IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON INTELLIGENT BIOMETRIC SYSTEMS FOR SECURE SOCIETIES

Received June 13, 2020, accepted June 22, 2020, date of publication June 30, 2020, date of current version July 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3006100

A Granular Level Feature Extraction Approach to Construct HR Image for Forensic Biometrics Using Small Training DataSet

KHALID SAEED[1], (Senior Member, IEEE), SOMA DATTA[2], AND NABENDU CHAKI[2], (Senior Member, IEEE)

[1]Faculty of Computer Science, Bialystok University of Technology, 15-351 Bialystok, Poland
[2]Department of Computer Science and Engineering, University of Calcutta, Kolkata 7000106, India

Corresponding author: Nabendu Chaki (nabendu@ieee.org)

This work was supported in part by the Bialystok University of Technology, Poland and in part by the Ministry of Science and Higher Education in Poland. During the progress of the research work, the authors got funding support from Visvesvaraya Ph.D. Scheme, Ministry of Electronics and IT, India.

ABSTRACT In forensic biometrics old x-ray images are often used for identification and verification. The number of homicide cases has increased manyfold along with population growth. Forensic odontology is a less expensive method to solve these cases. In forensic odontology, the old dental x-ray images of a victim are used as ante mortem data to compare the pattern for identification. The success of this post mortem biometrics process totally depends on the brain perception of forensic odontologist. Low resolution (LR) image may create human brain perception error at the time of making decision both in case of disease diagnosis and forensic biometrics. In such context, a software solution could help to reconstruct high resolution (HR) image from LR medical x-ray image. However, methods like convolution neural network (CNN) require high volume of training images to reconstruct HR images. Unfortunately, the available medical x-ray image repository does not offer a large volume of training dataset. This work aims to overcome this data related issue and presents a granular level feature based HR grayscale medical x-ray image reconstruction mechanism from LR image. This method uses machine learning for HR image reconstruction. The proposed granular level feature extraction method generates adequate amount of training set from limited amount of training image sets. These granular level features contain the influence of neighboring points separately and its direction. These features are highly immune to noises and preserve important properties like edges. Polynomial regression model is used for HR value reconstruction. This image reconstruction provides satisfactory results for image database like dental radiographs datasets and the average structural similarity index (SSIM) metric reaches 0.9326.

INDEX TERMS Forensic odontology, dental biometrics, granular level features, polynomial neural network, superresolution, image reconstruction, data sequencing, training data shuffling.

I. INTRODUCTION

The numbers of homicide cases throughout the year are increasing gradually. This alarming situation has been published by UN Office of Drug and Crime (UNODC) in their report ''Global Study on Homicide 2019'' [1]. In general, homicide cases are being reported couple of days/weeks later. Under these circumstances, identification of the dead bodies becomes very difficult because soft tissues based good physiological biometrics techniques (like finger print, iris, face pattern etc.) may not work properly. Therefore, identification process totally depends on hard tissue based forensic odontology or expensive DNA test methods. In forensic odontology dental radiographic images (antemortem) are used to extract features like dental restoration, dental features, teeth morphology, root morphology. These features are compared during post mortem. World Health Organization's (WHO) report [2] on oral health says that 90% of world population suffers from dental caries. Hence, most of the world population has to undergo dental radiography through their lifetime. Thus, availability of antemortem data regarding dental radiography is not an issue. At the time of ante mortem and

The associate editor coordinating the review of this manuscript and approving it for publication was Marina Gavrilova.

post mortem comparison from gray scale dental radiographic images, human brain perception error plays an important role. This perception error increases the difficulty level for particular pattern identification in forensic odontology. Sometimes comparatively high resolution images are preferred to reduce this error in human brain. In these circumstances, a software solution is preferred that will reconstruct High Resolution (HR) gray scale dental radiographic image from Low Resolution (LR) images. Most of the time these LR anti mortem images are taken earlier or obtained from old generation equipments. These reconstructed images should look very similar with actual high resolution images. Based on the implementation technique, high resolution image reconstruction mechanisms are broadly categorized into two classes. The former class is single image super resolution [3] and the other class is example based super resolution. In single image super resolution method high resolution image is reconstructed by estimating values from the corresponding low resolution image. These estimations are linear, nonlinear and distribution based.

Sometimes these methods reduce the aliasing effect, impact of noises etc. However, such methods do not provide satisfactory results when resolution increases beyond 1.6 times [4]. Example based super resolution methods are proposed to overcome these difficulties. Using these methods, a low and corresponding high resolution images are sliced into patches. At the time of training, a corelation mapping is built among the low and corresponding high resolution patches. After completion of training, a low resolution image is given into the system as an input. The system breaks the image into patches and finds the corresponding best matched high resolution patch. After that, these high resolution patches are stitched to get high resolution image. Machine learning provides satisfactory results in high resolution image reconstruction. However, the bottle neck of this method is that it requires huge training sets. Conventional machine learning methods for HR image reconstruction fail due to inadequate training image set. If granular level features are considered, then adequate number of training sets can be generated from limited number of training image datasets.

II. REVIEW ON HR IMAGE RECONSTRUCTION

In this section, we focus on the existing HR image reconstruction and critically analyze the merits and demerits of such methods. Our study reveals a few important gaps in the existing state of the art. These motivate us to identify the specific scope of work and to explore the solution thereafter.

A. STUDY AND ANALYSIS

Based on implementation procedure, HR image reconstruction methods are roughly categorized into four major groups. These are interpolation based, frequency domain based, regularization based and learning based approaches [5]. Each of these approaches has merits and demerits. Interpolation approaches are very straight forward and are practical in reducing aliasing effects significantly in text images. That is why they are suitable for handwritten biometrics. However, they have very high computational cost. Frequency domain based approaches migrate the entire computation into frequency domain. Hence, these methods are not suitable for real time applications. On the other hand, after completion of exhausted learning using predetermined model, the learning based approaches can work on real time basis.

The computational cost during learning is very high. Regularization based approaches are suitable for document images. In such methods, determination of the hyper parameter is very important and arbitrary selection of these parameters can deviate the result. Yang *et al.* [6] proposed a high resolution image reconstruction mechanism where the low resolution image is migrated into frequency domain using Discrete Fourier Transform (DFT). This is to combine it with aliased DFT coefficients and finally to revert back in the spatial domain. Woods *et al.* [7] introduced an interactive expectation maximization procedure to get high resolution image. Image registration, interpolation, deconvolution are used in the expectation maximization procedure. Iterative back projection is also used to reconstruct the HR image [8]. Multiple images of the same object are used in the reconstruction. Translation and rotation are used to describe local motion. In wavelet, an image can be splitted in different scale; this multi-scaling property is used during high resolution image reconstruction [9]. Bicubic, bilinear, neighborhood, cubic spline [10]–[14] are the most common example of interpolation based high resolution image reconstruction methods. In these methods, unknown pixel values within high resolution images are estimated from corresponding low resolution images. This process is smoothing the pixel value changes. Hence, interpolation method reduces the sharpness of the edges within the HR image. Image registration is one of the primary operations used in multi image super resolution method. This is essential to mitigate the alignment issue of the subject within multiple input images. Ur and Gross *et al.* [15] reported a uniform interpolation technique without considering image registration. Reconstruction-based methods mainly focus on the removal of aliasing effects that occur in the low-resolution images due to under-sampling process.

An increasing interest is found in the application of super-resolution to different biometrics modalities like finger prints, face, gait and iris. Fundamental limits of reconstruction-based super resolution algorithms come under local translation. Lin and Shum [16] proposed a technique that derived theoretical limits of reconstruction depended super-resolution methods. The authors have also proved that magnification factors are very low and less than 2. Moreover, these methods are only applicable for video sequences. These need several sequences of low-resolution images as input and fail to restore dynamic non-rigid objects such as faces, finger print. A framework based technique that extends reconstruction-based SR method has been proposed for the single image super-resolution problem [17]. This method is based on some observations. The patches of a natural image inclined to appear many times inside the image. Both of the

images emerge within the same scale and across different scales as well [18]. Research community is increasingly paying more attention to learning based super-resolution methods. Learning based methods recover more texture detail, attain higher magnification factors as well.

Kim *et al.* [19] proposed a method based on deep convolution network motivated by VGG-net. Here the accuracy significantly improves when network depth is increased. They have used 20 weighted layers at their final level. The authors cascaded small filters in a deep network structure. However, here convergence is one the of major critical issues during training time. Datsenko and Elad *et al.* [20] proposed a method that assigned many candidates to high quality patches at every pixel position in the LR image. At the end of their method, it is merged into Maximum A Posteriori probability (MAP) penalty cost function. In learning based approaches [21], the high frequency components from the given sample images are collected. In the next phase, the high resolution image is constructed using these high frequency components. Chang *et al.* [12] reported a method based on multiple nearest neighbours in the training set. Nguyen *et al.* [22], [23] worked upon the poor quality iris image reconstruction. The authors have given the effort upon the quality measures to compute the weights of reconstruction based SR technique.

In paper [19], the authors engaged on the focus level at each frame. The focus level is measured by evaluating the high frequency and total energy of the image. A high-resolution image was reconstructed using the focus-score, weighted average of the available frames. Highly de-focused frames were discarded for further processing. Remaining frames were fused to super-resolve the iris image. The authors have achieved an EER of 2.1 using the MBGC Portal database [24]. In paper [25], the authors combined focus, off-angle, motion blur, and illumination variation factors into a unified quality score for each frame. The authors employed the Depmster-Shafer theory. The novelty of this work [25] is that in spite of using the conventional weighted average of image reconstruction, the frames were fused by an exponential weighted average. Experimental results also showed that optimum number of frames were used for fusion. The work in [25] is not validated on different dataset and different image acquisition. Othman *et al.* [26] and Othman and Dorizzi [27] extended this idea by computing the quality of local image patches. They have estimated a Gaussian Mixture Model (GMM) of clean image texture distribution. At the time of fusion, each pixel was weighted individually. These are associated with the quality value of the corresponding local patch, instead of employing a single quality score for the whole image.

B. MOTIVATION

Different learning methods [28] provide satisfactory results in high resolution image reconstruction. However, these methods consume considerable time the during preparation of training model (creation of pre-trained model). Huge amount of proper training set is the prerequisite criteria of learning methods. The SR methods using machine learning require huge amount of low and corresponding high resolution image pair as a training set. But in reality, this huge number of image pair is not available in medical domain specially in dental radiography [28]–[30]. Hence, conventional ML methods cannot fit directly for high resolution image reconstruction. Moreover, single image super-resolution methods are not able to increase the resolution beyond 1.6 times [4].

C. SCOPE OF THE WORK

Under such circumstances, a research gap is identified towards developing high resolution x-ray image reconstruction that ensures a higher rate of accuracy. It also requires minimum training image set at the same time. The main contribution of this research work includes,

1) A granular level feature extraction mechanism is proposed to reconstruct a high resolution medical x-ray image.
2) The granular level features address both the local influence and its aggregate direction. The influence of neighbor is derived from laws of magnetism property of a ferromagnetic material.
3) A novel training data sequencing method has been proposed to avoid biased learning during training.

III. METHODOLOGY

A learning based high resolution medical x-ray image reconstruction mechanism is proposed in this research work. It is helpful for disease diagnosis and forensic biometrics as well. A granular level feature extraction mechanism has been proposed in this work. This helps to generate adequate amount of training dataset from small number of image pairs with low and corresponding high resolution images. Hence, limited amount of training pairs of a particular type medical images are sufficient to build the pre-trained model. This method works like the same way that an example based image reconstruction method works. The estimation of high resolution pixel value is calculated on the basis of neighboring pixels influence. The block diagram of the system is shown in figure 1. Typically, like most other machine learning applications, the sequence of actions are the training phase, testing phase and then deployment phase. In the first phase, the training image dataset is checked to confirm whether it has image pair of high and corresponding low resolution image. If the image pair is not available then input image is treated as high resolution image. A low resolution image synthesis mechanism is introduced to get corresponding low resolution image. After that high and low frequency features are extracted from low resolution image.

High level and low level features, contain the influence of neighbor pixels and direction of the collective influence. Then ''prepare the training dataset'' module combines these high and low level features and builds a vector of 14 elements. This feature vector is called granular level feature. At the

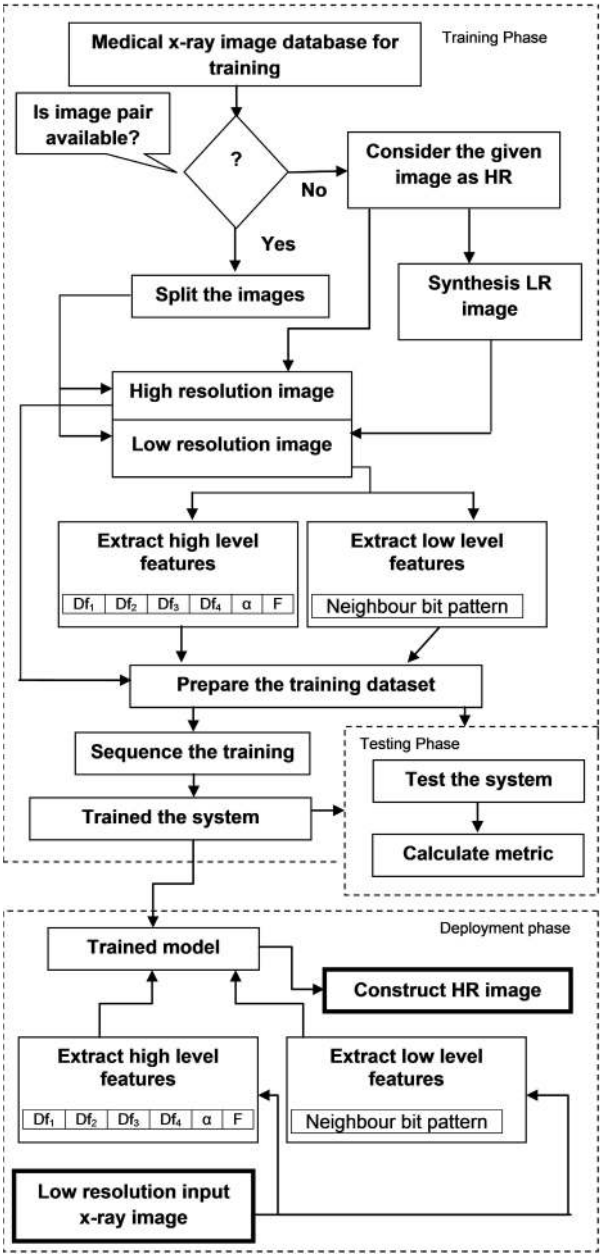


FIGURE 1. Block Diagram of the proposed method.

time of preparing training dataset, a training data sequencing mechanism is introduced to prevent biased training through ''sequence the training'' module. All training data are not used for preparing the trained model. Some portion of this training data are reserved for testing the performance of the system. The testing phase uses this reserved data and calculates the performance through metrics.

In deployment phase the high and low level frequency features of the input x-ray image are extracted. The input x-ray image should be of the same category from which training model is built. Next, these feature sets are passed to the trained model to generate corresponding output value. These output values are mapped as pixel values in the corresponding high resolution image. Some selected theoretical preliminaries are discussed before going into the details of the system. These are used as fundamental building blocks of the system.

A. PRILIMINARIES

In this section, a brief overview of the fundamental imaging in terms of the combination of illumination and reflection is presented. It's properties and correlation are also incorporated in this section. The property of magnetism and its influence on ferromagnetic materials are discussed. This property assists us to understand the collective influence of neighboring pixels. At the end, a robust feature set extraction mechanism is proposed by combining these properties.

1) OPTICAL COMPONENT OF MEDICAL X-RAY IMAGE

Horn [31] has established that the intensity level of any part of an image $I(x,y)$ is a product of illumination component $IL(x,y)$ and reflection component $Re(x,y)$ of that (x,y) point.

$$I(x, y) = IL(x, y)Re(x, y) \tag{1}$$

In the case of normal gray image, illumination component $IL(x,y)$ totally depends on the various source of light and it can be re-written as

$$IL(x, y) = \sum_{i=1}^{n} K_i(x, y)ILS_i \tag{2}$$

Here, n is the number of light sources, ILS_i denotes the illumination level of i^{th} light source. $K_i(x, y)$ is the factor that reduces the illumination of ILS_i at the point (x,y). On the other hand, $Re(x,y)$ depends on the surface of the object. The *"Re"* component contains actual information, like sudden change of value that indicates the presence of edge in the image. On the other hand, *"IL"* component is incident on the object, so it does not contain information regarding the object. In the case of machine vision application, sudden change of *"IL"* component is unusual [17], [32], [33]. However, it happens in the case of natural images, like landscape images. This makes the entire system very complex. This is explained in the next paragraph.

In terms of digital image processing, *"IL"* can be compared as low frequency component and *"Re"* as high frequency component. So it can be concluded that in a digital image the impact of *"Re"* is smoothen by *"IL"*. Hence, if *"IL(x,y)"* eliminated from $I(x,y)$, then $I(x,y)$ contains more detailed information of local neighbourhood. In case of natural images, the value of n in equation 2 tends to infinity. In a sunny day, it is easy to find the shadow of an object placed in an open space. The *"IL"* component of both places which are in shadow and not in shadow are different. These have become almost same when the direct source of *"IL"* is absent. Consider the case of a cloudy day where direct source of *"IL"* is absent. Hence, simple implementation of equation 1 is difficult in most of the natural images. Fortunately, this situation does not arise in medical x-ray imaging.

In the case of medical x-ray imaging, incident x-ray penetrates through the body. However, the penetrated amount of

light differs on the basis of internal structure of the patient's body. This penetrated light is used for x-ray imaging. Hence, instead of reflection component, the penetration component is being considered for x-ray images. On the other hand, in medical x-ray images, only one source of x-ray is used for imaging. Ambient light from natural sources of light does not have the capacity of penetration properly like x-ray. Thus, for x-ray image, equation 2 is rewritten as

$$IL(x, y) = K_1(x, y)ILS_1 \tag{3}$$

The k value is almost constant for all (x,y) coordinates of the x-ray images. Finally, equation 1 is re-defined as

$$I(x, y) = IL(x, y)P_e(x, y) \tag{4}$$

$P_e(x, y)$ denotes the penetrated component of incident x-ray. The variation of the intensity levels in two consecutive points of the x-ray image indicates the change of internal structure of the organ. These x-ray images are used to study the structure of the organ.

B. EXTRACTION OF HIGH LEVEL FEATURES

Equation 4 is transformed into logarithmic domain to replace the multiplicative model with additive model. This will reduce the dilution effect of *"IL"* on *"Pe"* as discussed earlier.

$$\begin{aligned} \log(I) &= \log(IL) + \log(P_e) \\ \log(P_e) &= \log(I) - \log(IL) \end{aligned} \tag{5}$$

The advantage of this model is that Illumination-invariant component *Pe* can be obtained by subtracting Illumination from the x-ray image. It is much easier than division. Let *df* be the difference between two consecutive pixel intensity level in x-ray image. Thus, *df* can be expressed in logarithmic domain as,

$$df = \log(I_N) - \log(I_C) \tag{6}$$

I_C is some source point and I_N is the neighbourhood point.

$$\begin{aligned} df &= \log(IL_N) + \log(Pe_N) - \log(IL_C) - \log(Pe_C) \\ df &= \log(Pe_N) - \log(Pe_C) \end{aligned}$$

as $IL_N = IL_C$ (as discussed earlier, x-ray illumination is almost constant for all points)

$$df = \log(Pe_N/Pe_C) \tag{7}$$

Equation 7 provides a rate of change calculation method for intensity value without considering illumination component of a point *"C"*. Experimental results of the literature [28], [34]–[37] justify this approach. Equation 7 sometimes provides undefined result when denominators tend to zero. This issue can be fixed by adding 1 to all points of the image matrix. Here, *"df"* is an important feature that can be used for local pattern prediction. This feature can be used to estimate the direction of the influence of a pixel from its neighbour pixels. We infer from equation 7 that df finds the change of *"Pe"* component with respect to a point *"C"*. But *"C"* point is surrounded by other points. Hence, a set of *"df"* is generated for point *"C"*. Four neighbor points are considered for this application. These points are marked as *N1, N2, N3, N4* in figure 2.

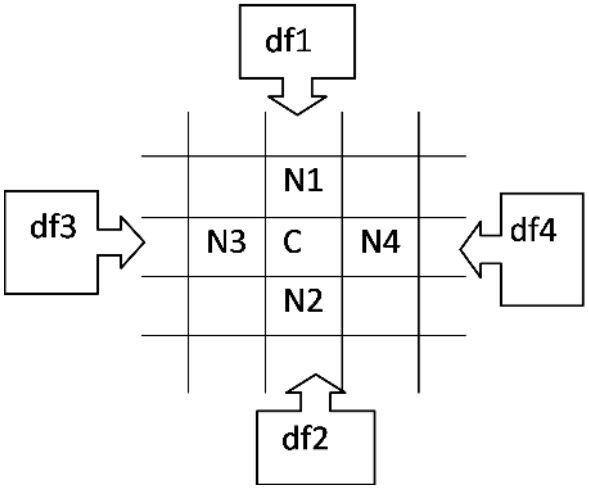


FIGURE 2. Neighbour points and corresponding df.

1) MAGNETISM PROPERTIES

In super resolution, each pixel in low resolution image is mapped into a distinct set of $n \times n$ pixels in high resolution image. Here $n \geq 1$ and n is an integer. If all elements in the derived set of $n \times n$ pixels have the same value, then there is no practical difference with LR image and the corresponding derived HR image. The objective of super resolution is to estimate values for all points in $n \times n$ from derived pixels in such a way that the derived HR image is quite similar with actual HR image. During this estimation, influence from neighbour pixels is important. Analytical observation shows that the nearest pixels have more influence than furthest pixels. This phenomenon is quite similar to the magnetism property. The effect of magnetism over a magnetic particle is inversely proportional to the distance between these points and mathematically written as,

$$f = k * \frac{M_p}{d^2} \tag{8}$$

Here, M_p denotes the conjugate power of magnetism of two particles and d denotes the distance. Here, *"f"* measures the force upon the particle and *"k"* is a constant. Again, mass of a particle and the energy are reversible. Pixel intensity is a form of measurement of energy. Hence, equation 8 can be used to study the influences of neighbouring pixels upon a pixel along with the direction of the influence.

Only 8 neighbour pixels are considered for this application. These pixels are marked as $P_1, P_2, \ldots . P_8$. This is shown in figure 3. The corresponding influences are towards the central pixel *("Ce")* are denoted as $FP_1, FP_2, \ldots . FP_8$. Let us consider Fx_{axis} is the force towards x-Axis and Fy_{axis} is the force towards y-Axis, as shown in figure 4. Then cumulative

P_4(x-1,y-1)	P_3(x-1,y)	P_2(x-1,y+1)
P_5(x,y-1)	C_e(x,y)	P_1(x,y+1)
P_6(x+1,y-1)	P_7(x+1,y)	P_8(x+1,y+1))

FIGURE 3. Centre pixel and its neighbor pixels.

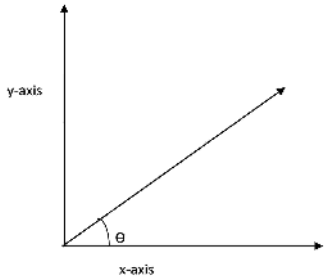


FIGURE 4. Vector division of force FP_i.

force towards *x and y* axis are defined as follows.

$$Fx_{axis} = \sum_{i=1}^{8} FP_i \times \cos(\theta P_i) \tag{9}$$

$$Fy_{axis} = \sum_{i=1}^{8} FP_i \times \sin(\theta P_i) \tag{10}$$

(θP_i) is the angle between FP_i and Fx_{axis}. These two equations are re-written with the help of equation 8 as follows.

$$Fx_{axis} = \sum_{i=1}^{8} \left(\frac{Mp_i}{{d_i}^2}\right) \times \cos(\theta P_i) \tag{11}$$

$$Fy_{axis} = \sum_{i=1}^{8} \left(\frac{Mp_i}{{d_i}^2}\right) \times \sin(\theta P_i) \tag{12}$$

The value of d_i is either *1* or $\sqrt{2}$. Value of d_i is *1* for the points P_1, P_2, P_5 and P_7. The distance is $\sqrt{2}$ for the corner points and are calculated by applying the theorem of Pythagoras. The total amplitude of the neighbour pixel's influence is denoted as F and is defined as

$$F = \sqrt{{Fx_{axis}}^2 + {Fy_{axis}}^2} \tag{13}$$

And the direction is

$$\alpha = \tan^{-1}\left(\frac{Fy_{axis}}{Fx_{axis}}\right) \tag{14}$$

F and α have important roles to estimate the pixel value during reconstruction of high resolution images. The value of α is between *0* to 2Π. F and α can be treated as features. All of these features like *F, α and "df"* are high frequency features. These contain the minute variation details. However, low frequency features are also important to construct the base line value of a pixel on which high frequency features will be applicable.

C. EXTRACTION OF LOW LEVEL FEATURES

Surrounding row pixel values are considered for low level features. The centre pixel is *"Ce"* and the neighbour pixels

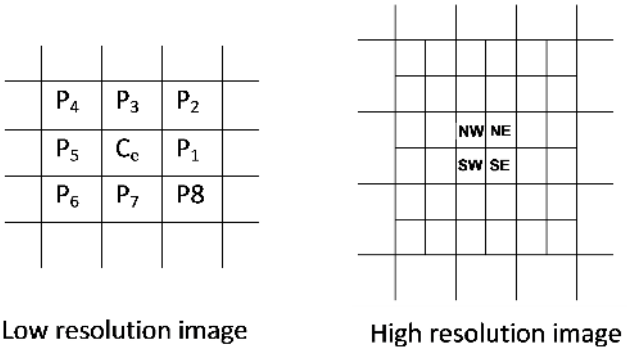


FIGURE 5. Mapping between LR and HR images.

would be $P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8$. In this work, a learning based high resolution image reconstruction system is described. In figure 5, the mapping between LR image and HR image are explained. The centre pixel *"Ce"* is divided into four child pixel in high resolution image. It is marked as *NW, SW, SE and NE*.

The upper left portion of *"Ce"* (which will be mapped with *NW)* pixel should have more influence of P_3 than P_8. Similarly, *SW, SE and NE* should have more influence by P_5, P_7 *and* P_1 respectively. Hence, at the time of preparation of training set, these positions of pixels should be prioritized. The value *(V)* within pattern sequence of low frequency features are restricted within *0* to $(2^8 - 1)$ and these are defined as $V \in \mathbb{Z}$ and $0 \leq V < 2^8$.

During the training set preparation, both high and low resolution images are required. All high frequency and low frequency features are extracted from LR images. Corresponding responses are obtained from the HR images. Upper left portion of *Ce* should have more influence by its nearest neighbour *P3, P4, P5*. This portion is mapped with the HR image's pixel and is marked with *"NW"*. Mapping sequences of *NW, SW, SE and NE* are different from each other due to this reason. The format of training set is the sequence of low frequency features, followed by high frequency features *F, α and "df"*.

P3,	P4,	P5,	P6,	P7,	P8,	P1,	P2,	F,	∝,	Df1	Df2	Df3	Df4	NE
P5,	P6,	P7,	P8,	P1,	P2,	P3,	P4,	F,	∝,	Df1	Df2	Df3	Df4	NW
P7,	P8,	P1,	P2,	P3,	P4,	P5,	P6,	F,	∝,	Df1	Df2	Df3	Df4	SE
P1,	P2,	P3,	P4,	P5,	P6,	P7,	P8,	F,	∝,	Df1	Df2	Df3	Df4	SW
Low Frequency Feature								High Frequency Feature						Desire output
Input feature set														

FIGURE 6. Mapping for low and high frequency features.

Figure 6 shows the training set data format pattern. Low resolution image (say *"Ce")* will produce four training sets for each pixel. In this figure, the first eight columns represent low frequency feature whilst columns 9 to 14 represent high frequency features. Column 15 shows the desired output obtained from the corresponding high resolution image. The *NE, NW, SE, and SW* each comprises of two categories of features - high frequency features and low frequency features.

The high frequency features are the same for *NW, SE, and SW*. However, low frequency features are different for *NW, SE, and SW*. Hence, after combination of these high and low frequency features, the *NW, SE, and SW* do not have the same feature value.

IV. PREPARATION OF TRAINING DataSet

Both low resolution and corresponding high resolution medical x-ray image datasets are required to prepare the training dataset. Unfortunately, the medical x-ray image repository hardly has high and corresponding low resolution image pair set. Hence, to meet this criteria for preparing training datasets, low resolution images are synthesized from the available medical x-ray images. In this work, Kaggle [28], [38] x-ray image repository is used. The repository contains more than thousand gray scale x-ray images. However, it does not have low and its corresponding high resolution image pair to prepare training dataset. The available images in the repository are considered as the high resolution images whilst the corresponding low resolution images are synthesized. The simplest way to get this synthesized LR image is the averaging method [39]. It is defined as,

$$LowResolutionI(x,y) = \frac{1}{m^2}\sum_{i=1}^{n}\sum_{j=1}^{n} HighResolutionI((x-1)\times m+i,(y-1)\times m+j) \tag{15}$$

LowResolutionI(x,y) represents a pixel in low resolution image whose row and column are *x,y* respectively. *HighResolutionI(x,y)* represents the high resolution image. The number of rows in *HighResolutionI* is m times greater than the row number in *LowResolutionI*. This is also applicable for column numbers. The synthesized low resolution image by using averaging method differs much from actual low resolution images that are captured by transducer. Thus, to overcome this issue an alternate approach is proposed to create LR image from HR image. In the proposed method, input image matrix, *Img* is converted into two sets of vectors. The former one is a row vector set and other one is a column vector set. One element in row vector is denoted byRow_i and Col_s. *"Img"* is the set of all elements in input HR image. The properties of these are defined as follows,

$$Row_i \in Img, \quad Row_i \neq \phi,\ i \in \mathbb{Z},\ i = 1,2,\ldots,n$$

where n is the number of row in the input image matrix.

$$Row_1 \cup Row_2 \cup Row_3 \cup Row_4 \ldots.. \cup Row_n = Img$$
$$Row_k \cap Row_j = \phi; k \neq j \quad k,j = 1,2,3\ldots n$$

Similarly, column vector properties are as follows.

$$Col_s \in Img, Col_s \neq \phi, s \in \mathbb{Z}; \quad s = 1,2,\ldots.m$$

m is the number of column in the input image matrix.

$$Col_1 \cup Col_2 \cup Col_3, \ldots. \cup Col_m = Img$$
$$Col_p \cap Col_q = \phi; p \neq q; \quad p,q = 1,2,\ldots.m$$

Row_i and Col_s have m and n number of elements respectively. Let us assume Row_i and Col_s are set, then for any value of i and s (within its range) is defined as $Z = Row_i \cap Col_s \neq \phi$ and the result set *(Z)* should have always one and only one element. Let us consider Row_{is} as the s^{th} element in Row_i set. Let us assume another row vector set Row' and its elements, which are also a set of elements, are denoted as Row'_i. The elements in Row'_i are estimated from the corresponding Row_i using interpolation method. This method is highly immune in abrupt change of single value (noise). The number of elements in each Row'_i is $m/2$. The set Row' will be used to construct low resolution image. The elements in Row'_i vector are estimated on the basis of corresponding neighbouring points in Row_i vector and it is defined as

$$Row'_i[k] = \sum_{\substack{p=0 \\ p\neq\frac{npoint}{2}}}^{npoint} Row_i[k\times step - \frac{npoint}{2} + p] \times \prod_{\substack{q=0 \\ q\neq npoint \\ q\neq p}}^{npoint} \frac{k\times step - (k\times step - \frac{npoint}{2} + q)}{(k\times step - \frac{npoint}{2} + p) - (k\times step - \frac{npoint}{2} + q)} \tag{16}$$

After simplification, equation 16 can be rewritten as equation 17.

$$Row'_i[k] = \sum_{\substack{p=0 \\ p\neq\frac{npoint}{2}}}^{npoint} Row_i[k\times step - \frac{npoint}{2} + p] \times \prod_{\substack{q=0 \\ q\neq npoint \\ q\neq p}}^{npoint} \frac{\frac{npoint}{2} - q}{p - q} \tag{17}$$

$k = 1,2,3,\ldots m/2$.

Here, *"npoint"* denotes the number of neighboring points to be considered at the time of estimation of $Row'_i[k]$ th value. *"Step"* is a constant that indicates the reduction of elements in each row. For example if $step = 2$ then, it indicates that the number of rows in HR image is 2 times greater than the number of rows in LR image.

In the same way Col_s*'s* components are calculated using equation 18.

$$Col'_s[k] = \sum_{\substack{p=0 \\ p\neq\frac{npoint}{2}}}^{npoint} Col_1[k\times step - \frac{npoint}{2} + p] \times \prod_{\substack{q=0 \\ q\neq npoint \\ q\neq p}}^{npoint} \frac{\frac{npoint}{2} - q}{p - q} \tag{18}$$

$k = 1,2,3,\ldots n/2$.

Once Row', Col' set of vector is considered as a matrix, these are used to construct the low resolution image. If x and y are the coordinates of a point in low resolution image *LowResolutionI* then the combination formula will be defined as follows:

$$LowResolutionI(x, y) = \frac{1}{2}(Row'(x \times step, y) + Col'(x, y \times step)) \quad (19)$$

Here $x = 1, 2, \ldots, m/2$ and $y = 1, 2, \ldots, n/2$.

This method provides a better result than the averaging method using equation 15. The reason behind is that in averaging method, noise has active participation at the time of calculating LR value. However, the proposed method checks the change of values for trained dataset during calculation. Hence, noise has very less contribution in the proposed low resolution image synthesis method.

The implementation of the above technique is explained in procedure 1. Procedure 1 is to construct a low resolution image from the corresponding high resolution image. This is essential for training dataset preparation.

The synthesized low resolution image using procedure 1 is quite similar with actual low resolution image. This method uses an estimation technique on the basis of neighbouring pixels. Hence, the effect of high entropy is smoothed. This synthesis is much better than the method discussed using equation 15. The computational time of this method is higher than the method for equation 15. However, this is one time computation to prepare the training dataset. This computational overhead is not part of the actual computation that is used for reconstruction of the HR image.

A. LEARNING FROM TRAINING DATASET

A fraction of training dataset is kept aside for testing the performance of learning system. Rest of the training data is used for training purpose. Almost all machine learning systems suffer a common problem that is biased learning. This happens when similar type of output valued training sets are given consecutively at the time of learning. It results in less robust learning. Hence, the ordering of training dataset at the time of training is very important to increase the performance of the system. Training data sequencing method is proposed in the next subsection.

1) REARRANGING THE TRAINING DATASET

Random selection of training dataset is widely used to overcome the bias learning issues. However, random selection is not a good solution in this case. Figure 7A is the input image for which training set is preferred. The distribution of quantity for each training pattern is shown in figure 7B, the distribution is un-uniform.

In the case of random selection the training patterns which are more in numbers are getting selected more frequently. This leads to bias learning. A novel hashing based training data sequencing method has been proposed to overcome this issue. In this method total training set is distributed among n numbers of bins. Mathematically, it is defined as the total training set (T_S) and divided into n numbers of bins *(Bin)*. The i^{th} bin is denoted as Bin_i.

Algorithm 1 Synthesis of LR Image From HR Image for Preparation of Training Dataset

Input: I input image matrix with m rows and n columns.
Output: Low resolution image matrix I_{LR}
Assumptions: $m/2 \in \mathbb{Z}$ and $n/2 \in \mathbb{Z}$
Step

1: Initialize two matix R'and C' with dimension of $m \times (n/2)$ and $(m/2) \times n$ respectively.
Subset_length = 7, step = 2
2: Repeat step 3 for all $i_1 = 1, 2, 3, \ldots\ldots\ldots\ldots m$
3: Repeat steps 4 to 9 for all $C_2 = 1, 2, \ldots\ldots n/step$
4: If $(C_2 \times step - \frac{subset\_length}{2}) \leq 0$ Then
$sub\_v\_x = [1, 2, \ldots\ldots, C_2 \times step - 1, C_2 \times step + 1, \ldots\ldots\ldots\ldots, C_2 \times step + \frac{subset\_length}{2}]$
$sub\_v\_y = I[i_1,$ values of $sub\_v\_x]$
Else
$sub\_v\_x = [C_2 \times step - \frac{subset\_length}{2}, \ldots\ldots\ldots, C_2 \times step - 1, C_2 \times step + 1, \ldots\ldots\ldots, C_2 \times step + \frac{subset\_length}{2}]$
$sub\_v\_y = I[i_1,$ values of $sub\_v\_x]$
End If
5: Repeat steps 6 to 8 for all elements in sub_v_y and index are denoted by i_2 $i_2 = 1, 2, \ldots\ldots(subset\_length - 1)$
6: Repeat step 7 for all elements in sub_v_x and index are denoted as j_2 $j_2 = 1, 2, \ldots\ldots\ldots(subset\_length - 1)$
7: $temp_1 = \frac{C_2 \times step - sub\_v\_x[j_2]}{sub\_v\_x[i_2] - sub\_v\_x[j_2]} \times temp_1$
$\forall i_2 \neq j_2$
8: $sum = (temp1 \times sub\_v\_y[i_2]) + sum$
9: $R'[i_1, c_2] = sum$
Reset sum
10: Repeat step 11 for all k = 1, 2,....n
11: Repeat steps 12 to 17 for all l = 1, 2,....$m/2$.
12: If $(l - \frac{subset\_length}{2}) \leq 0$ Then
$sub\_v\_x = [1, 2, \ldots\ldots, (l \times step - 1), (l \times step + 1), \ldots\ldots, (l \times step + \frac{subset\_length}{2})]$
$sub_v_y = I[$values of $sub\_v\_y, k]$
Else
$sub\_v\_x = [(l \times step - \frac{subset\_length}{2}), \ldots\ldots, (l \times step - 1), (l \times step + 1), \ldots\ldots, (l \times step + \frac{subset\_length}{2})]$
$sub_v_y = I[$values of $sub\_v\_x, k]$
End If
13: Repeat steps 14 to 16 for all elements in sub_v_y and indexes are denoted by $i_2 = 1, 2, \ldots\ldots, (subset\_length - 1)$
14: Repeat step 15 for all elements in sub_v_x and indexes are denoted by $j_2 = 1, 2, \ldots\ldots\ldots\ldots, (subset\_length - 1)$
15: $temp_2 = temp_2 \times \frac{l \times step - sub\_v\_x[j_2]}{sub\_v\_x[i_2] - sub\_v\_x[j_2]}$
$\forall i_2 \neq j_2$
16: $sum = temp_2 \times sub\_v\_y[i_2] + sum$
17: $C'[l, k] = sum$
Reset sum

Algorithm 1 *(Continued).* Synthesis of LR Image From HR Image for Preparation of Training Dataset

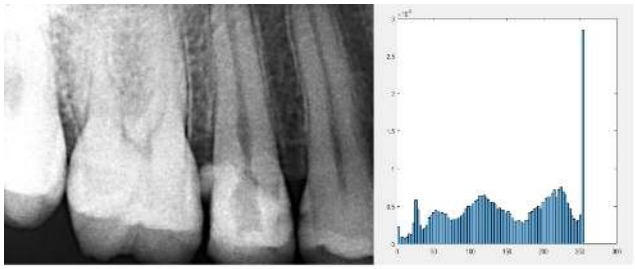
Repeat step 19 for all $x = 1, 2, ...m/2$

Repeat step 20 for all $y = 1, 2, ..., n/2$

$I_{LR}[x, y] = \frac{1}{2}(R'[x \times step, y] + C'[x + y \times step])$

Return low resolution image I_{LR}

Stop.



A B

FIGURE 7. Input image and corresponding histogram.

$Bin_i \in T_S$ *for all i=1,2,….,n;* $i \in \mathbb{Z}$, $Bin_1 \cup Bin_2 \cup Bin_3 \ldots\ldots \cup Bin_n = T_S$ $Bin_i \cap Bin_j = \phi$

Each Bin has a bin representative value denoted as *BinR*. Bin representative is used to calculate the similarity with corresponding bin. The k^{th} element of the training set T_S is denoted as T_k. Hence, $T_k \in T_S$ and $T_k(Y)$ is the output element of T_k.

$T_k \in Bin_i$ for which $\|T_k(y) - BinR_i\|$ is minimum and *i=1,2,….,n.* After distribution of all training data among *n* numbers of bins, the training data sequencing starts. This hashing based training data sequencing method reduces the chance of consecutive training datasets having similar type of output values. The number of bins should be greater than or equal to the total number of peaks in the distribution of output training datasets. Initially, *BinR* contains the peak values of the distribution. Change in derivatives is used to detect the peaks. This data sequencing mechanism is explained with the help of procedure 2. In this procedure, un-sequenced training dataset is given as input and it sequences the given training sets. The sequencing process is organized in such a way that average difference between output value $T_k(y)$ of two consecutive training data should be quite high.

In Procedure 2, steps 1 to 5 are used to define *"Bins"*. Steps 6 to 9 are used to categorize the training dataset as per the *"Bin"* property. Steps 10 to 14 describe the sequencing mechanism for which the average entropy of the output value of two consecutive training sets should be quite high. It is clear from figure 5, that some bins may have a lot of training set with similar type of output. This unbalanced dataset may bias the training procedure. However, steps 11 to 15 have mechanism to avoid such situation. This section discards additional training datasets.

Algorithm 2 Arrange the Sequence of Training Data to Reduce Biased Learning

Input: The set of training data along with count of dataset say R.

Output: Proper sequencing the training data for unbiased learning.

Assumptions: The training dataset must be complete; there is no missing data in any training set.

Step

1: Apply histogram on the output element of each training set; it is denoted as $T_k(Y)$. The result is stored in a vector V

2: Find the number of peaks and peak values in V by examining the change of derivative direction

3: Store the number of peak in "m" and peak values in "*peak_vector*"

4: If $m\%2 == 1$ Then
$n = m - 1$
Eliminate the least element from "*peak_vector*"
End If

5: For each element of "*peak_vector*" find the value of $T_k(Y)$ and store the result in BinRepresentative vector; Number of elements in "BinRepresentative" vector is equal to "n"

6: Repeat steps 7 to 9 for all elements in input training set $K = 1, 2, ..., R$

7: Repeats step 8 for all $j = 1, 2, ..., n$

8: $Distance(j) = ||T_k(Y) - BinRepresentative(j)||^2$

9: Find the index, i of the smallest value from distance vector and categorize the k^{th} dataset T_k as a member of that index Bin say Bin_i

10: count = 0

11: Repeat step 12 to 15 until (n-2) bins are empty

12: $BinIndex_1 = count\%(n/2) + 1$
$BinIndex_2 = BinIndex_1 + (n/2)$

13: If $Bin[BinIndex_1] \neq Empty$ Then
Fetch any arbitrary training set from $Bin(BinIndex_1)$ and sequence that training dataset
Else
Fetch any arbitrary training set from $Bin(BinIndex_1 + 1)$ and sequence the training dataset
End If

14: If $Bin[BinIndex_2] \neq Empty$ Then
Fetch any arbitrary training set from $Bin(BinIndex_2)$ and sequence that training dataset
Else If $(BinIndex_2 + 1) \leq n$ Then
Fetch any arbitrary training set from $Bin(BinIndex_2 + 1)$ and sequence the training dataset
End If

15: count = count + 1

16: Return the final sequence of training dataset

17: Stop.

2) LEARNING METHOD

Individual training data is a vector of 15 elements out of which 14 are the granular level features. These may be called as input and the last one is the desired output for the combination of the input values. As per the convention the output value must be an integer and its range varies between 0 and 255. The response from the learning system is denoted as ''the output''. It is treated as a continuous function in 14 dimensional feature plane. Polynomial neural network using group method of data handling (GMDH) [40], [41] is used to construct the pre-trained model from training dataset. GMDH model is suitable for single output with multiple inputs. This property satisfies the criteria for the application. Moreover, GMDH optimised the internal parameters that is beneficiary from implementation point of view. GMDH follows the Volterra Kolmogorov Gabor (VKG) polynomial at the time of constructing the polynomial equation internally. This complex polynomial looks like equation 20.

$$Y = A_0 + \sum_{i=1}^{M} A_i x^i + \sum_{i=1}^{M}\sum_{j=1}^{M} A_{ij} x^i x^j + \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{M} A_{ijk} x^i x^j x^k + \ldots\ldots \quad (20)$$

Here, M denotes the number of independent variable which is the same as the number of input features. $X_i(i = 1, 2, \ldots .M)$ are the independent input features. Y is the output response.$A_0, A_i, A_{ij} \ldots .$ are the constraints. It can be said that training set contains the value of independent input variables $X_i(i = 1, 2, \ldots .M)$ and corresponding response (Y). The values of constants $A_0, A_i, A_{ij} \ldots .$ are unknown. If adequate training set is available then, values of $A_0, A_i, A_{ij} \ldots .$ can be determined. Once these values are available then it is easy to determine the response (Y) for given input values $X_i(i = 1, 2, \ldots M)$. Higher order polynomial does not necessarily mean more accurate result. Sometimes higher order polynomial leads to ill condition. GMDH follows layer wise training and during training the neurons have good prediction power that are fed forwarded to next layer. The other neurons with less prediction power are discarded. The number of neurons should be provided during parameter setting. The default value of neuron is the same value as the number of input variables.

V. PERFORMANCE EVALUATION

This research work can be considered as an assembly of three blocks or sub-systems on the basis of implementation. These blocks are as follows:

1) Synthesis of low resolution image from high resolution image.
2) Learning from training dataset.
3) Reconstruction of high resolution image from low resolution input image.

TABLE 1. Comparison of PSNR for different methods.

Name of the sample image	PSNR of averaging method Denoted as Series1	PSNR value of proposed method Denoted as Series2	PSNR of the wavelet based method Denoted as Series3
0001.png	22.36	31.78	31.52
0006.png	28.98	32.89	31.92
0014.png	24.23	30.82	30.79
0020.png	21.98	32.89	27.98
0025.png	26.70	33.55	33.62
0027.png	23.52	34.59	30.01
0031.png	27.24	31.19	30.29
0036.png	20.89	34.87	33.82
0040.png	21.24	33.45	30.67
0044.png	26.20	34.28	32.92
0046.png	25.89	35.24	34.89
0048.png	23.67	34.56	36.78
0051.png	27.56	37.67	36.43
0054.png	28.34	38.5	37.45
0055.png	26.76	35.67	37.56
0057.png	23.45	35.5	33.78
0060.png	22.9	33.45	32.13
0062.png	25.45	36.78	35.67
0066.png	23.88	35.67	33.87
0070.png	25.78	39.65	37.65
Average PSNR	24.851	34.65	33.4875

The result section is also organized in that way. The performance of each block is evaluated separately into subsequent subsections.

A. PERFORMANCE OF LR IMAGE SYNTHESIS

The DIV2K image repository [42]–[44] is used to test the performance of the proposed subsystem. This repository contains normal high resolution images and corresponding low resolution images that are obtained from a physical transducer. In this dataset, ''Validation Data Track 1 x4'' is used as low resolution image set and corresponding high resolution images are in set ''Validation Data Track 1 x2''. The output from this subsystem is compared with other methods like averaging method, wavelet based method. The synthesized output images using the proposed method, are quite similar to the original low resolution images and this similarity is measured by PSNR. Table 1 shows the comparison of different low resolution images synthesized by different methods with actual ground truth low resolution images.

The average PSNR for the ''averaging method'' is 24.334. This value improves in wavelet based method and it is 31.354. However, the PSNR of the proposed low resolution image synthesis subsystem is 33.031. It is comparatively better than the other methods. Out of these ten samples, images like *''0020.png'', ''0027.png'', and ''0040.png''* have considerable amount of noises. The proposed subsystem is able to handle these noises much better as compared to the other two methods. Besides, the wavelet method takes considerable amount of time to compute.

B. PERFORMANCE OF THE TRAINING PROCEDURE

Dental x-ray image database [29], Kaggle [28] database are used in this section for performance benchmarking. These

TABLE 2. Performance of learning system.

Combination of parameters	Hidden Layers in Neural net-works	Perceptron in each hidden layer	Epoch	Score without sequenc-ing (without using procedure 2)	Score with training dataset sequenc-ing and after using procedure 2
1	2	5,5	5	0.9312	0.9354
2	2	5,5	10	0.9621	0.9620
3	2	5,5	20	0.9762	0.9869
4	2	5,5	30	0.9893	0.9894
5	2	5,5	50	0.9872	0.9931
6	2	10,10	5	0.9579	0.9605
7	2	10,10	10	0.9802	0.9860
8	2	10,10	20	0.9825	0.9895
9	2	10,10	30	0.9884	0.9909
10	2	10,10	50	0.9904	0.9938
11	3	5,5,5	5	0.9727	0.9807
12	3	5,5,5	10	0.9793	0.9846
13	3	5,5,5	20	0.9815	0.9876
14	3	5,5,5	50	0.9892	0.9934
15	3	5,10,5	10	0.9480	0.9483
16	3	5,10,5	20	0.9723	0.9779
17	3	5,10,5	50	0.9802	0.9849
18	3	10,10,10	50	0.9879	0.9932
19	3	10,10,10	200	0.9909	0.9941

databases have only high resolution images. Hence, images from these databases are passed through the proposed LR image synthesis method to obtain corresponding low resolution images. These low resolution images are used to extract low and high frequency feature set of the training data. The corresponding high resolution image is used to obtain output data of those training sets. The number of pixels in high resolution and corresponding low resolution images are 748×512 *and* 374×356 respectively. The number of valid training dataset for each pair of images are considerably large. All these datasets are used for the training and testing purpose. Table 2 shows the performance of the learning system with respect to different values of parameters. At the time of this performance evaluation, the number of training set is kept constant and its value is 400000. Huge computation time would be needed at the time of high value for number of epochs, number of neurons in each hidden layer. The last two columns show the score of the ''polynomial neural network'' model with and without data sequencing mechanism as proposed in section IV.A. The value of the score is between 0 and 1. A higher value denotes higher accuracy. The experimental result shows that the score improves when training data are sequenced using procedure 2. Performance increases when epoch increases althouogh it gets almost saturated after certain value.

Figure 8 shows the graphical representation of score with respect to a chronological increment of an epoch. The blue line represents row 1 to 5 of Table 2. Here the number of the hidden layers are 2 and the number of perceptrons in each layer is 5. Orange line represents row 6 to row 10. In this case, the number of the hidden layers are 2. It is the same as the previous one. However, the number of perceptrons in each layer is increased to 10. This configuration provides a better score than the blue line with low epoch value.

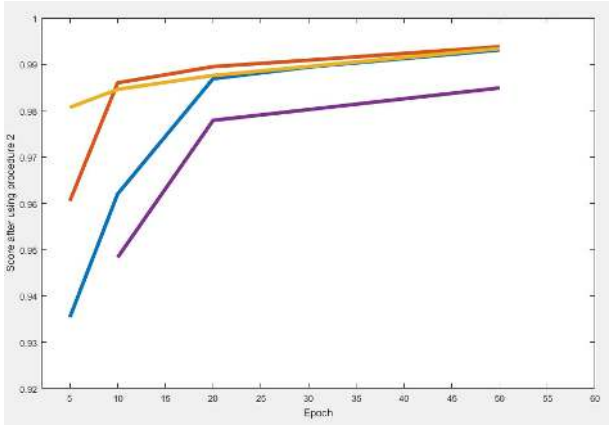


FIGURE 8. Score evaluation against increasing epoch.

The yellow line represents the performance of row 11 to row 14. This configuration uses 3 hidden layers with 5 perceptrons in each layer. This configuration provides a high score when the epoch is low. However, performance does not improve much when the epoch value is increased. The violet line represents row 15 to row 17. As mentioned earlier, an arbitrary selection of more number of perceptrons in a hidden layer may lead to the ill condition. The violet line shows that the previous statement is correct. In this configuration, the number of neurons in a particular hidden layer is increased to 10. This leads to a drop in performance. It is clear that when the epoch is 50 then the score for most of the configuration is almost the same. However, the total computational time is very high for epoch value 50. Considering all of these, the preferred parameter configuration for further processing is set as follows. The number of the hidden layers are 2 whilst the number of neurons in each layer is 10 and the epoch value is 20. The supplementary materials are uploaded in [45]. After selection of preferred learning parameters, the performance of the learning system is evaluated with respect to different volume of dataset. The performance is shown in figure 9. In this evaluation, the number of hidden layers are 2, neurons in each layer is 10 and epoch is 20. These values remain constant during evaluation. Initially, the score increases proportionally with the volume of data. Performance decreases when the volume of data is further increased. After the performance degrades to some extent, it again increases. The fluctuation of performance is not very major. Hence, it is not exactly the example of over fitting. Fifty images from Dental x-ray image database [29] are used to generate training dataset; out of which 400000 dataset are selected by the procedure 2 to build the training model.

C. PERFORMANCE OF THE HR RECONSTRUCTION

The Kaggle [28] image repository is used both for the training system and to evaluate the performance of the proposed system. The synthesized low resolution image is given as input. The reconstructed high resolution image is compared with the original high resolution image of the repository.

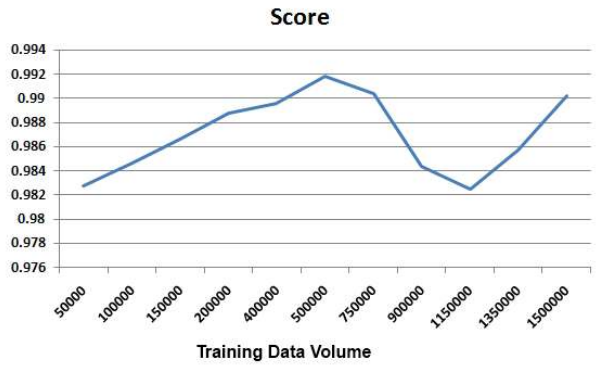


FIGURE 9. Performance evaluation against data volume.

Some standard metrics are used to compare the performance with other methods. These standard metrics [46]–[48] are Peak Signal to Noise Ratio (PSNR), Mean Squared Error(MSE), Root Mean Squared Error (RMSE), Structural Similarity Index (SSIM), Feature Similarity Index (FSIM), Smoothness, Normalized Cross Correlation (NCC) and Average Difference (AD).

MSE is calculated as,

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}||f(i,j) - g(i,j)||^2 \tag{21}$$

Here, PSNR is calculated as

$$PSNR = 20 * \log_{10}\frac{MAX_f}{\sqrt{MSE}} \tag{22}$$

f represents the original image, g represents the degraded image, m denotes numbers of rows of the images, i is the index of that row, n represents the number of columns of the image and j represents the index of that column. MAX_f denotes maximum value that exists in ground truth image.

SSIM helps compare the local patterns of pixel intensities between the ground truth image and the output images. The range varies between -1 and 1.The value closer to 1 indicates the output image and ground truth images are similar. SSIM measures the perceptual dissimilarity between two similar images. It is calculated as,

$$SSIM = \frac{(2 * \mu_x\mu_y + c_1) * (2\sigma_{xy} + c_2)}{(\mu_x{}^2\mu_y{}^2 + c_1)(\sigma_x{}^2 + \sigma_y{}^2 + c_2)} \tag{23}$$

Here, μ_x, μ_y are the average of x *and* y. σ_x *and* σ_y are the variance of x *and* y, σ_{xy} is covariance of x,y. c_1 *and* c_2 are two standard variable. RMSE finds the average magnitude of error. This metric takes the difference between ground truth image and output image and directly finds the variation in the pixel values. This metric is calculated as,

$$RMSE = \sqrt{\frac{1}{n} * \sum_{i=1}^{n}(Y_i - \widehat{Y}_i^2)} \tag{24}$$

n is the data points on all variables, Y_i is the vector of the observed values and Y_i' being the predicted values.

TABLE 3. Comparative performance for different methods.

Name of the Methods	PSNR	SSIM	Smoothness	FSIM
Aliasing Method	26.37	.7950	.5293	.8025
Yaung2013	24.29	.8350	.5832	.7500
Dong 2011	25.27	.8260	.3257	.6352
Bicubic	18.25	.7634	.2589	.7050
Iterative Back Projection (IBP)	23.78	.8112	.3457	.7318
Proposed	30.56	.9326	.6040	.8956

Normalized cross correlation measures the correlation between input image and newly constructed output image. It is calculated as,

$$NCC = \frac{\sum_{j=1}^{M}\sum_{k=1}^{N} x_{(j,k)}x'_{(j,k)}}{\sum_{j=1}^{M}\sum_{k=1}^{N} x_{(j,k)}x^2_{(j,k)}} \tag{25}$$

Average difference is the mean difference between the ground truth image and reconstructed output image. It is calculated as follows,

$$Avg_D = \frac{\sum_{j=1}^{M}\sum_{k=1}^{N} x_{(j,k)}x'_{(j,k)}}{MN} \tag{26}$$

Here $x_{j,k}$ are pixel values of the original image and $x'_{i,j}$ are the pixel values of the image after reconstruction. M,N are the number of columns and rows respectively.

The correlation between input image and reconstructed output image is normalized as this is the ratio of ''the cross-correlation'' and ''the summation of squares of the pixel values'' of the original image. Table 3 and 4 show the comparative performance study between different methods.

TABLE 4. Comparative study for different methods.

Name of the Methods	Value of NCC	Value of Avg_D	RMSE
Aliasing Method	.8158	.245	.2340
Yaung2013	.7829	.305	.1950
Dong 2011	.8343	.145	.2251
Bicubic	.7425	.348	.3532
Iterative Back Projection (IBP)	.8534	.102	.1892
Proposed	.9189	.098	.0967

We have compared the proposed method with five existing popular methods. These methods are aliasing based method [4], Yang proposed method [49], Dong proposed method [50], Iterative back projection [51] and bicubic [14]

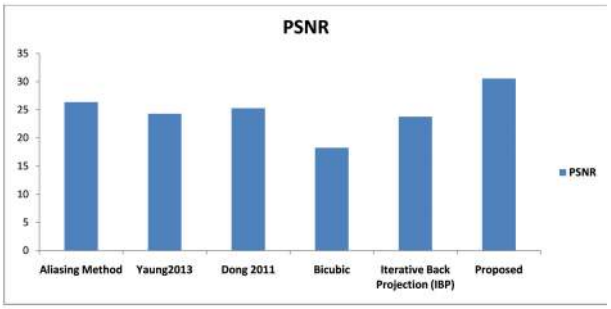


FIGURE 10. Performance evaluation for different methods.

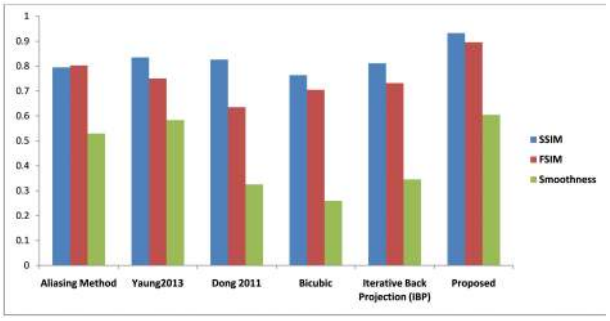


FIGURE 11. Comparison on SSIM, FSIM and Smoothness for different methods.

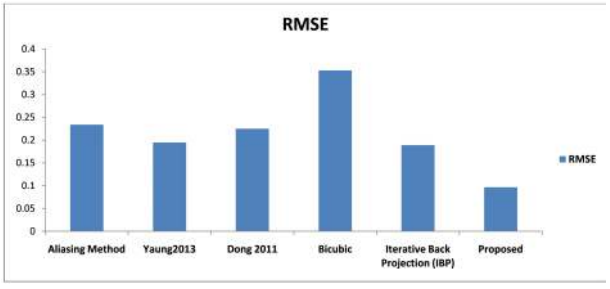


FIGURE 12. Comparison on RMSE for different methods.

method. The proposed method is compared with these high resolution images reconstruction mechanisms. The aliasing method and bicubic method estimate the values on the basis of neighbor values. These methods do not consider the trends of neighbor values. In these methods noise value also has active participation in estimating values. This degrades the performance of these methods. Other methods [49], [51] do not have the provision for input-output pattern analysis and exhaustive learning. Hence, the proposed method performs well with limited training image dataset. Figures 10 to 12 show the graphical presentation of different metrics.

D. CONCLUSIONS

The resolution of the medical x-ray image is an important factor to detect the abnormality, defect, injuries in the hard tissues. Modern x-ray systems that are used for medical diagnosis do not have resolution related issues. However, old medical diagnostic systems used in rural areas of developing countries have resolution related issues. The x-ray images are also important in forensic biometrics. A low-cost software solution is highly advantageous to solve this low resolution issue without changing the existing setup [52]. The available medical x-ray image repository does not have a huge number of training dataset. Hence, example based or CNN based super resolution techniques are inappropriate for these issues. On the other hand, single image super resolution methods do not provide a good result when magnification is greater than 1.6 times. Considering all these constrains, an alternative machine learning method is proposed in this research work. It uses existing image repository to generate a huge number of granular level feature dataset for training. These granular level feature sets are the combination of low frequency feature and high frequency feature. Low frequency feature contains the basic property and high frequency features contain the minute details. These granular level features along with their extraction mechanism is one of the contributions of this research work. The performance of this high resolution image reconstruction method is quite satisfactory. As the next phase of this work, authors are planning to localize the reconstruction mechanism within the region of interest [53], [54]. Consequently, the system itself will find the region of interest automatically for an unknown medical x-ray image. After that, it will reconstruct the high resolution image for those selected regions rather than the entire image. Such modified approach would also reduce the total computational cost.

ACKNOWLEDGMENT

Dr. Saumendu Datta, physician by profession, helped us to understand the basics of x-ray images from the medical domain perspective. The authors are thankful to him for this support. During the progress of the research work, the authors got infrastructural support from Visvesvaraya Ph.D. Scheme; Ministry of Electronics and IT, the Department of Computer Science and Engineering at University of Calcutta.

REFERENCES

[1] United Nations Office on Drugs and Crime. (2019). *Global Study on Homicide 2019*. [Online]. Available: https://www.unodc.org/unodc/en/data-and-analysis/global-study-on-homicide.html
[2] S. Datta, N. Chaki, and B. Modak, "A systematic review on the evolution of dental caries detection methods and its significance in data analysis perspective," in *Intelligent Data Analysis: From Data Gathering to Data Comprehension*. Hoboken, NJ, USA: Wiley, 2020.
[3] C.-Y. Yang, C. Ma, and M.-H. Yang, "Single-image super-resolution: A benchmark," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2014, pp. 372–386.
[4] S. Datta, N. Chaki, and K. Saeed, "Minimizing aliasing effects using faster super resolution technique on text images," in *Transactions on Computational Science*. Berlin, Germany: Springer, 2018, pp. 136–153.
[5] S. Borman and R. L. Stevenson, "Super-resolution from image sequences–A review," in *Proc. Midwest Symp. Circuits Syst.*, 1998, pp. 374–378.
[6] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.
[7] N. A. Woods, N. P. Galatsanos, and A. K. Katsaggelos, "Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 201–213, Jan. 2006.
[8] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP, Graph. Models Image Process.*, vol. 53, no. 3, pp. 231–239, May 1991.

[9] N. Nguyen and P. Milanfar, "A wavelet-based interpolation-restoration method for superresolution (wavelet superresolution)," *Circuits Syst. Signal Process.*, vol. 19, no. 4, pp. 321–338, Jul. 2000.
[10] A. Lukin, A. S. Krylov, and A. Nasonov, "Image interpolation by super-resolution," in *Proc. GraphiCon*, 2006, pp. 239–242.
[11] W.-C. Siu and K.-W. Hung, "Review of image interpolation and super-resolution," in *Proc. Asia Pacific Signal Inf. Process. Assoc. Annu. Summit Conf.*, Oct. 2012, pp. 1–10.
[12] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Jul. 2004, p. 1.
[13] F. Zhou, W. Yang, and Q. Liao, "Interpolation-based image super-resolution using multisurface fitting," *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3312–3318, Jul. 2012.
[14] Z. Pan, W. Chen, Z. Jiang, L. Tang, Y. Liu, and Z. Liu, "Performance of global look-up table strategy in digital image correlation with cubic b-spline interpolation and bicubic interpolation," *Theor. Appl. Mech. Lett.*, vol. 6, no. 3, pp. 126–130, 2016.
[15] H. Ur and D. Gross, "Improved resolution from subpixel shifted pictures," *CVGIP, Graph. Models Image Process.*, vol. 54, no. 2, pp. 181–186, 1992.
[16] Z. Lin and H.-Y. Shum, "Fundamental limits of reconstruction-based superresolution algorithms under local translation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 83–97, Jan. 2004.
[17] T. Zhang, Y. Yan Tang, B. Fang, Z. Shang, and X. Liu, "Face recognition under varying illumination using gradientfaces," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2599–2606, Nov. 2009.
[18] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 349–356.
[19] K. I. Kim, M. Franz, and B. Schölkopf, "Kernel hebbian algorithm for single-frame super-resolution," in *Proc. Workshop Stat. Learn. Comput. Vis.*, 2004, pp. 135–149.
[20] D. Datsenko and M. Elad, "Example-based single document image super-resolution: A global MAP approach with outlier rejection," *Multidimensional Syst. Signal Process.*, vol. 18, nos. 2–3, pp. 103–121, May 2007.
[21] M. V. Joshi, S. Chaudhuri, and R. Panuganti, "A learning-based method for image super-resolution from zoomed observations," *IEEE Trans. Syst., Man, Cybern. B. Cybern.*, vol. 35, no. 3, pp. 527–537, Jun. 2005.
[22] K. Nguyen, C. Fookes, S. Sridharan, and S. Denman, "Focus-score weighted super-resolution for uncooperative iris recognition at a distance and on the move," in *Proc. 25th Int. Conf. Image Vis. Comput.*, New Zealand, Oceania, Nov. 2010, pp. 1–8.
[23] K. Nguyen, C. Fookes, S. Sridharan, and S. Denman, "Quality-driven super-resolution for less constrained iris recognition at a distance and on the move," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 4, pp. 1248–1258, Dec. 2011.
[24] NIST. (2008). *National Institute of Standards and Technology: Multiple Biometric Grandchallenge (MBGT)*.[Online]. Available: http://face.nist.gov/mbgc
[25] F. Alonso-Fernandez, R. A. Farrugia, J. Bigun, J. Fierrez, and E. Gonzalez-Sosa, "A survey of super-resolution in iris biometrics with evaluation of dictionary-learning," *IEEE Access*, vol. 7, pp. 6519–6544, 2019.
[26] N. Othman, N. Houmani, and B. Dorizzi, "Improving video-based iris recognition via local quality weighted super resolution," in *Proc. ICPRAM*, 2013, pp. 623–629.
[27] N. Othman and B. Dorizzi, "Impact of quality-based fusion techniques for video-based iris recognition at a distance," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 8, pp. 1590–1602, Aug. 2015.
[28] A. E. W. Johnson, T. J. Pollard, N. R. Greenbaum, M. P. Lungren, C.-y. Deng, Y. Peng, Z. Lu, R. G. Mark, S. J. Berkowitz, and S. Horng, "MIMIC-CXR-JPG, a large publicly available database of labeled chest radiographs," 2019, *arXiv:1901.07042*. [Online]. Available: http://arxiv.org/abs/1901.07042
[29] Ehsanirada. (2015). *Digital Dental Periapical X-Ray Database for Caries Screening, Digital Dental X-Ray Image Dataset Along Ground Truth*. [Online]. Available: https://mynotebook.labarchives.com/share/Vahab/MjAuOHw4NTc2Mi8xNi9UcmVlTm9kZS83NzM5OTk2MDZ8NTIuOA
[30] P. Ghosh, D. Bhattacharjee, and M. Nasipuri, "Automatic system for plasmodium species identification from microscopic images of blood-smear samples," *J. Healthcare Informat. Res.*, vol. 1, no. 2, pp. 231–259, Dec. 2017.
[31] B. Horn, B. Klaus, and P. Horn, *Robot Vision*. Cambridge, MA, USA: MIT Press, 1986.
[32] E. H. Land and J. J. McCann, "Lightness and Retinex theory," *J. Opt. Soc. Amer.*, vol. 61, no. 1, pp. 1–11, 1971.
[33] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 249–256, Jul. 2002.
[34] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1635–1650, Jun. 2010.
[35] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
[36] A. Ramirez Rivera, J. Rojas Castillo, and O. Oksam Chae, "Local directional number pattern for face analysis: Face and expression recognition," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1740–1752, May 2013.
[37] F. Zhong and J. Zhang, "Face recognition with enhanced local directional patterns," *Neurocomputing*, vol. 119, pp. 375–384, Nov. 2013.
[38] P. Mooney. (2018). *Kaggle Database;Chest X-Ray Images (Pneumonia)*. [Online]. Available: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
[39] P. Ghosh, D. Bhattacharjee, and M. Nasipuri, "Intelligent toilet system for non-invasive estimation of blood-sugar level from urine," *IRBM*, vol. 41, no. 2, pp. 94–105, Apr. 2019.
[40] L.-L. Huang, A. Shimizu, Y. Hagihara, and H. Kobatake, "Face detection from cluttered images using a polynomial neural network," *Neurocomputing*, vol. 51, pp. 197–211, Apr. 2003.
[41] R. Ghazali, A. Hussain, and W. El-Deredy, "Application of ridge polynomial neural networks to financial time series prediction," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Dec. 2006, pp. 913–920.
[42] E. Agustsson and R. Timofte, "Ntire 2017 challenge on single image super-resolution: Dataset and study," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, Jul. 2017, pp. 126–135.
[43] A. Ignatov *et al.*, "PIRM challenge on perceptual image enhancement on smartphones: Report," in *Proc. Eur. Conf. Comput. Vis. (ECCV) Workshops*, Jan. 2019, pp. 315–333.
[44] R. Timofte, S. Gu, J. Wu, L. Van Gool, L. Zhang, M.-H. Yang, and M. Haris, "Ntire 2018 challenge on single image super-resolution: Methods and results," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit. (CVPR) Workshops*, June 2018, pp. 852–863.
[45] S. Datta. *LR to HR Image Reconstruction*. Accessed: Jan. 6, 2020. [Online]. Available: https://github.com/somadattaghosh/Training-dataset-for-LR-to-HR
[46] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
[47] S. Hu, L. Jin, H. Wang, Y. Zhang, S. Kwong, and C.-C.-J. Kuo, "Compressed image quality metric based on perceptually weighted distortion," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5594–5608, Dec. 2015.
[48] P. Jagalingam and A. V. Hegde, "A review of quality metrics for fused image," *Aquatic Procedia*, vol. 4, pp. 133–142, Oct. 2015.
[49] J. Yang, Z. Lin, and S. Cohen, "Fast image super-resolution based on in-place example regression," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1059–1066.
[50] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1838–1857, Jul. 2011.
[51] M. Nawaz, R. Xie, L. Zhang, M. Asfandyar, and M. Hussain, "Image super resolution by sparse linear regression and iterative back projection," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, Jun. 2016, pp. 1–6.
[52] P. Ghosh, D. Bhattacharjee, and M. Nasipuri, "Blood smear analyzer for white blood cell counting: A hybrid microscopic image analyzing technique," *Appl. Soft Comput.*, vol. 46, pp. 629–638, Sep. 2016.
[53] S. Datta, N. Chaki, and B. Modak, "Neutrosophic set-based caries lesion detection method to avoid perception error," *Social Netw. Comput. Sci.*, vol. 1, no. 1, p. 63, Jan. 2020.
[54] S. Datta, N. Chaki, and B. Modak, "A novel technique to detect caries lesion using isophote concepts," *IRBM*, vol. 40, no. 3, pp. 174–182, Jun. 2019.

KHALID SAEED (Senior Member, IEEE) received the B.Sc. degree in electrical and electronics engineering from Baghdad University, in 1976, and the M.Sc. and Ph.D. (distinguished) degrees from the Wroclaw University of Technology, Poland, in 1978 and 1981, respectively, and the D.Sc. degree (Habilitation) in computer science from the Polish Academy of Sciences, Warsaw, in 2007. He was with the Faculty of Mathematics and Information Sciences, Warsaw University of Technology, from 2014 to 2019. He was with AGH Krakow, from 2008 to 2014. He is currently a Full Professor of computer science with the Faculty of Computer Science, Bialystok University of Technology. He has published more than 245 publications, including 98 journal articles and book chapters and 87 peer-reviewed conference papers, edited 48 books, journals, and conference proceedings, and written 12 text and reference books (H-index 15 in the WoS base and 12 in the SCOPUS base). He has supervised more than 15 Ph.D. and 130 M.Sc. theses from and outside Poland. He is a member of more than 15 editorial boards of international journals and conferences. He gave 50 invited lectures and keynotes in different universities in Canada, China, Colombia, Czech, Germany, India, Japan, Serbia, Slovakia, and South Korea on biometrics, image processing and analysis. He received more than 30 academic awards. He was nominated by the President of Poland for the title of Professor in 2014. He has been selected as an IEEE Distinguished Speaker for period of 2011–2016. He is the Editor-in-Chief of the *International Journal of Biometrics* (Inderscience Publishers).

SOMA DATTA received the M.Sc. and M.Tech. degrees from the University of Calcutta, India, in 2005 and 2007, respectively.

She was with the Research and Development section, TATA Consultancy Service Limited as a System Engineer. After that, she joined the University of Calcutta, Kolkata, as a Ph.D. Research Fellow. She has a considerable number of paper publications in the image processing domain. Her research interests include medical imaging and pattern recognition.

NABENDU CHAKI (Senior Member, IEEE) is currently a Professor with the Department Computer Science and Engineering, University of Calcutta, Kolkata, India. He has served as a Visiting Professor in different places, including the U.S. Naval Postgraduate School, CA, and in different universities in Italy and Poland. Besides editing about 40 conference proceedings with Springer, he has authored eight text and research books with CRC Press, Springer Nature, and so on. He has published more than 200 Scopus Indexed research articles in journals and international conferences. He is the Founder Chair of the ACM Professional Chapter, Kolkata. He served the ACM Professional Chapter for three years. He has been active during 2009–2015 toward developing several international standards in software engineering and service science as a Global (GD) Member for ISO-IEC. He is the Editor-in-Chief of the Springer Nature book series on Services and Business Process Reengineering.