ADVANCE IN IMAGE AND AUDIO RESTORATION AND THEIR ASSESSMENTS: A REVIEW

Omar H. Mohammed and Basil Sh. Mahmood

Department of Computer Engineering, Mosul University, Nineveh, Iraq

ABSTRACT

Image restoration is the process of restoring the original image from a degraded one. Images can be affected by various types of noise, such as Gaussian noise, impulse noise, and affected by blurring, which is happened during image recordings like motion blur, Out-of-Focus Blur, and others. Image restoration techniques are used to reverse the effect of noise and blurring. Restoration of distorted images can be done using some information about noise and the blurring nature or without any knowledge about the image degradation process. Researchers have proposed many algorithms in this regard; in this paper, different noise and degradation models and restoration methods will be discussed and review some researches in this field.

KEYWORDS

Restoration, image, audio, signal, blurring & Noise.

1. INTRODUCTION

Digital signals today, such as audio and video, play an essential role in our lives. The digital signal may suffer from distortion during acquisition or transfer/storage stages, and essential information may be lost. Signal distortion can vary depending on the signal's nature, the acquisition device, and the transmission channel. One of the most famous distortions in the image is the blurring of the image. Blurring is a common, unwanted artifact associated with image formation; blurring is a deterministic process that can occur due to numerous reasons, such as atmospheric distortion motion, optical aberrations, and motion. In addition to blurring, image degradation is also caused by noise during image recording. Mathematically, image degradation due to blurring and noise can be modeled by a linear system, as shown in equations 1 and 2 [1]:

$$g\,(x,y) = A * f(x,y) + n(x,y) \qquad \ldots\ldots \quad (1)$$

$$f\,(x,y) = A^{-1} * g(x,y) - n(x,y) \qquad \ldots\ldots \quad (2)$$

Where (x,y) represents spatial coordinates, $g\,(x,y)$ is the acquired image with blur and noise, A is the blurring matrix defined by a certain point spread function (PSF), $f\,(x,y)$ is a true or the original image and $n\,(x,y)$ is noise. The problems associated with image restoration, which will be discussed in detail in the next sections, are choosing appropriate restoration techniques; some require an enormous dataset size, others required to build a precise distortion model, collecting or synthesizing datasets for fine-tuning the restoration model, and choosing suitable metrics for evaluating the restored result.

DOI: 10.5121/ijcses.2021.12201

2. BLURRING AND NOISE

The blurring is a distortion process on an image that happens in many scenarios, such as capturing an image when the camera or the captured object in motion, the camera is out-of-focus, or atmospheric turbulence or fog. The main types of blur are [2]:

A. Motion Blur: While capturing an object's image, this blurring type happens if the camera or the object in motion; motion blur has many possibilities such as rotation, translation, an abrupt change in the scale, or a combination of them.

B. Average Blur: Average blur has a distortion effect horizontally and vertically across the whole image.

C. Gaussian Blur: the Gaussian blur, also known as Gaussian smoothing, has an unequal effect on the image pixels; this blurring type following the bell-shaped curve means that pixels at the center have more effected than those at the edges.

D. Out-of-Focus Blur: this blurring happened when the captured scene is not centered at the camera's lens depth.

E. Atmospheric Turbulence Blur: images could be affected by this type of blur when captured from relative long-distance, due to the atmospheric turbulence, like when the temperature variation is considerable, wind speed, or during a sand storm.

Images are usually corrupted by undesirable effects known as noise from various sources during taking images or when transmitting. Noise randomly changes image pixels intensity; these changes could be visually seen as grain or speckles on the image. The main types of noise are [3]:

A. Gaussian noise: Gaussian noise degrades images by following Gaussian amplitude distribution; this noise is also known as normal noise.

B. Impulse noise: this noise, also known as salt and pepper noise, follows a random distribution; it Occurs during signal transmission or due to sensor malfunction; It is visible as white or black dots on the pictures.

C. Uniform noise: which also known as quantization noise, occurs due to quantizing pixels intensity to a set of discrete levels.

D. Poisson noise: also referred to as shot noise, when capturing low-light scenes, this noise occurs because capturing numbers of photons that may not have enough energy to be detectable by the camera's sensor.

E. Speckle noise: This is a multiplicative noise arising due to backscattered echo, usually found in medical images and synthetic aperture radar (SAR), medical images, and satellite images.\

An audio signal's degradation will be considered an undesirable modification to the audio signal, which occurs due to (or after) the recording process; for example, in a recording made, the degradations could include noise in the microphone and amplifier[4]. Signal restoration is recovering the original signal, which considers one of the fundamental and challenging digital signal processing tasks. Many techniques are used in this field, such as filtering approaches, signal transformation domains such as fast Fourier wavelet, statistical approaches, and artificial intelligence, such as neural network and fuzzy. Another approach deals with the 1D signal as a 2D

image, such as Mel-frequency cepstral coefficients (MFCCs), which help restore signal using image restoration techniques[5, 6].

3. RESTORATION TECHNIQUES

The main types of techniques using for image restoration are Non-Blind and Blind Deconvolution Techniques [7].

3.1. Non-Blind Deconvolution Techniques Margins

If there is some knowledge about the degradation process, the restoration process referred to as the non-blind deconvolution method.

A. Inverse Filtering: in this method, the Fourier transform of the original image $F(v,u)$ is found by division the corrupted image $G(v,u)$ by the degradation function $H(v,u)$ shown in equations 3, then Fourier inverse is used to obtain the original image[8].

$$F(v,u) = \frac{G(v,u)}{H(v,u)} \qquad \ldots\ldots \ (3)$$

The main problem is when the degradation image is affected by noise, this method will give a low restoration image quality, so if noise is introduced, other methods must be used [9].

B. Weiner Filter: The Wiener filter is the MSE-optimal stationary linear filter for images degraded by additive noise and blurring. The objective of this filter is to find $\hat{F}$ the estimate of the original image to minimize the mean square error. Wiener filter used in the frequency domain as in equation 4, where $G(u,v)$ is the degradation image, $H(u,v)$ is the blurring function, $H^*(u,v)$ is its complex conjugate, and $\frac{N(u,v)}{I(u,v)}$ is the signal-to-noise ratio [10].

$$\hat{F} = \left[\frac{H^*(u,v)}{|H(u,v)|^2 + \frac{N(u,v)}{I(u,v)}}\right] G(u,v) \qquad \ldots\ldots \ (4)$$

C. Constraint Least-Square Filter: it is similar to the Wiener filter, but without knowledge of noise property, it works as shown in equation 5, where γ is the parameter to be adjusted manually, and p is the Fourier transform of Laplacian $P(u,v)$ [11].

$$\hat{F} = \left[\frac{H^*(u,v)}{|H(u,v)|^2 + \gamma|P(u,v)|^2}\right] G(u,v) \qquad \ldots\ldots \ (5)$$

$$p(x,y) = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \qquad \ldots\ldots \ (6)$$

D. Lucy- Richardson Algorithm: is an iterative non-linear restoration method. This algorithm's advantage is not considering with knowledge of the type of noise and the original image. The new estimate of the original image $\hat{f}_{n+1}$obtained from the previous one $\hat{f}_n$ as in equation 7 [12].

$$\hat{f}_{n+1} = \hat{f}_n + (g - h^* \hat{f}_n) \qquad \ldots\ldots \ (7)$$

Where g is the distorted image and h^* is the blurring matrix.

3.2. Blind Deconvolution Techniques Margins

In this technique, an image is restored without having prior knowledge of the degradation process.

A. Mean Filter: The mean filter is a linear filter used to smooth the image by replacing each pixel with the average of its neighborhood pixels; this process is done by sliding a squared window with the pixel that wants to be replaced in the center process repeated to the all pixels in the image. Mean filter is a simple and powerful tool to remove Gaussian noise, but not useful when the degradation image is affected by salt and pepper noise and blurs [13].

B. Adaptive Mean Filter: The adaptive mean filter is similar to the mean filter, but the filter size can be changed according to the image's local area. This filter is capable of removing high-density noise from degradation images and preserves detail and smooth non-impulsive noise. The main drawback of dealing with salt and pepper noise cannot remove any blurring effect [14].

C. Order Static Filters: are non-linear filters depend on the ranking of neighborhood pixels' values; these family filters include median, min, max, and alpha trim filters. The Median filter works by replacing the pixel's value with the median value of its neighborhood pixels; this is done via applying the center of a window of size M×N (N and M must be odd values). The process is repeated to all pixels; this filter is used to remove the impulsive noise; it can also keep the edge characteristics but cannot remove the blurring effect. The alpha-trimmed mean filter is similar to the median and means filters, as, in the median filter, the neighborhood pixels are ranked, but the extreme pixel values are excluded (trimmed), then the remained pixels are averaged as done in the mean filter [15].

D. Statistical approach: the statistical methods try to find the estimated original image and blurring function by observing the distorted images. The maximum-a-posterior (MAP) most widely used in this approach; the MAP estimate the original image $\hat{f}$:

$$\hat{f} = \arg\ min_f\ (g - Af) + \lambda p\ (f) \ldots\ldots\ (8)$$

Where $(g - Af)$ is the likelihood loss function, and $\lambda p\ (f)$ is prior to the image space [16].

E. Deep Neural network: Neural networks can be used for either classification or regression(ex. output is an image), a deep learning network that output an image that usually contains transposed convolutional layers[17]. An autoencoder, as shown in figure 1, consists of two parts: encoder and decoder; the encoder takes input images passing it to convolutional layers to create a latent or compressed representation of the input images, follow by the decoder, which tries to reconstruct the input image from its latent features, the autoencoder could be used for the image to image regression, the other deep learning networks such as generative adversarial networks (GAN) share the same principles used in autoencoder [18].

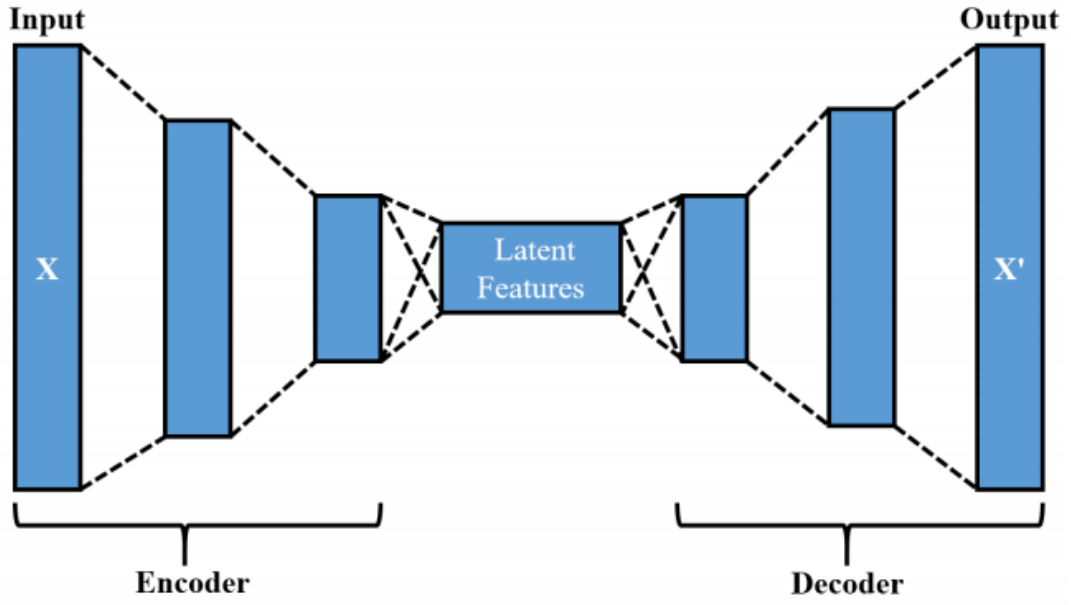


Figure 1. autoencoder architecture

Deep neural network (DNN) has recently been widely used in blind restoration by using two approaches: PSF estimation and end-to-end aproach[19]. The first approach estimates the kernel blurring function by using a fully-convolutional deep neural network (FCN), then by using the estimated kernel, the deconvolutional process is used to produce the non-blurry image closer to the original image. The FCN is training on a dataset containing distorted images with its blurring kernel, and then the FCN takes a new distorted image to produce its blurred kernel[20].

The second approach uses a deep deconvolutional neural network (DDNN), which is an end to end take the distorted image as input and produced an approximated original image; the process in DDNN consists of an encoder network that produces the visual feature of an image and decoder network to build the sharp image [21]. The end-to-end approach also could be implemented using conditional GAN, which could be used for image generation[22]. The network of general GAN consists of two networks the generator and discriminator (CNN); the generator generate randomly an image (referred to as a fake image), the discriminator work as a classifier to figure out if the image is real or fake(generated by the generator), the generator and discriminator work against each other, after training the generator will produce very realistic image[23]. In the conditional GAN used for image restoration, the generator takes a distorted image as input and generates an image wanted to be the nearest as possible to the original non-distorted image [24]. The discriminator part tells if the output image is sharp or distorted. During training, the generator and discriminator compete with each other, which leads to better production of the sharp image and distorted image detection. Also, an essential condition must be used in training the generator part, which is a content loss function that guided the generator not to generate uncorrelated images with the sharp image; MSE, SSIM, or Perceptual losses can be used as the content loss function[25, 26].

4. IMAGE QUALITY ASSESSMENT

Image quality assessment (IQA) is used to measure how the image restoration process is succeeded; it determines if the resulted image's quality is good enough. Depending on the availability of a reference Image, the IQA methods fall into three categories: full reference (FR), no-reference (NR), and reduced-reference (RR) methods [27]. The most methods used in FR IQA are Spatial

domain, which includes: mean squared error (MSE), root mean squared error (RMSE), signal to noise ratio (SNR), peak signal to noise ratio (PSNR), human visual system (HVS), structure similarity index method (SSIM), Features Similarity Index Matrix (FSIM), transformation domain, and learn the based method. The MSE is calculated, as shown in equation 9, by taking the summation of squared error between each pixel in restored $\hat{f}$ and original image f then divided by the total number of pixel in that image. The RMSE, as its name indicates, is produced by taking the root of MSE, RMSE=$\sqrt{MSE}$ [27].

$$MSE = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}\left(\hat{f}(x,y)-f(x,y)\right)^2}{\mathrm{M}\times\mathrm{N}} \quad \ldots\ldots \quad (9)$$

The SNR, as shown in equation 10, is measuring the signal power (the original image f) to the noise power, which is the difference between restored image $\hat{f}$ and the original image f from equation 10, it is noticed that noise is equal to MSE multiplies by (M × N). The PSNR is calculated by peak signal power to the distorting noise (MSE) power, as shown in equation 11; when the pixel is 8 bits, the peak value is 255 [19].

$$SNR = 10\,log_{10}\left(\frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(f(x,y))^2}{\sum_{x=1}^{M}\sum_{y=1}^{N}\left(\hat{f}(x,y)-f(x,y)\right)^2}\right) \quad \ldots\ldots \quad (10)$$

$$PSNR = 10\,log_{10}\left(\frac{peakvalue^2}{\mathrm{MSE}}\right) \ldots\ldots \quad (11)$$

The human visual system-based metrics (HVS) are trying to evaluate the restored, and distorted images as human eyes perceptually see the difference between two images; this requires modeling the human eye, which required many physiological and psychophysical experiments[28]. The HVS implemented by applying the discrete cosine transform (DCT) and then contrast sensitivity function (CSF) on both the reference and images need to be judged. The CSF is a band-pass filter associate with the human eye see images in the frequency domain. The HVS could be calculated by measuring, using the MSE, the difference between the reference image's CSF and distortion or restored image's CSF[28].

The structure similarity index method (SSIM) is a perception-based model; the SSIM measures the similarity between two images; this method is usually used to measure the compressed image's quality. Unlike MSE and PSNR measure the absolute error, the SSIM give more information about how two images are similar to each other by measuring the similarity between the two images by three terms luminance $l(x,y)$, contrast $c(x,y)$ and structure $s(x,y)$ [29]:

$$l(x,y) = \frac{2\mu_x\mu_y+C_1}{\mu_x{}^2+\mu_y{}^2+C_1} \quad \ldots\ldots \quad (12)$$

$$c(x,y) = \frac{2\sigma_x\sigma_y+C_2}{\sigma_x{}^2+\sigma_y{}^2+C_2} \quad \ldots\ldots \quad (13)$$

$$s(x,y) = \frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3} \quad \ldots\ldots \quad (14)$$

Where μ_x and μ_y are the average of original and restored images, respectively, σ_x and σ_y the standard deviations of the two images, σ_{xy} is the cross-covariance of the two images, and C_1, C_2 and C_3 are constants close to zero.

The general form of SSIM is calculated from the three terms as shown in equation 15:

$$SSIM\ (x,y) = \ [l(x,y)]^{\alpha}\ [l(x,y)]^{\beta}\ [l(x,y)]^{\gamma} \quad \ldots\ldots\ (15)$$

Where α, β, and γ are the positive constants, usually set to be α=β=γ=1, so the equation of 15 become equation 16:

$$SSIM\ (x,y) = \frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x{}^2+\mu_y{}^2+C_1)(\sigma_x{}^2+\sigma_y{}^2+C_2)} \ldots\ldots\ (16)$$

The features similarity index matrix (FSIM) measures the similarity between two images based on phase congruency (PC) and gradient magnitude (GM); the phase congruency is a frequency-based algorithm that gives the critical feature of an image without affected by contrast. the gradient magnitude (GM) of an image is given as [29]:

$$G = \sqrt{G_x{}^2 + G_y{}^2} \ldots\ldots\ (17)$$

Where G_x and G_y are respectively the horizontal and vertical gradients. If the phase congruency feature result PC_1 and PC_2 for two images f_1 and f_2 respectively, and G_1 and G_2 are the gradient magnitude of f_1 and f_2, the similarity of the two images is:

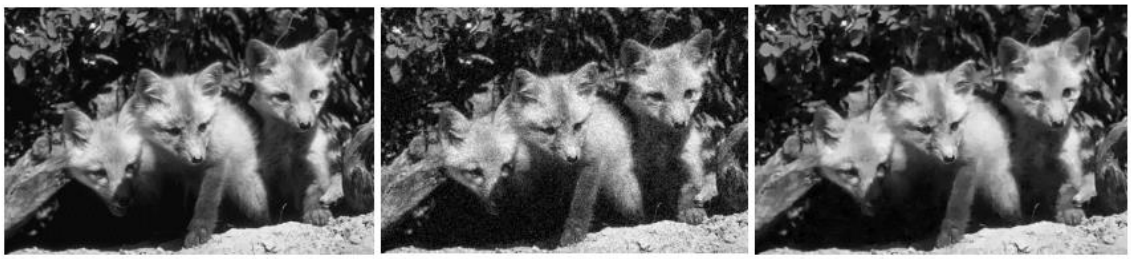
$$S_L = S_{PC}(x)^{\alpha}.S_G(x)^{\beta} \quad \ldots\ldots\ (18)$$

Where α and β are adjusting parameter similar used in SSIM above, S_{PC} and S_G are respectively measure the similarity using PC_1 and PC_2And the similarity by G_1 and G_2, as shown in equation 19 and 20, where $T_1$and T_2 are constants close to zero.

$$S_{PC} = \frac{2\ PC_1PC_2+T_1}{PC_1{}^2+PC_2{}^2+T_1} \quad \ldots\ldots\ (19)$$

$$S_G = \frac{2\ G_1G_2+T_2}{G_1{}^2+G_2{}^2+T_2} \quad \ldots\ldots\ (20)$$

Image distortion could be seen in the wavelet and DCT transformations domain, so these transformations could be used to evaluate the restored image by extracting the feature in the wavelet or DCT domain [30]. Figure 2 shows an example for image metrics MSE, PSNR, and SSIM for the reference(sharp), distorted, and the restored images[31].



Reference image	Distorted image	Restored image
	MSE=538	MSE=104
	PSNR=20.8	PSNR=27.95
	SSIM=0.46	SSIM=0.81

Figure 2. Image metrics for reference, distorted, and restored images.

In practical application, the image reference image is not available, so no-reference (NR) IQA is needed to be used. There are mainly two NR IQA, the feature-based approach by building a model for a specific image artifact and a learning-based approach such by using neural networks. A deep neural network (DNN) is supplied by standard image (not distorted) and distorted with various artifacts and blurring types with its IQA value (manually) after the DNN trained well; it could be used to calculate the IQA and gives the artifact types [32, 33].

The Perceptual Evaluation of Audio Quality (PEAQ) is usually used to determine the decoded audio signal's quality by different coding methods by comparing non-coded audio signals with decoded ones; however, the PEAQ could be used for general audio signal restoration. Briefly, PEAQ compares two audio signals by first pre-processing the two signals, then using the short-time Fourier transform (STFT) or filtering bank, aided by a psychoacoustic model form features. The quality grade is found by applying these features to a regression model[34]. Recently deep learning is used for measuring audio quality that could be implemented in a various but straightforward method is to feed audio signal onto a CNN after converted to spectrograms, to get features (after activation function), these features then compare with clean signal's features to calculate the perceptual audio quality[35].

5. DATASET

Datasets play an essential role in developing and evaluating restoration algorithms, extraordinarily for deep learning-based algorithms. Datasets used in this domain problem usually require distorted images and clean or sharp images used as ground truth images. It is a great challenge to create these pairs of images, except for some cases, the distortion process needs to be modeling; distorted images are synthesized from clean images by adding the modeled distortion. The synthesized distorted image must be similar to the real one; else, algorithms could not succeed in the real-world image[36]. For deep learning-based gigantic datasets are required for training, but this will tremendous computational power; for this reason, many restoration networks use for some parts pre-trained networks if available or, by apply transfer learning technique where a pre-trained network modified to fit the problem by removing the last layers[29].

6. LITERATURE REVIEW

Here, a literature review of related works is discussed.

HONG-XIA DOU1 et al. [37] proposed an algorithm for image restoration using modified Tikhonov regularization (MTR) based on standard Tikhonov regularization accelerated by using the Lanczos preconditioner technique. A preconditioned MTR (PMTR) algorithm is then designed based on the conjugate gradient least squares (CGLS) method to solve the ill-posed problem. Their results showed that the proposed algorithm has a good and stable performance in terms of PSNR and relative error (ReErr).

Chidananda Murthy M V et al. [38]. This paper made a comparative analysis for image restoration for images degraded by several types of artifacts like Gaussian noise, motion blur, and out-of-focus blur; the analysis applied several methods, including Weiner filter, Lucy Richardson Constrained Least Square(CLS), and blind deconvolution methods. Their results and analysis show that the CLS filter outperformance other filters, by testing it against blurring and noise combinations, gives the best performance according to PSNR, MSE, and Correlation index (CI) metrics.

Arun Kumar Patel et al. [39] described a method to remove the motion blur present in the image taken from any camera. To minimize the computational complexity, they implemented an algorithm by using a Wiener filter and wavelet-based image fusion. They found, according to their

result, that the proposed algorithm gives better performance compared Wiener filter and other blind convolution methods in terms of SNR and RMSE.

Madhuri Kathale1 et al. [40] discussed various motion deblurring techniques and proposed image restoration using statistical local and non-local approaches by using adaptive discrete cosine transform(DCT). The statistical approach to produce restored image with local smoothness and the non-local approach to maintenance self-similarity with the original image. The split Bregman-based algorithm was used to solve distorted structure images due to inverse problems.

Charu Khare et al. [41] discussed practical algorithm techniques for image restoration such as histogram Adaptive Fuzzy (HAF), weighted fuzzy mean (WFM) filter, Minimum-maximum detector Based (MDB) filter, adaptive fuzzy median (AFMF) filter, and others; then presented a novel algorithm for analysis different filtering methods image enhancement and restoration using. They applied these restoration techniques on the LENA image distorted by many artifacts in different levels; their result shows that HAF gives about 5 dB PNSR more than the closet performance technique (AFMF).

Sheelu Mishra et al. [42] proposed an image restoration and image enhancement technique that will be used to restore the original image from a fog degraded image by using an integrated technique that will integrate the non-linear enhancement technique with the gamma correction and dynamic restoration technique. The proposed work claimed to be very simple but effective for restoring clear images from foggy ones in real-time. Their proposed algorithm worked by using an adaptive median filter after re-blurring the image with two Gaussian kernels, producing two images, then calculating the average of gaussian blurring to estimate the absolute image depth. The proposed method was applied on many hazy images; for the village image, the contrast improvement by 0.4030 with 0.084% saturated pixel.

Poonam Baruah et al. [43] presented an image restoration by denoising based on soft thresholding. The process recovers the degraded images by adopting a dynamic wavelet transform to minimize the error to the extent that it helps achieve satisfactory, quality, and suitable forms for specific medical applications. Several types of noise are used to distort images, such as Gaussian, impulse, and Poisson noise, in their work. The restored image was produced by removing noise using an adaptive dynamic process of wavelet transform (WT); this process is made dynamic by continuing the process until it achieves the desired goal. Their proposed approach provides a 2.45 to 9.92% PSNR improvement for restored images.

Fazil Altinel et al. [44] proposed an image inpainting method using an energy-based model to learn the structural relationship in an image between non-masked and the masked regions. Their proposed claimed to be outperformance, by PSNR of 2.23 dB, the GANs based methods. In their work, the CNN is used to find an energy function to learn the structural relationship between missing and non-missing parts. They designed a CNN of two paths, the first one takes distorted image by missing part, and the latter takes the restored image produced by the algorithm; during the training phase, using the L1 distance between the estimated image and its reference (non-distorted image), the CNN learns to provide energy function produce restored image similar to the ground truth image.

H. Lin et al. [45] created two datasets KADID-10K and KADIS-700K, which could be very useful for evaluating restoration algorithms based on the IQA, specifically for the deep learning-based algorithm. The KADID-10K has only 81 images and 140,000 for KADID-700K; they applied 25 types of image distortion, which includes: brightness changing, color shifting, compression, and different kind of noise, by creating new methods for distortion and utilize the existing ones. Images in their dataset are distorted by these 25 distortion types in five levels range based on visual quality,

level 5 will give the worst distortion, and level 1 is the less one. Also, they evaluated eighteen IQA methods, seven for NR-IQA and the rest for FR-IQA.

Y. Li et al. [46] introduced an audio-to-audio solution for denoising historical music using the GAN network by changing the domain problem from audio to image-like, using short-time Fourier transform (STFT). The GAN network consists of a generator and two discriminators. The generator model based on the U-Net network takes the STFT of the noisy music as input and tries to produce a clean one using inverse STFT. The first discriminator based on STFT has an architecture similar to the generator's encoder part, the second discriminator in the waveform domain using the same architecture of MelGAN. A dataset was created to train the GAN network, from old digitized music of two sources, one available on Public Domain Project, and the latter a collection of CD-quality; the dataset contains pairs of clean audio music and its noisy version, the clean one is a modern recording used as ground truth, and the noisy version is synthesized by adding noise. The noise is extracted from the old music's time domain at the low-energy segment of 100 ms, which represents the pause of music, then synthesized noisy music is created by mixing the clean music with overlapping these noise segments.

T. Hsieh et al. [47] proposed an end-to-end for speech enhancement and restoration; the proposed method is waveform-mapping-based, without using signal transformation like STFT or MFCC. They built a deep neural network called WaveCRN that utilizes both the input waveform's local and sequential features. The waveform input X goes through a 1-D CNN to produce a local feature mapped F. The computationally of the CNN is reduced by choosing the convolution strive to be half the kernel size. The sequential features are captured using the bidirectional simple recurrent units (Bi-SRU) instead of long short-term memory (LSTM) because the SRU is more efficient for utilizing parallel computing. The Bi-SRU create sequential features by using the features produced by the 1-D CNN in both directions to produce restricted feature mask (RFM) M, then masked feature map F is produced by multiply M by F. Finally, the speech waveform of the same length as the input waveform is regenerated by using a transposed 1D Convolution.

O. Kupyn et al. [48] introduced a conditional GAN(DeblurGAN) for an end-to-end for image restoration of motion blurring. The DeblurGAN architecture, as shown in figure 3, like other GANs, consist of a generator and discriminator (critic).

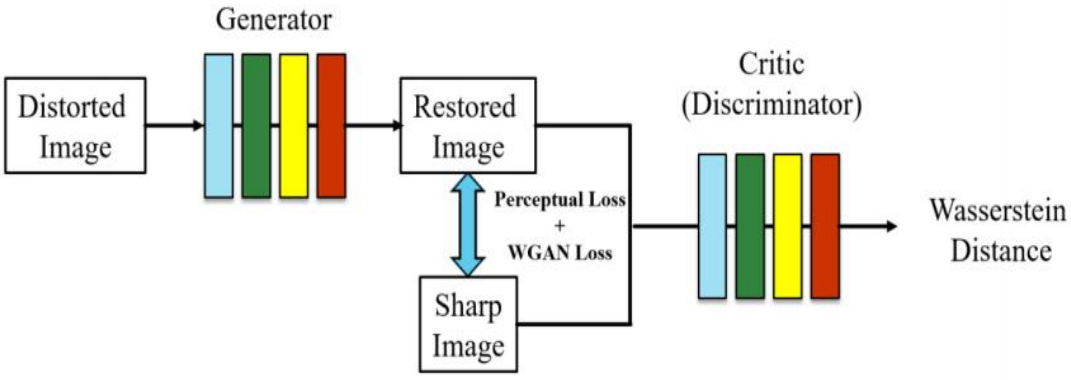


Figure 3. DeblurGAN Architecture

The generator takes a distorted image and produces an approximation of the sharp image; the discriminator tries to determine if it is produced by a generator (fake image) or a real image. They

use a combination of two-loss functions to train the network, as shown in equation 21, one for adversarial loss function, and the latter is content loss function.

$$L = L_{GAN} + \lambda \cdot L_X \ \ldots\ldots \quad (21)$$

Where L_{GAN} is the adversarial loss focus on texture detail by using Wasserstein GAN, λ is a parameter equals to 100, and L_X the content loss used for producing a restored image has similar general content to the distorted image. They use a feature map of the convolution layer to calculate the loss content instead of MSE or MAE using the activation function from VGG33 of the pre-trained VGG19.

O. Kupyn et al. [49], based on their succeeded network the DeblurGAN for motion blur restoration, built DeblurGAN-v2 according to their experiential results is state-of-the-art in both restoration quality, and network's complexity also introduced flexibility capabilities by using bulky or light CNNs as network's core based on targeted results efficiency or quality. The generator model in the GAN network is based on the feature pyramid network (FPN).

P. Jia et al. [50] proposed Cycle-GAN deep neural network, a conditional GAN, to restore astronomical images and estimate the PSF. Their Cycle-GAN contains two generative models, the first generator (PSF-Gen), to estimate the PSF and the other (Dec-GAN) for image restoration by learning the deconvolution kernel. They tested their method with real astronomical images taken by solar telescope and small aperture telescopes; the Dec-Gen gives convenient restored images for solar images and can reduce the PSF variation of smaller telescopes. The PSF-Gen provides a non-parametric PSF model for short exposure images.

Y. Zhou et al. [51] introduced a novel solution for image restoration of a degraded image taken by the under-display camera(UDC) due to lower light transmission rate and diffraction effects. Image restoration for this task required removing noise, apply anti blurring technique, and low-light enhancement. They used a camera with two types of displays: 4K Transparent OLED (T-OLED) and a phone Pentile OLED (P-OLED), to create two pairing datasets (sharp and degraded image), one with real image pairing and the latter is synthesized near-realistic paring by creating a model of the degradation process. They analyzed the two types of display according to light transmission rate, point spread function (PSF), and modulation transfer function (MTF). The light transmission rate was found to be 20% and 2.9% for T-OLED and P-OLED, respectively, measured by using a spectrophotometer and white light source. The PSF is measured by using a laser of (λ = 650nm). In T-OLED, the MTF effect on horizontal direction contrasts by nearly lost in the mid-band frequency due to diffraction, but for P-OLED, the MTF has zero effect, the MTF analyzed by recording sinusoidal patterns with increasing frequency in both lateral dimensions. Image restoration was done using a convolutional neural network with a similar structure to UNet and two sub-encoder. SSIM and perception loss evaluated the obtained results.

Table 1 below compares some restoration techniques using the IQA: PSNR, MSE, and SSIM (comparing restructured and reference images).

Table 1. Comparing between methods for image restoration

Method name	Dataset	Tools	objection	PSNR	MSE	SSIM
Pix2Pix[52]	Kohler[53]	Conditional GAN, Generator(U-Net), Discriminator(PatchGAN)	Motion deblur	25.41	188.5	0.810
DeblurGAN[48]		Conditional GAN, Generator(Resnet bloc), Discriminator(PatchGAN)	Motion deblur	26.10	160.8	0.816
DeblurGAN-v2[49]		Conditional GAN with two discriminators, FPN	Motion deblur	26.72	139.4	0.836
Sun et al.[20]		Blur kernel estimation by CNN	Motion deblur	25.22	197	0.773
DeepDeblur[54]		multi-scale CNN(ResBlock)	Motion deblur	26.48	147.3	0.807
DuRN-U[55]	GoPro[54]	Dual Residual Block	Motion deblur	29.9	67.06	0.91
DeepDeblur [54]		multi-scale CNN(ResBlock)	Motion deblur	28.3	96.93	0.92
Sun et al.[20]		Blur kernel estimation by CNN	Motion deblur	24.6	227.2	0.84
[56]		Motion flow estimation	Motion deblur	23.64	283.4	0.823
Y. Zhou[51]	-	UNet based architecture(two sub-encoder)	Image restoration under-display camera(T-OLED)	36.69	14.04	0.971

Table 2 shows some methods used for speech enhancement by comparing the improvement speech with the noisy one, using metrics *Perceptual Evaluation of Speech Quality* (*PESQ*), which is similar to PEAQ but for speech, and segmental SNR(SSNR).

Table 2: Comparing methods for speech enhancement.

Methods name	Dataset	Tools	ΔPESQ	Δ SSNR
Wiener filter	[47, 57]	-	0.25	3.39
SEGAN[58]		Conintial GAN, 1D Generator	0.19	6.05
WaveCRN[47]		CNN, SRU	0.67	8.58
Wave-U-Net		Wave-U-Net	0.65	8.37
NMF	[59, 60]	-	0.14	1.53
LMMSE		-	0.20	3.15
NRPCA		-	0.25	27.15
DNNP[60]		MFCC,DNN	0.49	4.4

7. CONCLUSIONS

Restoration technologies play an important role in the DSP field, as all methods attempt to obtain a restored version that looks as similar as possible to the original versions, and there is no perfect algorithm to restore the exact original version in the case of real-world problems. All restoration research approximately share the same steps, including preparing appropriate datasets, building a restoration model and fine-tuning it with train datasets, and finally evaluating the restoration results using test datasets with metrics such as PSNR, SSIM, and others. In this research, we explain the theories and models for image and signal restoration and try to cover most of the terms and methods

in this field; then, several types of research are discussed, including the most recent. Deep learning has become the predominant approach for the restoration process, almost bypassing traditional methods such as wiener filer and others; because deep learning methods give fair and realistic results compared to traditional techniques, and in the case of using GANs could be inpainting the unrestorable parts with most other parts and appear realistic. More researches are needed to reinvest in the traditional method by somehow combining it with deep neural networks.

REFERENCES

[1] D. Perrone and P. Favaro, "A clearer picture of total variation blind deconvolution," IEEE transactions on pattern analysis and machine intelligence, vol. 38, pp. 1041-1055, 2015.

[2] D. Singh and M. R. Sahu, "A survey on various image deblurring techniques," International Journal of Advanced Research in Computer and Communication Engineering, vol. 2, pp. 4736-4739, 2013.

[3] R. Verma and J. Ali, "A comparative study of various types of image noise and efficient noise removal techniques," International Journal of advanced research in computer science and software engineering, vol. 3, pp. 617-622, 2013.

[4] E. Vincent, S. Watanabe, A. A. Nugraha, J. Barker, and R. Marxer, "An analysis of environment, microphone and data simulation mismatches in robust speech recognition," Computer Speech & Language, vol. 46, pp. 535-557, 2017.

[5] J. SIMON, R. GODSILL, and J. PETER, DIGITAL AUDIO RESTORATION: SPRINGER LONDON Limited, 2013.

[6] D. M. Nogueira, C. A. Ferreira, E. F. Gomes, and A. M. Jorge, "Classifying heart sounds using images of motifs, MFCC and temporal features," Journal of medical systems, vol. 43, p. 168, 2019.

[7] S. Jain and M. S. Goswami, "A Comparative Study of Various Image Restoration Techniques With Different Types of Blur," International Journal of Research in Computer Applications and Robotics, vol. 3, pp. 54-60, 2015.

[8] A. Swarnambiga, Medical Image Processing for Improved Clinical Diagnosis: IGI Global, 2018.

[9] M. Kathale and A. Deshpande, "REVIEW PAPER ON IMAGE RESTORATION USING STATISTICAL MODELING," IJRET: International Journal of Research in Engineering and Technology, vol. 05, pp. 235-238
2016.

[10] M. R, "Restoration Of Blurred Images Using Wiener Filtering," International Journal Of Electrical Electronics And Data Communication, vol. 5, pp. 2320-2084.

[11] Y. Zhao, X. Sun, C. Zhang, and Y. Zhao, "Using Markov constraint and constrained least square filter to develop a novel method of passive terahertz image restoration," in Journal of Physics: Conference Series, 2019, p. 042094.

[12] H.-L. Yang, P.-H. Huang, and S.-H. Lai, "A novel gradient attenuation Richardson–Lucy algorithm for image motion deblurring," Signal Processing, vol. 103, pp. 399-414, 2014.

[13] P. K. Patidar and P. Dadheech, "Performance of Fuzzy Filter and Mean Filter for Removing Gaussian Noise," International Journal of Computer Applications, vol. 975, pp. 29-35, 2019.

[14] A. Shah, J. I. Bangash, A. W. Khan, I. Ahmed, A. Khan, A. Khan, et al., "Comparative Analysis of Median Filter and its Variants for Removal of Impulse Noise from Gray Scale Images," Journal of King Saud University-Computer and Information Sciences, 2020.

[15] H. Joshi and D. J. Sheetlani, "Image Restoration Techniques in Image Processing: An Illustrative Review," International Journal of Advance Research in Science and Engineering, vol. 6, 2017.

[16] E. Y. Lam and J. W. Goodman, "Iterative statistical approach to blind image deconvolution," JOSA A, vol. 17, pp. 1177-1184, 2000.

[17] H. Gao, H. Yuan, Z. Wang, and S. Ji, "Pixel transposed convolutional networks," IEEE transactions on pattern analysis and machine intelligence, vol. 42, pp. 1218-1227, 2019.

[18] S. Ladjal, A. Newson, and C.-H. Pham, "A PCA-like autoencoder," arXiv preprint arXiv:1904.01277, 2019.

[19] X. Xu, J. Pan, Y.-J. Zhang, and M.-H. Yang, "Motion blur kernel estimation via deep learning," IEEE Transactions on Image Processing, vol. 27, pp. 194-205, 2017.

[20] J. Sun, W. Cao, Z. Xu, and J. Ponce, "Learning a convolutional neural network for non-uniform motion blur removal," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 769-777.

[21] X. Mao, C. Shen, and Y.-B. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in Advances in neural information processing systems, 2016, pp. 2802-2810.
[22] S. Vyas, "Brain Computer Interfaces Employing Machine Learning Methods : A Systematic Review," International Journal of Computer Science & Engineering Survey (IJCSES), vol. 11, p. 1, April 2020.
[23] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al. (2014, June 01, 2014). Generative Adversarial Networks. arXiv:1406.2661. Available: https://ui.adsabs.harvard.edu/abs/2014arXiv1406.2661G
[24] D. W. Kim, J. R. Chung, J. Kim, D. Y. Lee, S. Y. Jeong, and S. W. Jung, "Constrained adversarial loss for generative adversarial network-based faithful image restoration," ETRI Journal, vol. 41, pp. 415-425, 2019.
[25] Z. Hong, X. Fan, T. Jiang, and J. Feng, "End-to-End Unpaired Image Denoising with Conditional Adversarial Networks," in AAAI, 2020, pp. 4140-4149.
[26] M. K. Lenka, A. Pandey, and A. Mittal, "Blind Deblurring Using GANs," arXiv preprint arXiv:1907.11880, 2019.
[27] S. Athar and Z. Wang, "A Comprehensive Performance Evaluation of Image Quality Assessment Algorithms," Ieee Access, vol. 7, pp. 140030-140070, 2019.
[28] Y. A. Al-Najjar and D. C. Soong, "Comparison of image quality assessment: PSNR, HVS, SSIM, UIQI," Int. J. Sci. Eng. Res, vol. 3, pp. 1-5, 2012.
[29] U. Sara, M. Akter, and M. S. Uddin, "Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study," Journal of Computer and Communications, vol. 7, pp. 8-18, 2019.
[30] G. Zhai and X. Min, "Perceptual image quality assessment: a survey," SCIENCE CHINA Information Sciences, vol. 63, p. 211301, 2020.
[31] L. Xiao, F. Heide, W. Heidrich, B. Schölkopf, and M. Hirsch, "Discriminative transfer learning for general image restoration," IEEE Transactions on Image Processing, vol. 27, pp. 4091-4104, 2018.
[32] S. Bosse, D. Maniry, K.-R. Müller, T. Wiegand, and W. Samek, "Deep neural networks for no-reference and full-reference image quality assessment," IEEE Transactions on Image Processing, vol. 27, pp. 206-219, 2017.
[33] D. Yang, V. Peltoketo, and J. Kämäräinen, "CNN-Based Cross-Dataset No-Reference Image Quality Assessment," in 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 2019, pp. 3913-3921.
[34] P. M. Delgado and J. Herre, "Can We Still Use PEAQ? A Performance Analysis of the ITU Standard for the Objective Assessment of Perceived Audio Quality," in 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX), 2020, pp. 1-6.
[35] G. Sharma, K. Umapathy, and S. Krishnan, "Trends in audio signal feature extraction methods," Applied Acoustics, vol. 158, p. 107020, 2020.
[36] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, M.-H. Yang, et al., "CycleISP: Real Image Restoration via Improved Data Synthesis," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 2696-2705.
[37] H.-X. Dou, H.-B. Li, Q.-Y. Fan, and Y.-C. Chen, "Signal Restoration Combining Modified Tikhonov Regularization and Preconditioning Technology," IEEE Access, vol. 5, pp. 24275-24283, 2017.
[38] C. Murthy, M. Kurian, and H. Guruprasad, "Performance evaluation of image restoration methods for comparative analysis with and without noise," in 2015 International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), 2015, pp. 282-287.
[39] A. K. Patel, N. Muchhal, and R. Yadav, "Method for image restoration using wavelet based image fusion," International Journal of Computer Applications, vol. 39, pp. 18-23, 2012.
[40] M. Kathale and A. Deshpande, "Image Restoration Using Two Statistical Modeling Methods," International Journal of Science and Research (IJSR), vol. 5, pp. 82–985, 2016.
[41] C. Khare and K. K. Nagwanshi, "Image restoration technique with non linear filter," International Journal of Advanced Science and Technology, vol. 39, pp. 67-74, 2012.
[42] S. Mishra and M. T. Sharma, "Image restoration technique for fog degraded image," International Journal of Computer Trends and Technology (IJCTT), vol. 18, pp. 208-213, 2014.
[43] P. Baruah and K. K. Sarma, "Dynamic Wavelet Thresholding based Image Restoration," International Journal of Computer Applications, vol. 74, pp. 24–29, 2013.
[44] F. Altinel, M. Ozay, and T. Okatani, "Deep structured energy-based image inpainting," in 2018 24th International Conference on Pattern Recognition (ICPR), 2018, pp. 423-428.

[45] H. Lin, V. Hosu, and D. Saupe, "KADID-10k: A large-scale artificially distorted IQA database," in 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), 2019, pp. 1-3.
[46] Y. Li, B. Gfeller, M. Tagliasacchi, and D. Roblek, "Learning to denoise historical music," arXiv preprint arXiv:2008.02027, 2020.
[47] T.-A. Hsieh, H.-M. Wang, X. Lu, and Y. Tsao, "WaveCRN: An Efficient Convolutional Recurrent Neural Network for End-to-end Speech Enhancement," arXiv preprint arXiv:2004.04098, 2020.
[48] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "Deblurgan: Blind motion deblurring using conditional adversarial networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 8183-8192.
[49] O. Kupyn, T. Martyniuk, J. Wu, and Z. Wang, "Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better," in Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 8878-8887.
[50] J. Peng, X. Wu, X. Yang, Y. Huang, B. Cai, and D. Cai, "Astronomical image restoration and point spread function estimation with deep neural networks," in Advances in Optical Astronomical Instrumentation 2019, 2020, p. 112030Q.
[51] Y. Zhou, D. Ren, N. Emerton, S. Lim, and T. Large, "Image restoration for under-display camera," arXiv preprint arXiv:2003.04857, 2020.
[52] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1125-1134.
[53] R. Köhler, M. Hirsch, B. Mohler, B. Schölkopf, and S. Harmeling, "Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database," in European conference on computer vision, 2012, pp. 27-40.
[54] S. Nah, T. Hyun Kim, and K. Mu Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 3883-3891.
[55] X. Liu, M. Suganuma, Z. Sun, and T. Okatani, "Dual residual networks leveraging the potential of paired operations for image restoration," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 7007-7016.
[56] T. Hyun Kim and K. Mu Lee, "Segmentation-free dynamic scene deblurring," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2766-2773.
[57] C. Veaux, J. Yamagishi, and S. King, "The voice bank corpus: Design, collection and data analysis of a large regional accent speech database," in 2013 International Conference Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE), 2013, pp. 1-4.
[58] S. Pascual, A. Bonafonte, and J. Serra, "SEGAN: Speech enhancement generative adversarial network," arXiv preprint arXiv:1703.09452, 2017.
[59] H.-G. Hirsch and D. Pearce, "The Aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions," in ASR2000-Automatic speech recognition: challenges for the new Millenium ISCA tutorial and research workshop (ITRW), 2000.
[60] N. Saleem, M. I. Khattak, and A. B. Qazi, "Supervised speech enhancement based on deep neural network," Journal of Intelligent & Fuzzy Systems, vol. 37, pp. 5187-5201, 2019.

Authors

Omar Hatif Mohammed received the B.Sc. and M.Sc. degree in Computer Engineering Technology from Northern Technical University, Iraq, in 2011 and 2015. He is currently a Ph.D. candidate in computer engineering at the University of Mosul, Iraq. His researches interests include pattern recognition, image restoration, and signal processing.

Basil Sh. Mahmood was born in 1953 in Mosul/ Iraq, graduated in 1976 from the University of Mosul/ Electrical department, and the M.Sc. degree in Electronics and Communications in 1979. Then he joined in Computer Center of the same university as an assistant lecturer then after he got the degree of Ph.D. On microprocessors architecture in 1996. Now, he is a microprocessors and computer architecture professor in the Computer Engineering Department/the University of Mosul. He published with others four books and more than 50 research papers in many journals and conferences. He supervised more than 22 M.Sc. and 14 Ph.D. Students. His interests are in microprocessors, computer architectures, image and signal processing, modern methods of Artificial Intelligence. He awarded many prizes and Medals.