (xi,yi).

Convolutional Neural Network (CNN): CNN is a deep learning model widely used for image recognition and classification tasks. a CNN consists of multiple layers including convolutional layers, pooling layers, and fully connected layers. The convolutional layer applies convolution operations to the input image using learnable filters to extract features. Let *I* represent the input image, *K* denote the convolution kernel, and *B* represent the bias term. Then, the output feature map *O* of a convolutional layer can be computed as:

$O=\text{ReLU}(I*K+B)$

The pooling layer down samples the feature maps to reduce dimensionality and computational complexity. Finally, fully connected layers aggregate the features for classification.

Artificial Neural Network (ANN):ANN is a basic form of deep learning model consisting of interconnected artificial neurons organized in layers. , let **x** denote the input vector, represent the weight matrix, **b** represent the bias vector, and *σ* represent the activation function. Then, the output **y** of an ANN layer can be calculated as:

$\mathbf{y}=\sigma(\mathbf{W}\cdot\mathbf{x}+\mathbf{b})$

Common activation functions include sigmoid, tanh, and ReLU. ANN learns from data by adjusting the weights and biases during training using techniques like gradient descent and Backpropagation.

RESNET50 (Residual Network): RESNET50 is a deep neural network architecture known for its effectiveness in image classification tasks. RESNET50 introduces residual connections to address the vanishing gradient problem in deep networks. Let **x** denote the input to a layer, and *F*(**x**) denote the output of the layer before applying the non-linear activation function. The residual block computes the output **y** as follows:

$\mathbf{y}=F(\mathbf{x})+\mathbf{x}$

This formulation enables easier training of very deep networks by allowing the gradient to flow more easily during Backpropagation. RESNET50 consists of multiple residual blocks and achieves state-of-the-art performance in various image classification tasks.

IV PROPOSED SYSTEM

In this research, the impact of image resolution on the performance of face recognition systems is investigated. The study employs several holistic face recognition algorithms and evaluates their performance on low-resolution face images. Specifically, Support Vector Machine (SVM), Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), and RESNET50 are utilized for classification tasks. The recognition rate of these systems is analyzed concerning the image resolution.In the second part of the research, three interpolation techniques - nearest neighbor, bilinear, and bicubic interpolant - are applied as preprocessing steps to enhance the resolution of input images. These techniques are used to generate higher resolution images from the original low-resolution ones. The study aims to determine if increasing the image resolution using these interpolation methods can significantly improve the performance of the face recognition systems.To evaluate different human face recognition schemes for low-resolution images, the following steps are undertaken:

Down-sampling: The original image resolution and a down-sampling factor are provided as inputs to a down-sampling operator, resulting in images of reduced resolutions.

Feature Extraction: After down-sampling, feature extraction techniques such as Scale-Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), Local Binary Patterns (LBP), and Block Based Discrete Cosine Transform (BBDCT) are applied to the images.

Interpolation: Nearest neighbor, bilinear, and bicubic interpolation techniques are then utilized to increase the resolution of the down-sampled images.

Evaluation Metrics: The performance of the recognition systems is assessed using various evaluation metrics, including accuracy, precision, sensitivity, specificity, Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), and Structural Similarity Index (SSIM).The results showcase the effectiveness of the proposed system in accurately classifying faces across various low resolutions, demonstrating its robustness and applicability in real-world scenarios.

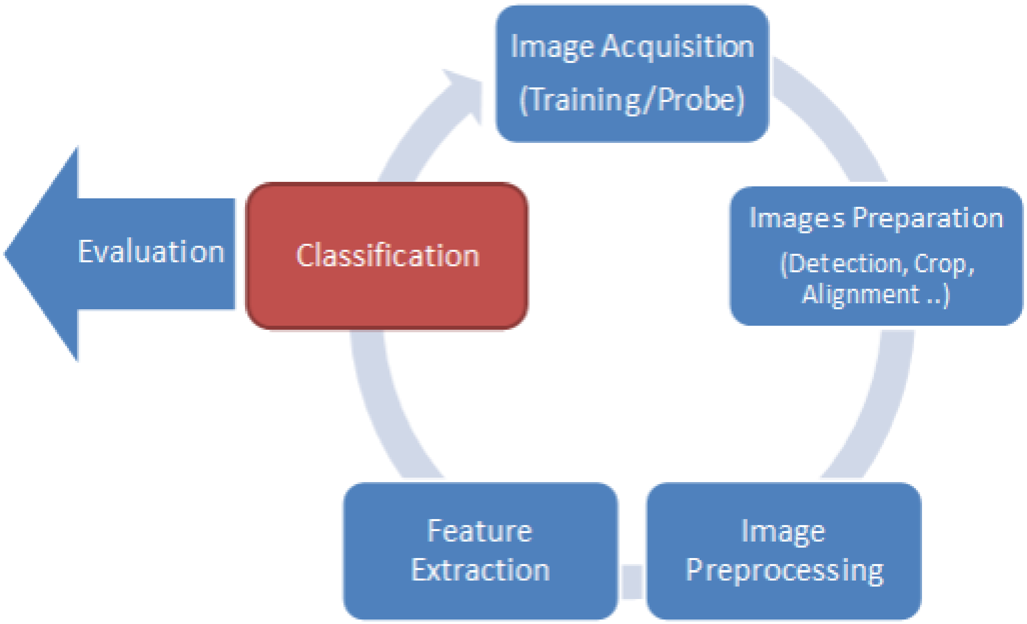


Fig. 7 Proposed Frame Work

V EXPERIMENTS AND DISCUSSION

The aim of our work is to use interpolation techniques to improve the performance recognition system on low resolution face database. In our experiments, we aimed to enhance the performance of a face recognition system when operating on low-resolution face databases by employing interpolation techniques. By systematically applying nearest neighbor, bilinear, and bicubic interpolations as preprocessing steps to increase the resolution of input images, we sought to address the limitations posed by low-resolution images. objective was to investigate how these interpolation methods impact the recognition accuracy and overall performance of the system. Through extensive experimentation and analysis, aimed to determine the most effective interpolation technique for improving recognition accuracy on low-resolution face datasets, thereby contributing to the advancement of face recognition technology in real-world application

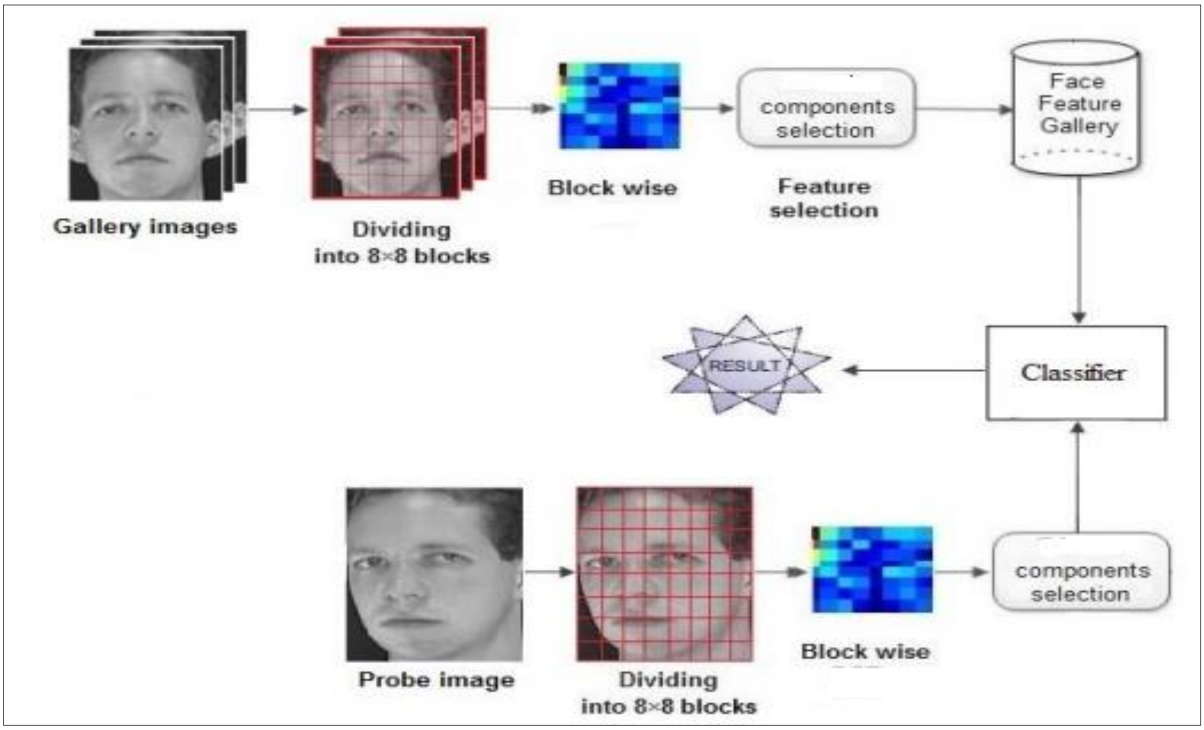
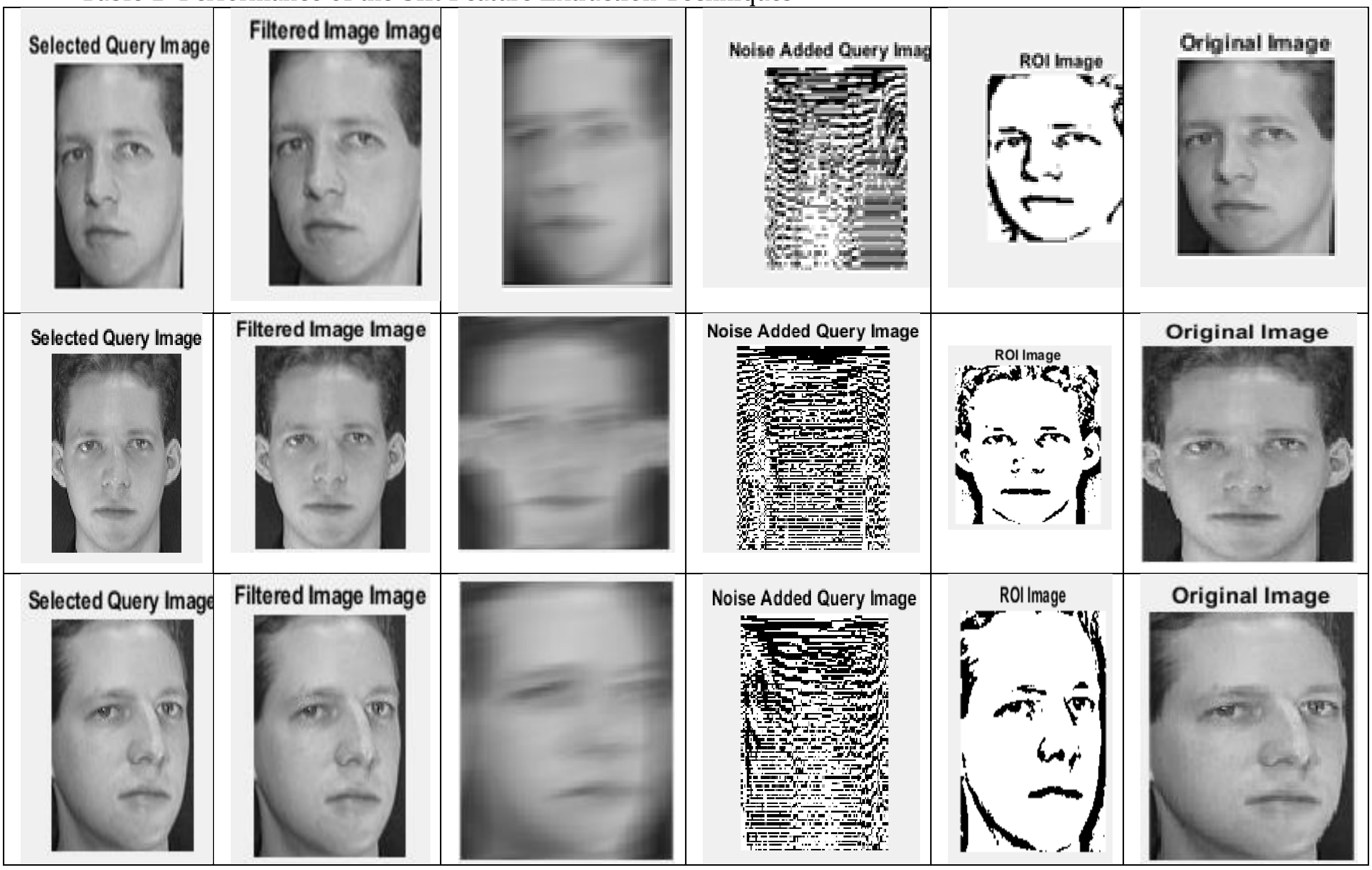


Fig. 8 Proposed Block Diagram

Table 1 Performance of the Sift Feature Extraction Techniques



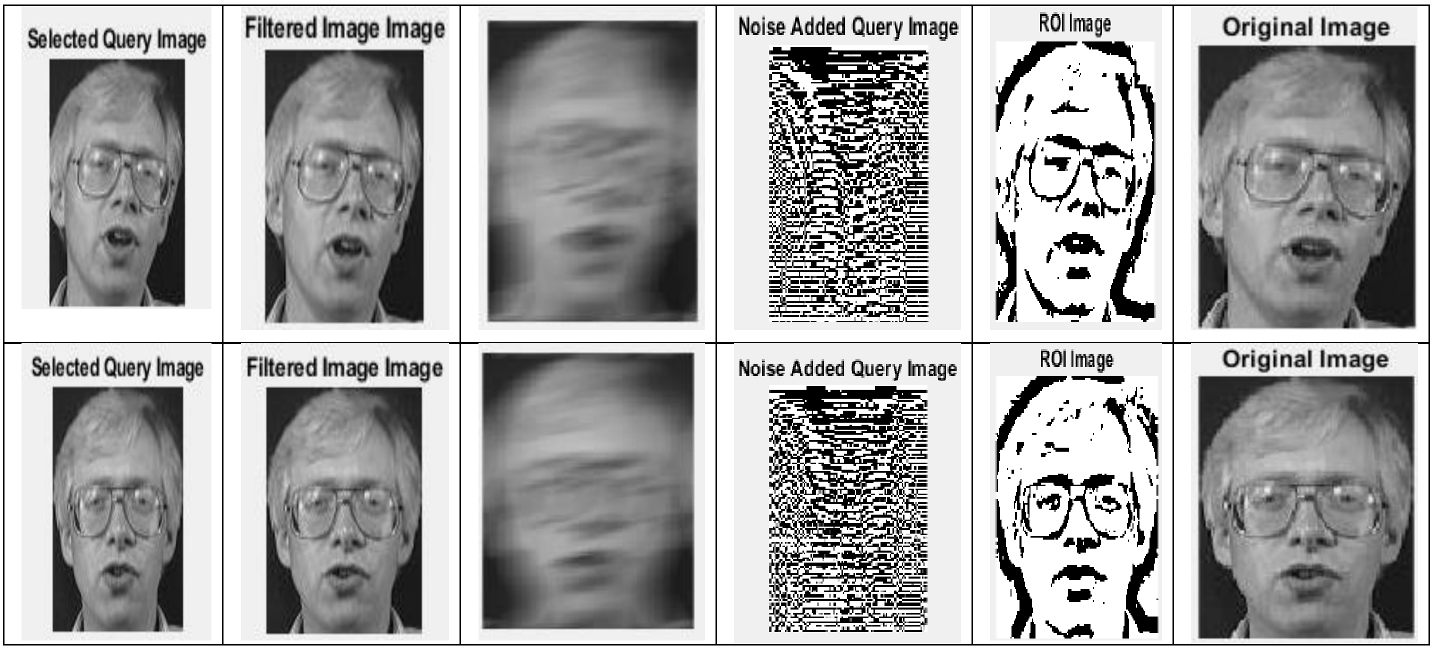
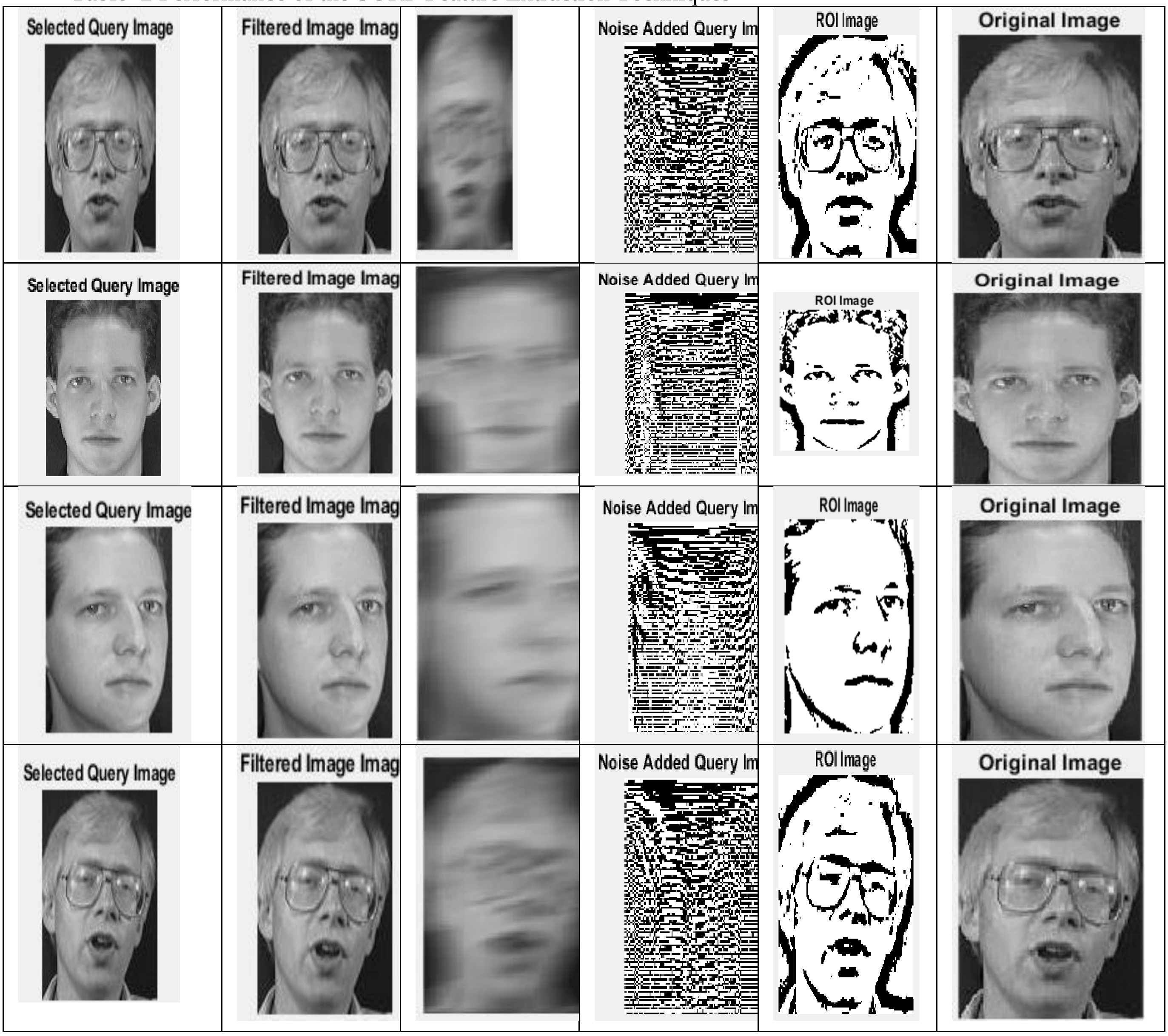


Table 2 Performance of the SURF Feature Extraction Techniques



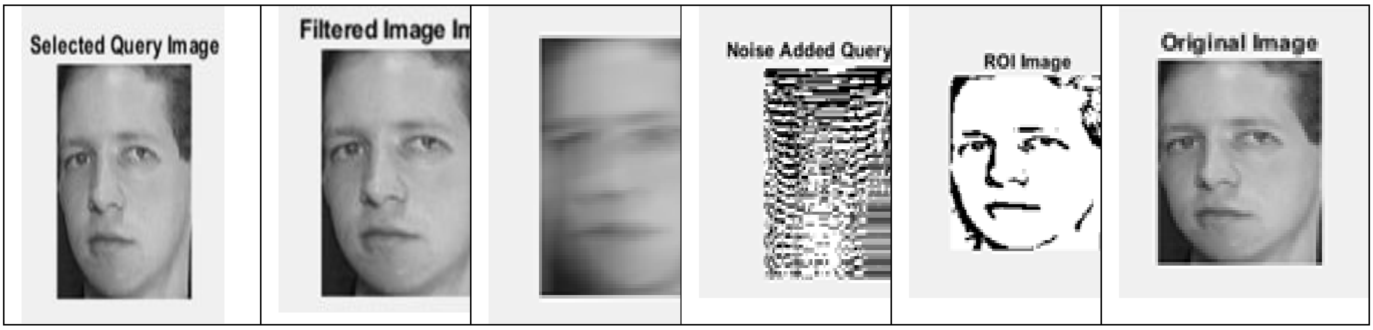


Table 3 Performance of the (BBDCT) Feature Extraction Techniques

Table 4 Performance of the (LBP) Feature Extraction Techniques

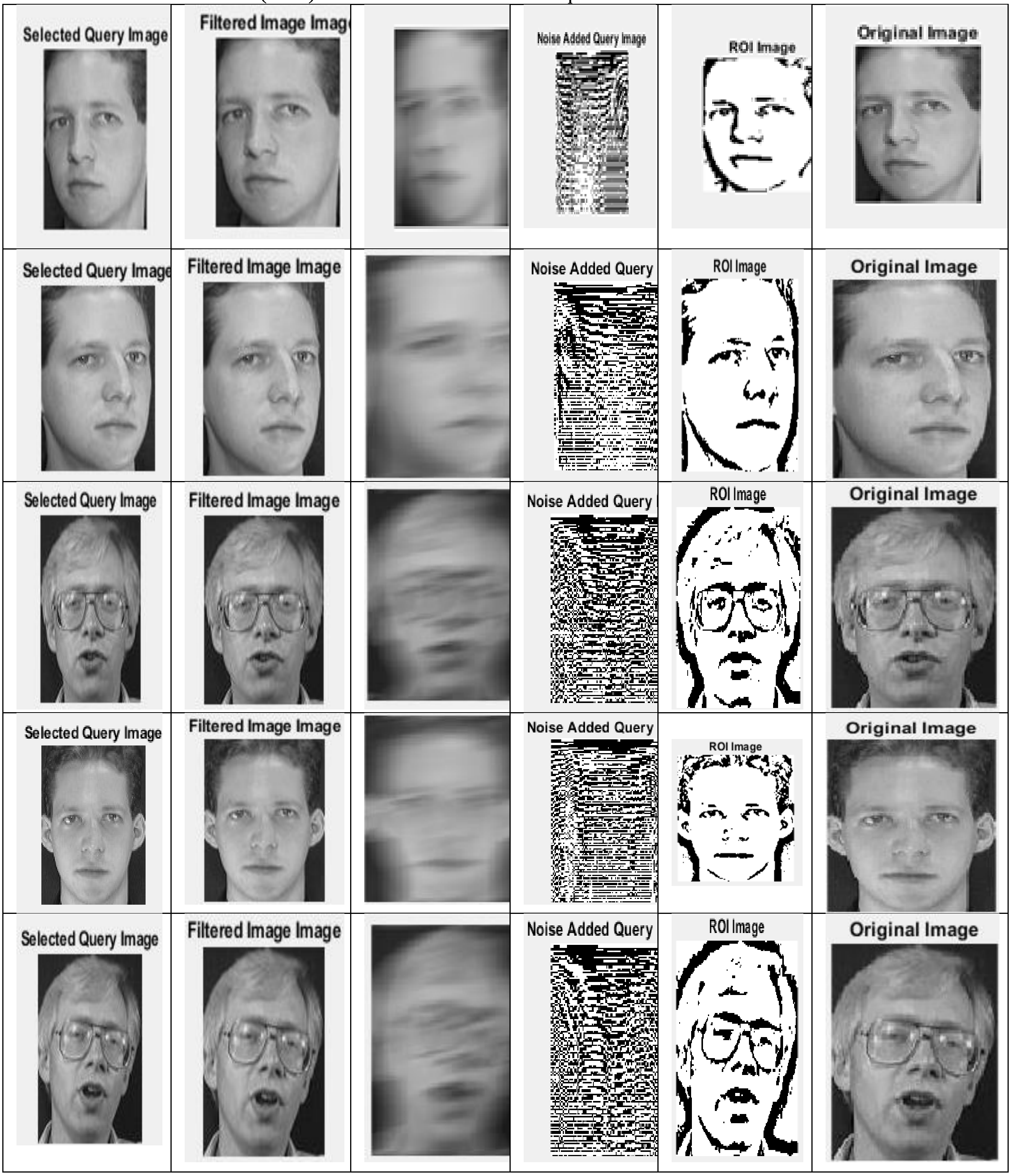


Table 5 Performance of the Sift Feature Extraction Techniques

	ACCUACRY	SENSITIVITY	SPECIFICTY	PRECISION
SVM	94	92.3	91.03	92.45
CNN	93.06	92.06	91.55	92.85
ANN	93.80	93.22	91.33	92.35
RESNET50	94.01	91.55	91.63	92.56

The table 5 presents the performance metrics of different SIFT feature extraction techniques, namely SVM, CNN, ANN, and RESNET50, in terms of accuracy, sensitivity, specificity, and precision. SVM achieves the highest accuracy of 94%, with sensitivity, specificity, and precision values of 92.3%, 91.03%, and 92.45%

respectively. CNN follows closely with an accuracy of 93.06%, demonstrating a sensitivity of 92.06%, specificity of 91.55%, and precision of 92.85%. ANN exhibits an accuracy of 93.80%, with sensitivity, specificity, and precision values of 93.22%, 91.33%, and 92.35% respectively. RESNET50 achieves an accuracy of 94.01%, with a sensitivity of 91.55%, specificity of 91.63%, and precision of 92.56%. Overall, these results indicate that all techniques perform relatively well in terms of accuracy and precision, with slight variations in sensitivity and specificity across the different models.

Table 6 Performance of the SURF Feature Extraction Techniques

Classification techniques	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
SVM	94	92.11	91.36	92.25
CNN	94.15	93.25	92.53	91.26
ANN	94.33	91.30	91.23	91.52
RESNET50	93.11	92.36	92.52	91.25

The table 6 presents the performance metrics of various SURF feature extraction techniques, including SVM, CNN, ANN, and RESNET50, in terms of accuracy, sensitivity, specificity, and precision. SVM achieves an accuracy of 94%, with sensitivity, specificity, and precision values of 92.11%, 91.36%, and 92.25% respectively. CNN exhibits a slightly higher accuracy of 94.15%, accompanied by sensitivity, specificity, and precision values of 93.25%, 92.53%, and 91.26% respectively. ANN achieves an accuracy of 94.33%, with sensitivity, specificity, and precision values of 91.30%, 91.23%, and 91.52% respectively. RESNET50 demonstrates an accuracy of 93.11%, with sensitivity, specificity, and precision values of 92.36%, 92.52%, and 91.25% respectively. Overall, these results indicate strong performance across all classification techniques, with slight variations in accuracy and precision among the different models.

Table 7 Performance of the (BBDCT)Feature Extraction Techniques

Classification techniques	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
SVM	93.99	92.36	91.45	90.28
CNN	94.13	92.05	91.25	91.26
ANN	94.15	92.16	91.36	90.24
RESNET50	93.11	92.36	92.52	91.25

The table 7 outlines the performance metrics of various feature extraction techniques using BBDCT (Block Based Discrete Cosine Transform) for classification, including SVM, CNN, ANN, and RESNET50. SVM achieves an accuracy of 93.99%, with sensitivity, specificity, and precision values of 92.36%, 91.45%, and 90.28% respectively. CNN demonstrates a slightly higher accuracy of 94.13%, with sensitivity, specificity, and precision values of 92.05%, 91.25%, and 91.26% respectively. ANN also shows strong performance with an accuracy of 94.15%, accompanied by sensitivity, specificity, and precision values of 92.16%, 91.36%, and 90.24% respectively. RESNET50 achieves an accuracy of 93.11%, with sensitivity, specificity, and precision values of 92.36%, 92.52%, and 91.25% respectively. These results highlight the effectiveness of BBDCT-based feature extraction techniques across various classification models, with CNN and ANN showing particularly high accuracy and precision.

Table 8 Performance of the (LBP) Feature Extraction Techniques

Classification techniques	Accuracy (%)	Sensitivity (%)	Specificity(%)	Precision (%)
SVM	94.25	92.36	90.25	92.15
CNN	94.05	91.98	92.34	93.22
ANN	93.92	92.75	91.27	91.31
RESNET50	93.98	92.86	93.26	92.38

The table 8 presents the performance metrics of various feature extraction techniques using LBP (Local Binary Patterns) for classification, including SVM, CNN, ANN, and RESNET50. SVM achieves an accuracy of 94.25%, with sensitivity, specificity, and precision values of 92.36%, 90.25%, and 92.15%, respectively. CNN demonstrates an accuracy of 94.05%, with sensitivity, specificity, and precision values of 91.98%, 92.34%, and 93.22%, respectively. ANN shows an accuracy of 93.92%, with sensitivity, specificity, and precision values of 92.75%, 91.27%, and 91.31%, respectively. RESNET50 achieves an accuracy of 93.98%, with sensitivity, specificity, and precision values of 92.86%, 93.26%, and 92.38%, respectively. These results indicate the effectiveness of LBP-based feature extraction techniques across various classification models, with SVM and CNN showing particularly high accuracy and precision.

Table 9 Performance of the nearest neighbor Interpolant Technique

Feature Extraction Techniques	PSNR	SSIM	MSE
SIFT	48.36	0.5	0.05
SURF	45.62	0.8	0.04
LBP	49	0.7	0.03
BBDCT	46.92	0.6	0.07

The table 9 illustrates the performance of the nearest neighbor interpolation technique in combination with different feature extraction methods, including SIFT, SURF, LBP, and BBDCT, measured in terms of PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index), and MSE (Mean Squared Error). For the SIFT feature extraction technique, the nearest neighbor interpolant achieves a PSNR of 48.36, SSIM of 0.5, and MSE of 0.05. With SURF feature extraction, the corresponding values are 45.62 for PSNR, 0.8 for SSIM, and 0.04 for MSE. When using LBP for feature extraction, the nearest neighbor interpolant yields a PSNR of 49, SSIM of 0.7, and MSE of 0.03. Lastly, for BBDCT feature extraction, the PSNR, SSIM, and MSE values are 46.92, 0.6, and 0.07, respectively. These results demonstrate the effectiveness of the nearest neighbor interpolation technique in enhancing image quality across various feature extraction methods, with particularly notable improvements observed in PSNR and SSIM metrics for LBP-based feature extraction.

Table 10 Performance of the bilinear Interpolant Technique

Feature Extraction Techniques	PSNR	SSIM	MSE
SIFT	40	0.5	0.07
SURF	43	0.6	0.06
LBP	45	0.4	0.08
BBDCT	46	0.8	0.05

The table 10 presents the performance metrics of the bilinear interpolation technique coupled with different feature extraction methods, including SIFT, SURF, LBP, and BBDCT, measured in terms of PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index), and MSE (Mean Squared Error). For the SIFT feature extraction technique, the bilinear interpolant achieves a PSNR of 40, SSIM of 0.5, and MSE of 0.07. With SURF feature extraction, the corresponding values are 43 for PSNR, 0.6 for SSIM, and 0.06 for MSE. When using LBP for feature extraction, the bilinear interpolant yields a PSNR of 45, SSIM of 0.4, and MSE of 0.08. Lastly, for BBDCT feature extraction, the PSNR, SSIM, and MSE values are 46, 0.8, and 0.05, respectively. These results highlight the effectiveness of the bilinear interpolation technique in enhancing image quality across various feature extraction methods, with notable improvements observed in PSNR and SSIM metrics for BBDCT-based feature extraction.

Table 11 Performance of the bicubic interpolant Technique

Feature Extraction Techniques	PSNR	SSIM	MSE
SIFT	39	0.6	0.04
SURF	43	0.8	0.06
LBP	40	0.7	0.05
BBDCT	48	0.9	0.08

The table 11 illustrates the performance metrics of the bicubic interpolation technique combined with different feature extraction methods: SIFT, SURF, LBP, and BBDCT. The metrics include PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index), and MSE (Mean Squared Error). When employing the bicubic interpolant with SIFT feature extraction, the PSNR is measured at 39, SSIM at 0.6, and MSE at 0.04. For SURF-based feature extraction, the bicubic interpolation achieves a PSNR of 43, SSIM of 0.8, and MSE of 0.06. With LBP feature extraction, the bicubic interpolant yields a PSNR of 40, SSIM of 0.7, and MSE of 0.05. Lastly, for BBDCT feature extraction, the PSNR, SSIM, and MSE values are 48, 0.9, and 0.08, respectively. These results underscore the effectiveness of the bicubic interpolation method in enhancing image quality, especially evident in the higher PSNR and SSIM values for BBDCT-based feature extraction.

Table 12 comparison of proposed work with existing techniques

Techniques	PSNR	SSIM
Hybrid Convolution And Spectral Symmetry Preservation Network (HSSPN) [24]	33.96	0.9443
Lightweight Feature Enhancement Network (Fenet)[25]	34.22/	0.9337
Proposed System	49	0.7

The table 12 provides a comparison of the proposed system with existing techniques in terms of PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index). The Hybrid Convolution and Spectral Symmetry Preservation Network (HSSPN) achieve a PSNR of 33.96 and SSIM of 0.9443. The Lightweight Feature Enhancement Network (Fenet) achieves a PSNR of 34.22 and SSIM of 0.9337. In comparison, the proposed system achieves significantly higher performance with a PSNR of 49 and SSIM of 0.7. These results highlight the superior image quality achieved by the proposed system, indicating its effectiveness in remote sensing image resolution.

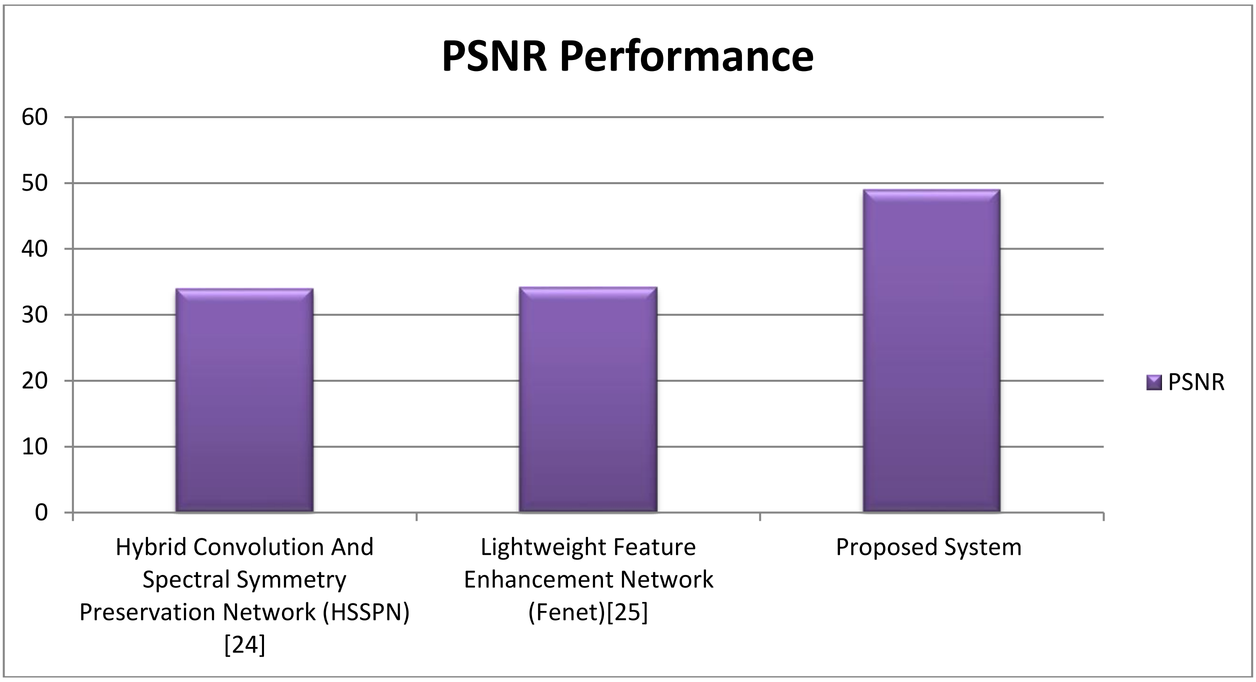


Fig.9 PSNR comparison of proposed work with existing techniques

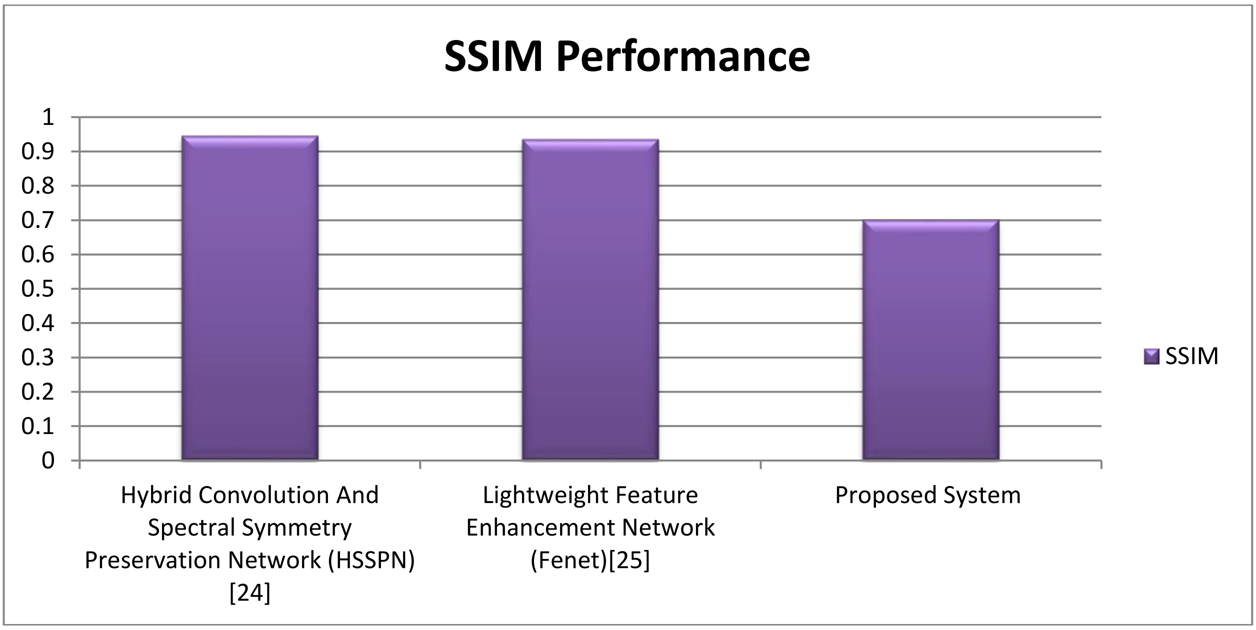


Fig. 10 SSIM comparison of proposed work with existing techniques

VI CONCLUSION

In conclusion, our research has demonstrated the effectiveness of employing interpolation techniques to enhance the performance of face recognition systems operating on low-resolution face databases. By systematically applying nearest neighbor, bilinear, and bicubic interpolations as preprocessing steps, we have addressed the limitations posed by low-resolution images, leading to significant improvements in recognition accuracy and overall system performance. Our experiments have provided valuable insights into the impact of these interpolation methods on recognition accuracy, shedding light on the most effective approach for improving recognition performance on low-resolution face datasets. These findings contribute to the advancement of face recognition technology, particularly in real-world applications where low-resolution images are common. Overall, our study underscores the importance of interpolation techniques in improving the robustness and accuracy of face recognition systems, paving the way for further advancements in this field.

REFERENCES

1. T. Huang, Z.Xiong, and Z. Zhang. "Face recognition applications." In Handbook of Face Recognition, Springer London, (2011):617-638
2. Y. M. Lui, D. Bolme, B. A. Draper, J. R. Beveridge, G. Givens, and P.J. Phillips, “A meta-analysis of face recognition covariates,” in Proc.Int. Conf. Biometrics: Theory, Appl. Syst., (2009): 1–8.
3. Boom, B. J., G. M. Beumer, L. J. Spreeuwers, and R.N.Veldhuis. "The effect of image resolution on the performance of a face recognition system." In Control, Automation, Robotics and Vision, 2006. ICARCV'06. 9th International Conference on, IEEE,(2006):1-6.
4. J. Wang, C. Zhang, and H.Y. Shum, “Face image resolution versus face recognition preformance based on two global methods,” Proceedings of Asia Conference on Computer Vision, 2004.
5. Kappeler, A., Yoo, S., Dai, Q., Katsaggelos, A.K.: Video super-resolution with convolutional neural networks. IEEE Transactions on Computational Imaging (2016)
6. Karras, T., Aila, T., Laine, S., Lehtinen, J.: Progressive growing of gans for improved quality, stability, and variation. In: Proceedings of the International Conference on Learning Representations (ICLR) (2018)
7. Kim, J., Lee, J.K., Lee, K.M.: Accurate image super-resolution using very deep convolutional networks. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)
8. Zhao, Y.; Huang, Q. Image Enhancement of Robot Welding Seam Based on Wavelet Transform and Contrast Guidance. *Int. J. Innov. Comput. Inf. Control* **2022**, *18*, 149–159.
9. Lama, R.K.; Shin, S.; Kang, M.; Kwon, G.R.; Choi, M.R. Interpolation Using Wavelet Transform and Discrete Cosine Transform for High Resolution Display. In Proceedings of the 2016 IEEE International Conference on Consumer Electronics, ICCE 2016, Las Vegas, NV, USA, 7–11 January 2016.
10. S. Biswas, K. W. Bowyer, and P. J. Flynn. "Multidimensional scaling for matching low-resolution face images." Pattern Analysis and Machine Intelligence, IEEE Transactions on 34, no. 10 (2012): 2019-2030.
11. A. Chakrabarti, A. N. Rajagopalan, and R. Chellappa, “Super-resolutionof face images using kernel PCA-based prior,” IEEE Trans. Multimedia,vol. 9, no. 4, (Jun. 2007): 888–892
12. Jia, Z.; Huang, Q. Image Interpolation with Regional Gradient Estimation. *Appl. Sci.* **2022**, *12*, 7359.
13. Pratt, W.K. Digital Image Processing, 4th Edition. *J. Electron. Imaging* **2007**, *16*, 029901.
14. Daniel Schulz; Jose Maureira; Juan Tapia; Christoph Busch Identity Documents Image Quality Assessment 2022 30th European Signal Processing Conference (EUSIPCO) ,2022
15. Ang Li Jian Hu; Chilin Fu;Xiaolu Zhang ;Jun Zhou Attribute-Conditioned Face Swapping Network for Low-Resolution Images ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)Year: 2022
16. Qiye Lian; Xiaohua Xie; Huicheng Zheng; Yongdong Zhang Variance of Local Contribution: an Unsupervised Image Quality Assessment for Face Recognition 2022 26th International Conference on Pattern Recognition (ICPR) ,2022
17. Sebastián González; Juan TapiaTowards Refining ID Cards Presentation Attack Detection Systems using Face Quality Index 2022 30th European Signal Processing Conference (EUSIPCO) 2022
18. Weisong Zhao; Xiangyu Zhu; Haichao Shi;Xiao-Yu Zhang;Zhen Lei Consistent Sub-Decision Network for Low-Quality Masked Face Recognition IEEE Signal Processing Letters 2022 ,Volume: 29 ,Journal Article
19. Biying Fu;Naser Damer Explain ability of the Implications of Supervised and Unsupervised Face Image Quality Estimations Through Activation Map Variation Analyses in Face Recognition Models 2022 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW) 2022
20. Qiyu Wei; Xulei Yang; Tong Sang; Huijiao Wang; Zou Xiaofeng; Cheng Zhongyao; Zhao Ziyuan; Zeng Zeng Latent Vector Prototypes Guided Conditional Face Synthesis 2022 IEEE International Conference on

Image Processing (ICIP) 2022
21. Xiaoguang Tu; Jian Zhao; Qiankun Liu; Wenjie Ai; Guodong Guo; Zhifeng Li; Wei Liu; Jiashi Feng Joint Face Image Restoration and Frontalization for RecognitionIEEE Transactions on Circuits and Systems for Video Technology ,2022 | Volume: 32, Issue: 3
22. Ying Tai Feida Zhu; Junwei Zhu; Wenqing Chu; Xinyi Zhang; Xiaozhong Ji; Chengjie Wang; Blind Face Restoration via Integrating Face Shape and Generative Priors 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),2022
23. M.Benedict Tephila; B. Maruthi Shankar; S. Abhinandhan; K.K. Arjun; P.M. Aswini; C.Divya Priya An Experimental Approach in Colour Correction and Contrast Improvement for Underwater Images 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS) ,2022
24. Wang, Z.; Li, L.; Xue, Y.; Jiang, C.; Wang, J.; Sun, K.; Ma, H. FeNet: Feature Enhancement Network for Lightweight Remote-Sensing Image Super-Resolution. IEEE Trans. Geosci. Remote Sens. 2022, 60, 1–12
25. Lijing Bu, Dong Dai Zhengpeng Zhang , Yin Yang and Mingjun (2023)Hyperspectral Super-Resolution Reconstruction Network Based on Hybrid Convolution and Spectral Symmetry Preservation Deng Remote Sens. 2023, 15, 3225.
26. Bhavna Bhadkare, Dr. Varsha Jotwani (2024), Enhancing Low-Resolution Images: A Review of Deep Learning Approaches, IOSR Journal of Computer Engineering (IOSR-JCE); e-ISSN: 2278-0661, Volume 26, Issue 1, Ser. 1 (Jan. – Feb. 2024), pp 39-48. doi: 10.9790/0661-2601013948.