Volume (1) No (12), 2025
Journal of Global Research in Multidisciplinary Studies (JGRMS)
Review Paper
Available online at https://saanvipublications.com/journals/index.php/jgrms/index

Facial Emotion Detection in Low Resolution Image: Review and Challenges

Mr. A.R. Landge
PhD. Scholar, Electrical and Electronics Engineering
Mandsaur University (M.P)
amollandge04@gmail.com

Dr. Manish Jain
Professor, Electrical and Electronics Engineering
Mandsaur University (M.P)
manish.jain@meu.edu.in

***Abstract*—Face emotion is a form of non-verbal communication extensively used by human beings. Detecting emotion to ascertain a human's mental state is a vital research area. Emotion detection for images with low resolution is an area where research is going as practical real-world images will be mostly low resolution. Real-world applications like surveillance, forensic and drone imagery capture low-resolution images due to camera resolution, distance of face, and motion of the camera as well as face and may contain noise. Most of time, researchers consider image less than 32* 32 pixels as low-resolution image as there is no fixed definition for categorizing low-resolution images. Human face emotion detection performance depends on factors such as lighting conditions, facial occlusion, and facial movements. The loss of resolution will reduce the image's vital facial details. Performance on facial emotion recognition is improving with newer deep learning models, but model accuracy remains a major issue.**

This paper will elaborate methods and algorithms used in face emotion detection in low resolution conditions. The paper will summarize different datasets, methodology and performance parameters and challenges.

Keywords—Machine Learning (ML), Risk Assessment, Healthcare Analytics, Insurance Fraud Detection, Predictive Modeling, Patient Stratification, Claims Management.

I. Introduction

During the past few decades, automatic facial expression recognition has been an extensive research area in computer vision and pattern recognition. It plays an important role in applications like surveillance and security, human computer interaction (HCI), and data driven animation . The main functions in surveillance include identifying criminals in crowds, in busy places like stadiums , airports, malls and railway stations. Other useful applications are attendance of employees and students , finding people in databases such as drivers license , PAN and Aadhaar . The face possesses some discriminating features for human beings which are used for identification. Face detection is a complex task due to the variations caused by pose, illumination, and so on but complexity increases as resolution of image decreases [1]. In real world scenario images captured by cameras may be of low resolution due lighting condition, blurriness, distance of camera etc. face recognition performance of LR images decreases due lack of availability of discriminating features to extract the features .

Most of time researchers consider image less than 32* 32 pixels as a low resolution image as there is no fixed definition for categorizing low resolution images [2]. Capturing images in real world scenarios using cameras , surveillance and mobile or any other device may result in degradation of image quality and loss of significant spatial details [3][4]. Resolution of these images is affected due to sub optimal lighting conditions , facial occlusion or camera sensor capabilities resulting in generation of LR images which lack distinctive details and clarity. Accuracy is the major challenge face detection in LR images,

The very first step in any face detection system is image acquisition using image sensors in imaging devices. Image acquisition can be done in well lite and controlled environment generating good quality images with clear face features or it can be done in a dim lite and uncontrolled environment which generally produce LR images with face occlusions [5]. Preprocessing is the next step which does processes like noise reduction, resizing , normalization and color conversion on captured images. Face detection algorithms scans images and extracts key facial features which are discriminative. It also removes redundant features [6][7]. Techniques lie SVM , neural network or traditional techniques like euclidean distance are used for determining if a face is detected and further for classification of image.

II. Background and Related Work

High resolution images respond better to face detection algorithms but low resolution images performance response decreases sharply. The main issue in LR images is less number of pixels describing the face , so face features are not readily available as compared to HR images [8]. Other major factors are condition of light, resolution of capturing device, focus ,blurriness that can degrade performance. Low resolution Face recognition systems contain traditional stages like face detection , feature mapping, face recognition along with some specialized stages like super resolution for low resolution problems.

A. Face Detection

Traditional face detection algorithms calculate similarities between different images with techniques like LBP, SURFand HOG. Deep learning algorithms like CNN, YOLO, SSD and faster CNN capture variances of face when the training dataset is large enough. the real time variations like position of head, blurriness , lighting condition makes face detection difficult [9]. The three stage CNN progressively refines facial regions with increasing computational complexity improving efficiency of face detection. To improve face detection calibration stage based on CNN is introduced. But traditional calibration nets are trained independently, potentially leading to suboptimal overall performance [10]. In Jointly Trained Cascaded CNN (JDA) framework for face detection where multiple CNNs in a cascade are trained end-to-end in face detection [11]. A face detection algorithm that uses a cascade of convolutional neural networks (CNNs) with specific improvements to progressively filter and refine candidate face regions to enhance detection accuracy, speed, and robustness in challenging conditions.

Faster R-CNN is a landmark object detection framework that introduces Region Proposal Networks (RPNs) to replace traditional, slow region proposal methods (like Selective Search), enabling end-to-end, real-time object detection. A face detection approach leveraging the high accuracy and robustness on the Faster R-CNN framework, adapts the general-purpose object detector for the specific task of face detection in complex environments, such as varying poses, lighting, and occlusions [12]. Faster R-CNN with targeted modifications of its components and training strategies, like tuning anchors, feature extractors, and training data significantly improve performance with better accuracy and robustness in challenging scenarios [13]. The Single Shot MultiBox Detector (SSD) is a real-time object detector performs object localization and classification in a single forward pass, making it significantly faster while maintaining competitive accuracy. SSD significantly proved that one-stage detectors could achieve both high speed and strong accuracy.SSH (Single Stage Headless) is a real-time face detection that eliminates fully connected layers and applies detection directly on convolutional feature maps. SSH is a lightweight, single-stage detector that can rival or outperform more complex models like Faster R-CNN for face detection without sacrificing speed. Face SSD, an optimized face detection model, is based on the Single Shot MultiBox Detector (SSD) architecture. It strikes a balance between the accuracy and the efficiency needed for deployment in real-time scenarios especially on embedded or resource-constrained devices. A hybrid method combining an improved SSD with a target tracking algorithm for face detection in video shows how combining lightweight detection with intelligent tracking can address the challenges of face detection in real-time face detection with accuracy

An automatic face detection system integrates classical geometric analysis modern deep learning methods like YOLO object detection algorithm improves the accuracy and reliability of face detection in real-time, real-world environments where accuracy, speed, and resilience to visual noise are essential [14] [15]. A real-time face detection system for live video streams, combining the YOLO with a VGG16's deep feature extraction capabilities can significantly enhance feature extraction by improving both detection accuracy and processing speed [16] [17]. An enhanced face detection method that combines YOLOv5 with integrating image enhancement techniques a super-resolution reconstruction (SRR) module improves detection accuracy, especially in low-resolution or blurry images. Dual Shot Face Detector(i.e., two detection passes) uses Feature Enhance Module (FEM) to take original feature maps and enhance them via dilated convolutions and context aggregation. Progressive Anchor Loss (PAL) uses the first shot uses a coarser set of anchors and the second shot refines using enhanced features and a different anchor distribution in progressive manner [18] .Improved Anchor Matching (IAM) module does regressor's initialization and reduces mismatch between face sizes and anchor boxes

B. Feature Mapping Techniques

Feature mapping refers to the process of transforming raw data into a new representation space called a "feature space" to make it more suitable for tasks like classification, clustering,Feature mapping transforms input data into a space that is more discriminative or separable [19]. Mapping bridge the gap between HR features to LR features for effective recognition. Different techniques like scaling, dimension reduction, or transformations reduce the feature size but preserve important information [20].

III. Methodologies

A. Traditional Methods

Local Binary Pattern (LBP) as an effective texture descriptor. It outperform more complex methods despite its computational simplicity [21][22]. LBP did so by capturing local spatial patterns and gray-scale contrast, relying on the statistical distribution of local pixel patterns for classification. Figure 1 : Working of LBP operator [3]. A novel method utilizing LBP with modifications including different scales and temporal fusion for local feature matching face recognition to low-resolution video, for faces as small as 8x8 pixels [23]. The approach tries to overcome limitations of applying local feature approaches to low-resolution images. A kernel-based manifold method that integrates gait features with low-resolution face features using Sparse LBP histograms with performance achievement of 78 % on VidTIMIT and 84 % on the HONDA dataset.It maps nonlinear features from faces and gait energy images into a common subspace to minimize distance between features from the same person, demonstrating improved accuracy over other methods. A knowledge distillation framework extracts features from low-resolution (LR) images for face recognition using knowledge gained from high-resolution (HR) images to improve performance of face recognition. Knowledge gained from informative features from an HR-trained network and transfers to an LR-trained network by decreasing the distance between them. To find out face similarities a cosine similarity measures and aligns the HR and LR features. The approach effectively. Achieved a 3% improvement over the previous benchmark. The performance is recognised on the CASIA and ORL dataset.

A neural networks to classifies low resolution images after upscaling. Interpolation techniques are used to recover missing details of the low-resolution images to upscale them before recognition [24]. Iterpolation techniques along with feature extraction methods like Adaptive DCT , Block-based DCT and DWT. Interpolation has improved recognition performance of low resolution images [25]. A 2D face synthetic data generator using 3D face models capturing 3D-rendered synthetic data that allows for fully controlled and automatic manipulation of features like pose, scale, facial occlusion, background, and illumination. A raster scan technique to extract block DCT for object features like facial occlusion blur and low resolution. The effective use of synthetic data can be done to train more face detectors, to improve their performance making them more robust [26]. An example-based image super-resolution (SR), Sparse Neighbor Selection (SNS) scheme that simultaneously selects the optimal neighbors and determines the reconstruction weights, making the process more efficient and optimal. An extended

Robust-SL0 algorithm to perform sparse selection and reconstruction [27]. Traditional Neighbor-Embedding (NE) algorithms performed neighbor search and weight calculation as two separate, sub-optimal processes for synthesizing high-resolution image patches. To enhance the quality and speed of neighbor search. Local structural information. Specifically, Histograms of Oriented Gradients (HoG) of low-resolution patches are used to cluster the training data as it is not sensitive to noise. This ensures that the chosen neighbors have similar local geometric structures[28]. By adaptively selecting neighbors from an associated, smaller subset based on HoG features, the proposed SNE method significantly improves the speed of HR image synthesis while achieving competitive or superior SR quality compared to existing state-of-the-art methods.

B. Deep Learning-Based Methods

Deep learning methods have improved face emotion detection methods accuracy to a large extent. they have improved low resolution face emotion detection substantially [8]. Xia Zhang et .al investigated the application of deep learning models, particularly Deep Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), in face feature extraction and generation. Analysis of the importance of large-scale datasets and how data augmentation techniques are essential for improving model performance and robustness. Deep learning models, like Convolutional Neural Networks (CNNs), have automated the extraction of complex facial features, surpassing traditional methods like LBP and Eigenfaces. Deep learning models can learn and capture rich features in face images, achieving more accurate and stable face recognition. Major steps in FER systems include face detection, alignment, feature representation, and matching, with feature extraction has a vital role that affects the overall system performance[29]. Feature Extraction via Non-Negative Matrix Factorization (NMF) extracts fundamental features from Several Angular Faces Using a Deep Learning-Based Fusion Technique for Face Recognition. Generative Adversarial Networks (GANs) fuses the extracted features from multiple angular views , to synthesize a comprehensive feature vector representative of a frontal face [30]. An advanced convolutional neural network (CNN) model designed to improve facial emotion recognition, particularly under challenging conditions such as occlusions, scale variations, and illumination changes. The model achieves a notable accuracy of 74.92% on the FER-2013 dataset, 99.47% accuracy on the CK+ dataset and 98.5% on the FERG dataset. A DeepLabv3+ backbone is integrated with a Swin-Conv-Dspp (SCD) Transformer architecture to mitigate “hole” effects with GLTB .It helps extract important details even when parts of faces are occluded or only limited pixels are visible and Crocodile Search Algorithm (CSA) fine-tunes model parameters to improve classification accuracy.

C. Super-Resolution Methods

Low-resolution (LR) images such as those from surveillance footage, low-quality webcams, or compressed videos lack discriminative features like wrinkles, muscle movements, or subtle expressions [31]. One of the ways to address these challenges in face emotion detection of low resolution images is to convert them to high resolution images .Super-resolution (SR) methods are mostly used for this LR to HR image conversion .Applying super-resolution can enhance facial details, making emotion.

1) Multi-Image Super Resolution

The multiple consecutive frames instead of a single frame to reconstruct a higher-resolution, sharper face image using Active Appearance Model (AAM) Fitting to form a super-resolved image. The super-resolved images lead to higher identification rates compared to the original low-resolution frames [32] [33]. A deep network that simultaneously handles feature extraction, non-rigid alignment, super-resolution, and identity preservation together. Ulti-frame data uses additional information from adjacent frames that can help reconstruct a higher-fidelity face image. Wrapping networks predict transforms to align adjacent frames to the central frame; warp feature maps accordingly. Reconstruction network uses Fuse warped features with central frame features to reconstruct the high-resolution version of the central frame. YouTube Faces (YTF) dataset is used for multi-frame sequences super-resolution face alignment.

Multi-frame image restoration tasks, is reparametrizes the maximum a posteriori (MAP) formulation into a learned feature (latent) space, rather than optimizing in image space directly. Model uses technique like learned error metrics , Latent representation and decoder and Feature-space degradation [16]The certainty weighting helps reduce influence of poorly aligned or noisy regions. Same structure is used for super-resolution and denoising tasks. Burst / multi-frame super-resolution (MFSR) process reconstruct a high-resolution RGB image given a burst of low-resolution RAW images . The Enhanced Burst Super-Resolution, a pipeline that splits the task into three stages: alignment, fusion, and reconstruction, with tailored modules for each [34] [35] . It uses Feature Enhanced Pyramid Cascading & Deformable Convolution , Cross Non-Local Fusion , reconstruction is done using Long Range Concatenation Network (LRCN) and Cascaded Residual Pathway (CR).EBSR achieves strong results: in the NTIRE21 BurstSR Challenge, it ranked 1st place in the real track and 2nd place in the synthetic track.

2) Map-based Methods

Using MAP-based (Maximum a Posteriori) methods for facial emotion recognition via a super-resolution (SR) approach is a powerful strategy to enhance performance, especially when working with low-resolution or degraded facial images. A method for video super-resolution, generating high-resolution video from low-resolution video input by a process called as video hallucination. The redundant information across multiple frames in video sequences is used to reconstruct fine high-resolution details that are not explicitly present in any single frame. The spatio-temporal regularization maintains no flickering and sharp details without artifacts & helps suppress noise and errors from motion estimation [36].DCT with face hallucination reconstructs a high-resolution (HR) face image from a low-resolution (LR) input [37]. The DCT preserves lower-frequency coefficients in the LR image (after appropriate scaling), but the high-frequency details are largely lost or down-sampled. Learning-based Mapping / Regression captures the correlation between existing low-frequency content and missing higher-frequency detail [38]. Inverse DCT reconstruction, Patch-Based Enhancement and Regularization gives sharper results in restoring edges or texture detail [39]. Super-resolution image reconstruction from multiple low-resolution images is done using probabilistic image models. Bayesian MAP technique along with Learned image priors using Markov Random Fields (MRFs) better captures the natural statistics of images and improve reconstruction quality [39]. Multiple-Model Learned Image Super-Resolution (MMSR) Benefits From Class-Specific Image Priors model .Instead of choosing just one model, they use a bank of SR class-specific models is chosen . MMSR with generic fusion outperforms generic SR or any single class-specific model. MMSR also reduces the performance variance (across images) compared to generic models, making results more stable [40]. Visual quality is also better: edges (text edges etc.) preserved, textures, etc. Fusion helps to bring in the strengths of different specialized models.

3) Example-based Methods

Example-based super-resolution (SR) techniques use external or internal image examples to learn mappings from low-resolution (LR) to high-resolution (HR) image patches. These methods do not rely on explicit modeling of the imaging process (e.g., blur, downsampling kernels), but rather learn from data how high-frequency details look and how they relate to low-resolution counterparts. A joint, sparse neighbor embedding approach selects neighbors and solve weights so the reconstruction is more consistent and less prone to mismatch artifacts limits candidate sets helps speed up the search without sacrificing reconstruction quality [41][42]. A partial supervision improves SR performance over pure NE methods, especially when the training set does not well cover the test patch variety.The improvement is more pronounced for underrepresented patterns which benefit from the geometry preservation term. A single-image super-resolution (SISR) recovers a HR image from a single LR input by leveraging a training set of LR–HR image pairs. The coefficients that reconstruct a LR patch from its LR neighbors can be reused to reconstruct the corresponding HR patch from its HR neighbors [43]. This neighbor-embedding method produces plausible HR images with sharper details compared to simpler interpolation or direct lookup methods. The face hallucination from a single low-resolution (LR) image[44]. The combines example-based methods with morphable face models and error correction to produce better hallucinated (HR) faces.They evaluation is done on standard face datasets including MPI, XM2VTS, and KF using Metrics like reconstruction error, Structural Similarity Index (SSIM), and face recognition rate .

4) FFD-based Methods

FFD-based methods in super-resolution uses Free-Form Deformation (FFD) for aligning or modeling geometric variations in images — particularly useful in tasks like face hallucination, non-rigid object alignment, or multi-view SR [25] Super-resolution reconstruction and face recognition from LR input are improved by combining nonrigid registration with multi-frame SR fusion to reconstruct a higher-quality face image that is better suited for recognition. method shows improved registration accuracy, more coherent SR reconstructions under face deformation / expression changes, and higher face recognition accuracy in the LR-to-HR pipeline [45]. Sharp edges and identity are preserved in the super-resolved face, as facial components are distorted when upscaling from a low-resolution (LR) image. Method improves on preserving facial component clarity compared to existing models. Edges are enhanced or preserved not by changing the loss function, but by modifying the network architecture to explicitly integrate edge detection and merging. Results show sharper edges (contours, object boundaries) than comparable SR methods that do not explicitly fuse edge information. A n edge-informed SR approach that decouples the reconstruction of structure (edges) and texture (color / content) to improve final image fidelity. method reformulates the SISR problem as a kind of inpainting task [46] when upsampling, “new pixels” appear between existing pixels, and those can be thought of as missing regions to be filled in (inpainted) using both structural (edges) and textural cues.the method produces sharper edges and better structural delineation — the decoupling of edges and textures results in images with more distinct boundaries relative to many baselines.

D. Face Recognition

Face emotion detection systems often struggle with low-resolution (LR) images due to the lack of fine facial details required for accurate emotion classification. This is particularly challenging in real-world scenarios like surveillance footage or video calls under poor network conditions., super-resolution (SR) techniques like deep learning model (e.g., SRCNN, ESRGAN) upsamples and reconstructs a high-resolution face image enhancing the quality of low-resolution input. These SR-enhanced images preserve important facial features such as eyes, mouth, and wrinkles, which are crucial for emotion recognition[47].A convolutional neural network (CNN) or other classifier detects emotions (e.g., happy, sad, angry) from the super-resolved image[48]. A parallel deep architecture augmented with customized, learnable filters extracts discriminative features even from degraded face images. A bank of filters with multiple scales and orientations captures scale- and rotation-invariant features. Parallel Deep Network Streams architecture uses parallel branches (streams) to process features at different levels / resolutions [30]. A CNN-based model that can directly recognize faces from noisy inputs, being robust to various noise types and intensities .NR-Network outperforms hand crafted FLBP and NRLBP variants, as well as CNN baseline BN1 and BN2 baselines. GPEN (a face-focused super-resolution / enhancement model) to convert a low-resolution face image into a higher-resolution version that better preserves facial details. GPEN + FaceNet combined model showed improved recognition accuracy compared to using FaceNet alone on low-resolution inputs, validating the effectiveness of the super-resolution step[49]. A combination knowledge distillation + an identity-preserving network to emphasize low-frequency discriminative features and bridge the LR–HR feature gap. A ResNet structure with low-pass filtering via discrete wavelet transform (DWT) filters out high-frequency noise forcing the network to focus more on low-frequency components Applying SR methods (OpenCV DNN SR, SRGAN) to LR images substantially improved metrics like PSNR, SSIM, and MSE compared to bicubic interpolation. Among the SR methods they tested, the OpenCV DNN SR method often produced better PSNR / SSIM than SRGAN or bicubic interpolation [50]. A resolution-invariant face recognition method in uncontrolled environments [51][52]. The result suggests that integrating resolution-aware design (multi-scale feature extraction, alignment losses) helps mitigate the drop in accuracy due to resolution differences.

E. Datasets & Performance Measures

To perform analysis of models using low resolution images requires different datasets. The table summarizes different types of datasets used in Low resolution images. The consistent performance measures used in analyzing the performance of LRFR models are Accuracy, Recall, Precision, mAp(mean Average Precision), Frames per Seconds (FPS [53].

TABLE I. SUMMARY OF DATASET USED IN LOW RESOLUTION FACE EMOTION DETECTION

Dataset	HR/LR	Number of images	Variations/Details
FDDB [54]	HR	2845 images with a total of 5171 faces	• gray-scale and color images; • Low-quality images containing occlusions, difficult poses, and low-resolution and out-of-focus faces
Surveillance Cameras Face Database (SCface)) [55]	HR+LR	4160 images from 130 subjects	• the visible and infrared spectrum • contains nine different pose images suitable for head pose
UCCSFace [56]	HR	6337 images	• pose, illumination, scale, expressions, occlusion
UTKFace [57]	HR+LR	20k images	• pose, facial expression, illumination,occlusion, resolution • Face images with annotations of age, gender, and ethnicity.
Unconstrained Face Detection Dataset (UFFD) [58]	HR+LR	6424 images	• Face detection with distractors like fog, blur,etc.
QMUL-TinyFace [59]	LR	169,403 images from 5139 subjects	• unconstrained low-resolution face images collected from different web sources
QMUL-SurvFace [60]	LR	467,507 images from 15,573 subjects	• unconstrained low-resolution face images • collected from surveillance cameras
CASIA-WebFace [61]	LR	494,414 face images of 10,575 real Identities Collected from web	• CASIA-WebFace is processed by face detection, face landmarking, and alignment.
AgeDB	HR	16, 488 images from 568 subjects	• Images captured under completely uncontrolled, real-world conditions having different poses, containing noise, bearing various expressions, containing occlusion, etc. • Images of actors/actresses, writers, scientists,
IJB-S	HR+LR	Over 3 M images from 202 subjects; 350 surveillance videos spanning 30 h in total, 5, 656 enrollment images, and 202 enrollment videos.	• Surveillance videos are collected at the Department of Defense (DoD) training facility
WiderFace	HR	2203 images and label 393,703 faces	• Variability in scale, pose, and occlusion
Labeled faces in the wild (LFW)	HR	13,000 images from 5749 subjects	• Collected from web
VGGFace	HR	2.6 million face images of 2622 people	• Images of celebrities, public figures, actors, and politicians
Chokepoint	LR	48 video sequences and 64,204 face images.	• Variations in terms of illumination conditions pose, sharpness, as well as misalignment • Captured from three different cameras
CalTech 10k	HR	10,524 human faces from 7092 images	• portrait images, groups of people • The average image resolution is 304 × 312 pixels
Our Database of Faces (ORL)	HR	400 images from 40 distinct subjects	• each image is 92 × 112 pixels • images were taken at different times, varying the lighting, facial expressions • Now called as AT&T dataset
MegaFace	HR	4753,320 faces of 672,057 identities from 3311,471 photos	• All the images are collected from Flickr (Yahoo's dataset) • Now this dataset has been retracted and is not available for public use
Large-scale CelebFaces Attributes (CelebA)	HR	202,599 faces of 10,177 identities	• Collected from web • 40 binary labels indicating facial attributes like hair color, gender, and age.
CMU Multi-PIE [1]	HR	750,000 images of 337 people	• Varying poses and illuminations by using a system of 15 cameras and 18 flashes

IV. CHALLENGES

We summarize some challenges that affect the performance of Low Resolution Face emotion Detection.

- Dataset: The most important part of achieving better performance of low resolution face emotion detection is the requirement of a number of low resolution databases. A few datasets of LR face images are available. The other approach is to train the model on HR dataset and map it to LR images but to make this approach work there is need of a strong resolution algorithm [62].
- Lack of discriminative information: major drawback of low resolution images is absence of distinctive features in images due to issues such as lighting conditions, variations in pose, occlusion, and blurriness.
- Efficiency of LRFR: due to lack of databases, to improve performance of LRFR addition of a super resolution method becomes mandatory[63]. But now

the performance of the SR method determines the performance of LRFR.

- Variety in dimensions of the face: major issues with images captured by camera is the distance of face from camera giving altogether different dimensions to face based on camera position and distance[64]. The degradation in performance is attributed to deviation and distance.
- Non-aligned faces: non alignment in faces is an issue even in HR images ,but inLR images it becomes very difficult impacting the performance.

V. Conclusion

LRFR has a wide variety of applications particularly in areas like security. So a lot of research is going on to improve the performance of LRFR. The paper gives detailed methodology used in LRFR. It underscores the importance of the super resolution method for mapping high resolution images with low resolution images .Paper lists out different methodologies, different datasets and challenges of LRFR approach. The dataset mentioned can help researchers carry out experimentation and check performance of their models. Many solutions are proposed for LRFR but lack LR databases are a major hindrance for researchers.

References

[1] N. Khan, A. Singh, and R. Agrawal, "Enhancing Feature Extraction Technique Through Spatial Deep Learning Model for Facial Emotion Detection," *Ann. Emerg. Technol. Comput.*, vol. 7, no. 2, pp. 9–22, Apr. 2023, doi: 10.33166/AETiC.2023.02.002.

[2] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognit.*, 1996, doi: 10.1016/0031-3203(95)00067-4.

[3] C. Herrmann, "Extending a local matching face recognition approach to low-resolution video," in *2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance*, IEEE, Aug. 2013, pp. 460–465. doi: 10.1109/AVSS.2013.6636683.

[4] X. Ben, J. Mingyan, Y. J. Wu, and W. Meng, "Gait feature coupling for low-resolution face recognition," *Electron. Lett.*, vol. 48, pp. 488–489, 2012, doi: 10.1049/el.2011.4041.

[5] A. Elazhari and M. Ahmadi, "A neural network based human face recognition of low resolution images," in *2014 World Automation Congress (WAC)*, 2014, pp. 185–190. doi: 10.1109/WAC.2014.6935767.

[6] J. Han, S. Karaoglu, H.-A. Le, and T. Gevers, "Object features and face detection performance: Analyses with 3D-rendered synthetic data," in *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021, pp. 9959–9966. doi: 10.1109/ICPR48806.2021.9412915.

[7] Xinbo Gao, Kaibing Zhang, Dacheng Tao, and Xuelong Li, "Image Super-Resolution With Sparse Neighbor Embedding," *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3194–3205, Jul. 2012, doi: 10.1109/TIP.2012.2190080.

[8] X. Zhang, "Research on face recognition based on deep learning algorithm," *Int. Conf. Internet Things Mach. Learn.*, pp. 320--324, 2023.

[9] Y. Wong, S. Chen, S. Mau, C. Sanderson, and B. C. Lovell, "Patch-based probabilistic image quality assessment for face selection and improved video-based face recognition," in *CVPR 2011 WORKSHOPS*, 2011, pp. 74–81. doi: 10.1109/CVPRW.2011.5981881.

[10] E. Charoqdouz and H. Hassanpour, "Feature Extraction from Several Angular Faces Using a Deep Learning Based Fusion Technique for Face Recognition," *Int. J. Eng.*, vol. 36, no. 8, pp. 1548–1555, Mar. 2023, doi: 10.5829/IJE.2023.36.08B.14.

[11] I. Masi, Y. Wu, T. Hassner, and P. Natarajan, "Deep Face Recognition: A Survey," in *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, IEEE, Oct. 2018, pp. 471–478. doi: 10.1109/SIBGRAPI.2018.00067.

[12] S. Yang, P. Luo, C. C. Loy, and X. Tang, "WIDER FACE: A Face Detection Benchmark," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 5525–5533. doi: 10.1109/CVPR.2016.596.

[13] A. Angelova, Y. Abu-Mostafa, and P. Perona, "Pruning Training Sets for Learning of Object Categories," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, IEEE, 2018, pp. 494–501. doi: 10.1109/CVPR.2005.283.

[14] F. W. Wheeler, X. Liu, and P. H. Tu, "Multi-Frame Super-Resolution for Face Recognition," in *2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems*, IEEE, Sep. 2007, pp. 1–6. doi: 10.1109/BTAS.2007.4401949.

[15] B. Jeganathan, "Exploring the Power of Generative Adversarial Networks (GANs) for Image Generation: A Case Study on the MNIST Dataset," *Int. J. Adv. Eng. Manag.*, vol. 7, no. 01, 2025.

[16] G. Bhat, M. Danelljan, F. Yu, L. Van Gool, and R. Timofte, "Deep Reparametrization of Multi-Frame Super-Resolution and Denoising," *arXiv*, Aug. 2021.

[17] S. K. Chintagunta, "The Role of Artificial Intelligence in Software Engineering: A Review of Frameworks, and Impact on the Software Development Life Cycle," *IJERET*, vol. 6, no. 4, pp. 72–79, Oct. 2025.

[18] F. S. Samaria and A. C. Harter, "Parameterisation of a stochastic model for human face identification," in *Proceedings of 1994 IEEE Workshop on Applications of Computer Vision*, IEEE Comput. Soc. Press, 1994, pp. 138–142. doi: 10.1109/ACV.1994.341300.

[19] Z. Luo *et al.*, "EBSR: Feature Enhanced Burst Super-Resolution with Deformable Alignment," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2021, pp. 471–478. doi: 10.1109/CVPRW53098.2021.00058.

[20] G. Dedeoglu, T. Kanade, and J. August, "High-zoom video hallucination by exploiting spatio-temporal regularities," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, IEEE, 2021, pp. 151–158. doi: 10.1109/CVPR.2004.1315157.

[21] S. K. Davuluri, V. Challagulla, V. Mudapaka, and U. Konka, "Telcoformrix: An AI-Augmented Framework for Declarative and Scalable Provisioning of Real-Time Communication Infrastructure," in *2025 IEEE International Conference and Expo on Real Time Communications at IIT (RTC)*, 2025, pp. 1–4. doi: 10.1109/RTC66985.2025.11211725.

[22] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, pp. 51–59, 1996.

[23] Kishore Subramanya Hebbar, "AI-Driven Code Review: A Real-Time Feedback System for Secure and Maintainable Software Development," *J. Inf. Syst. Eng. Manag.*, 2024.

[24] K. Lee, "ENHANCING LOW-RESOLUTION FACE RECOGNITION WITH FEATURE SIMILARITY KNOWLEDGE DISTILLATION," *arXiv*, pp. 1–13, 2023.

[25] S. Achouche, U. B. Yalamanchi, and N. Raveendran, "Method, Apparatus, and Computer-Readable Medium for Performing a Data Exchange on a Data Exchange Framework," US 10,387,195 B2, Aug. 2019

[26] X. Gao, K. Zhang, D. Tao, and X. Li, "Image Super-Resolution With Sparse Neighbor Embedding," *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3194–3205, 2012, doi: 10.1109/TIP.2012.2190080.

[27] D. Patel, "AI-Enhanced Natural Language Processing for Improving Web Page Classification Accuracy," vol. 4, no. 1, pp. 133–140, 2024, doi: 10.56472/25832646/JETA-V4I1P119.

[28] R. P. Mahajan, "Optimizing Pneumonia Identification in Chest X-Rays Using Deep Learning Pre-Trained Architecture for Image Reconstruction in Medical Imaging," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 5, no. 1, pp. 52–63, Apr. 2025, doi: 10.48175/ijarsct-24808.

[29] R. P. Mahajan and N. Jain, "Enhancing the Deep Learning-Based Pet Imaging Super-Resolution for Facial Expression Images," in *2025 International Conference on Intelligent and Cloud Computing (ICoICC)*, 2025, pp. 1–6. doi:

10.1109/ICoICC64033.2025.11052125.

[30] R. P. Mahajan and N. Jain, “Early Detection of Breast Cancer through Automated Radiology Image Classification using Machine Learning Model,” in *2025 4th International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, 2025, pp. 1–7. doi: 10.1109/ICDCECE65353.2025.11034849.

[31] F. W. Wheeler, X. Liu, and P. H. Tu, “Multi-Frame Super-Resolution for Face Recognition,” in *2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems*, IEEE, Sep. 2007, pp. 1–6. doi: 10.1109/BTAS.2007.4401949.

[32] E. Ustinova and V. Lempitsky, “Deep multi-frame face super-resolution,” Oct. 2017.

[33] R. Palwe, “Adaptive human: AI decision support for high-stakes financial advice,” *Int. J. Comput. Artif. Intell.*, vol. 6, no. 2, 2025.

[34] D. Capel and A. Zisserman, “Super-resolution from multiple views using learnt image models,” in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, 2001, pp. II–II. doi: 10.1109/CVPR.2001.991022.

[35] V. M. Prajapati, Vikas, “A Review of AR/VR Technologies in Simulation-Based Learning: Current Trends and Future Directions Article Sidebar,” *J. Glob. Res. Electron. Commun.*, vol. 1, no. 1, pp. 1–6, 2025.

[36] C. Korkmaz, A. M. Tekalp, and Z. Dogan, “MMSR: Multiple-Model Learned Image Super-Resolution Benefiting From Class-Specific Image Priors,” Sep. 2022.

[37] V. PAL, “Federated Contrastive Learning for Privacy- Preserving Medical Image Analysis,” vol. 9, no. 1, pp. 601–606, 2022.

[38] S. Moschoglou, A. Papaioannou, C. Sagonas, J. Deng, I. Kotsia, and S. Zafeiriou, “AgeDB: The First Manually Collected, In-the-Wild Age Database,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017, pp. 1997–2005. doi: 10.1109/CVPRW.2017.250.

[39] J. Kim, G. Li, I. Yun, C. Jung, and J. Kim, “Edge and identity preserving network for face super-resolution,” *Neurocomputing*, vol. 446, no. 24, pp. 11–22, Jul. 2021, doi: Machine learning for numerical weather and climate modelling: a review.

[40] Z. Zhang, Y. Song, and H. Qi, “Age Progression/Regression by Conditional Adversarial Autoencoder,” Mar. 2017.

[41] H. Nada, V. A. Sindagi, H. Zhang, and V. M. Patel, “Pushing the Limits of Unconstrained Face Detection: a Challenge Dataset and Baseline Results,” Aug. 2018.

[42] Z. Cheng, X. Zhu, and S. Gong, “Surveillance Face Recognition Challenge,” Aug. 2018.

[43] D. Yi, Z. Lei, S. Liao, and S. Z. Li, “Learning Face Representation from Scratch,” *arXiv*, 2014.

[44] R. Moghekar and S. Ahuja, “Impact of applying super resolution to low resolution face images on the performance of deep neural networks,” *J. Phys. Conf. Ser.*, vol. 1950, no. 1, p. 012050, Aug. 2021, doi: 10.1088/1742-6596/1950/1/012050.

[45] C. Patel, “Customer Experience Optimization Using Machine Learning : A Systematic Review,” *ESP J. Eng. Technol. Adv.*, vol. 3, no. 4, pp. 176–187, 2023, doi: 10.56472/25832646/JETA-V3I8P120.

[46] A. Sapkota and T. E. Boult, “Large scale unconstrained open set face database,” in *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2013, pp. 1–8. doi: 10.1109/BTAS.2013.6712756.

[47] M. Grgic, K. Delac, and S. Grgic, “SCface – surveillance cameras face database,” *Multimed. Tools Appl.*, vol. 51, no. 3, pp. 863–879, Feb. 2011, doi: 10.1007/s11042-009-0417-2.

[48] N. D. Kalka *et al.*, “IJB–S: IARPA Janus Surveillance Video Benchmark,” in *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, 2018, pp. 1–9. doi: 10.1109/BTAS.2018.8698584.

[49] S. Thangavel, “Precision Agriculture Robot with Image Processing,” *2024*.

[50] Dan Zeng, Hu Chen, and Qijun Zhao, “Towards resolution invariant face recognition in uncontrolled scenarios,” in *2016 International Conference on Biometrics (ICB)*, IEEE, Jun. 2016, pp. 1–8. doi: 10.1109/ICB.2016.7550087.

[51] P. J. Phillips *et al.*, “Overview of the Multiple Biometrics Grand Challenge,” 2009, pp. 705–714. doi: 10.1007/978-3-642-01793-3_72.

[52] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, “A convolutional neural network cascade for face detection,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2015, pp. 5325–5334. doi: 10.1109/CVPR.2015.7299170.

[53] Y. Liu, R. Liu, S. Wang, D. Yan, B. Peng, and T. Zhang, “Video Face Detection Based on Improved SSD Model and Target Tracking Algorithm,” *J. Web Eng.*, pp. 21–37, Jan. 2022, doi: 10.13052/jwe1540-9589.21218.

[54] I. Kemelmacher-Shlizerman, S. Seitz, D. Miller, and E. Brossard, “The MegaFace Benchmark: 1 Million Faces for Recognition at Scale,” *arXiv Comput. Sci.*, Dec. 2015.

[55] Z. Liu, P. Luo, X. Wang, and X. Tang, “Deep Learning Face Attributes in the Wild,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Dec. 2015, pp. 3730–3738. doi: 10.1109/ICCV.2015.425.

[56] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, “Automatic Face & Gesture Recognition,” in *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition*, IEEE, Sep. 2008, pp. 1–8. doi: 10.1109/AFGR.2008.4813399.

[57] P. J. Phillips, P. Grother, R. Micheals, D. M. Blackburn, E. Tabassi, and M. Bone, “Face recognition vendor test 2002,” in *2003 IEEE International SOI Conference. Proceedings (Cat. No.03CH37443)*, IEEE Comput. Soc, p. 44. doi: 10.1109/AMFG.2003.1240822.

[58] H. Qin, J. Yan, X. Li, and X. Hu, “Joint Training of Cascaded CNN for Face Detection,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 3456–3465. doi: 10.1109/CVPR.2016.376.

[59] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[60] H. Yan, X. Wang, Y. Liu, Y. Zhang, and H. Li, “A new face detection method based on Faster RCNN,” *J. Phys. Conf. Ser.*, vol. 1754, no. 1, p. 012209, Feb. 2021, doi: 10.1088/1742-6596/1754/1/012209.

[61] B. Ye, Y. Shi, H. Li, L. Li, and S. Tong, “Face SSD: A Real-time Face Detector based on SSD,” in *2021 40th Chinese Control Conference (CCC)*, IEEE, Jul. 2021, pp. 8445–8450. doi: 10.23919/CCC52363.2021.9550294.

[62] H. Aung, A. V. Bobkov, and N. L. Tun, “Face Detection in Real Time Live Video Using Yolo Algorithm Based on Vgg16 Convolutional Neural Network,” in *2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, IEEE, May 2021, pp. 697–702. doi: 10.1109/ICIEAM51226.2021.9446291.

[63] Q. Xu, Z. Zhu, H. Ge, Z. Zhang, and X. Zang, “Effective Face Detector Based on YOLOv5 and Superresolution Reconstruction,” *Comput. Math. Methods Med.*, vol. 2021, pp. 1–9, Nov. 2021, doi: 10.1155/2021/7748350.

[64] J. Li *et al.*, “DSFD: Dual Shot Face Detector,” *IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Apr. 2019.

