Available online at www.ijpe-online.com

International Journal of Performability Engineering

vol. 18, no. 12, December 2022, pp. 885-892
DOI: 10.23940/ijpe.22.12.p6.885892

CNN and PCA in Image Fusion: A Comparative Statistical Analysis

Ashi Agarwal*

ABES Engineering College, Ghaziabad, 201009, India

Abstract

In technical terms, an image is the visual representation of a thing or person created using optical technology (such as a mirror or lens) or a technological apparatus. 2 distinct domains of multiresolution image fusion, i.e., PCA and CNN, are discussed in this study. Multiresolution image fusion techniques amalgamate more than two images covering optical unclear and blurred parts to produce an image covering all the focused areas or information. Based on the study, in both PCA and CNN. PCA is more straightforward among all image fusion approaches; meanwhile, according to the study conducted in this paper, it produces less effective results. On the other hand, CNN gives more effective results, but it is complex to handle. Also, the boundary pixels of the fused image has some mismatching problems, i.e., unrecognizable pixels. The effectiveness of the results is measured based on some statistical image quality parameters.

Keywords: CNN; image fusion; multi-focus image fusion; mean square error; PCA; PSNR; SNR

(Submitted on September 20, 2022; Revised on October 16, 2022; Accepted on November 12, 2022)

1. Introduction

An image is a 2D, 3D, or multidimensional representation of a person or any object, and an image can be analog or digital. In today's scenario, digital illustration is trending so very much. A digital image is a collection of pixels that can be handled or stored by the digital computer; a digital image is nothing but visual perception of something in binary numbers, i.e., 0,1. Every digital image is of different quality as some are different from angles; some are different in resolution based on features, etc. An image fusion process is required to have an image that focuses on each part effectively, i.e., to preserve all the features' information and quality. The image fusion process takes photos from the same perspective under various conditions and settings of the optical device as input and produces a final informative fused image [1]. A greater view of human or automated perception is provided by the fused image [2]. In terms of performance metrics like peak signal-to-noise ratio (PSNR), normalised correlation (NC), and mean square error, using the simple primitive technique will not result in the recovery of an outstanding fused image (MSE).

There are three levels of image fusion algorithms: low, middle, and high, or pixel, feature, and decision levels [3]. Pixel/Data-based image fusion blends the preliminary data of several source images to form a distinguished image, which is anticipated to be more elucidative than each input or source image data. Feature-level image fusion extracts numerous features, like edges, color, lines, etc., from distinctive image data sources and, at the moment, merges them into one or more feature maps that may replace the original data for further processing. Decision-level image fusion comprises blending data at a most significant level of deliberation [4], consolidating the outcomes from different calculations to yield a finishing up of the fused image. Key images are treated for the process separately for in-order extraction. The attained information is, in that case, merged by applying evaluation regulations to fortify the usual interpretation.

The principle component analysis (PCA) is a statistical technique that separates multiple associated variables into several unrelated variables. We should pick the main components from the MultiSpectral image bands to sort out this approach. From that point forward, the primary standard module, which contains the most picture data, is subbed by the panchromatic picture. Lastly, the inverse principal component transform is prepared to understand a multispectral image's new RGB color bands with the help of principal components. It reveals the internal structure of data unbiasedly [5].

* Corresponding author.
E-mail address: ashiagarwal31@gmail.com

CNN is a kind of neural network that incorporates the grouping of pictures during handling, it takes a shot at a layering premise, and CNN includes more than one layer, including concealed layers. Primarily, an input image is provided into the convolution layer. After choosing the parameters, filters are applied with strides and padding whenever required, then the convolution of the image is performed, and the ReLU application is applied to the matrix. Then pooling is performed to decrease the dimensionality size. We can add as many layers as we want until we got satisfied with the performance. Whenever satisfaction with layers, the output is flattened and fed into the FCL – Fully Connected Layer. In the end, an activation function is used for the output class [6-8]. Image fusion has its wide spectrum of application like medical diagnosis, pattern recognition, security purpose, steganography [9-13] in order to provide a user with multiple quality parameters. Wearable devices are also in tread now-a-days that help humans in their day to day work and to complete their assignments on time, these wearables also act as an health monitor to help humans in maintaining their health as well [14, 15], remembering the quote "Health Is Wealth".

2. Related Work

In recent years many researchers have found many fields of an image to be research for better fusion method. Due to the importance of multi-sensor data in many areas, such as remote sensing, medical and military imaging applications, picture fusion has become popular in the field of research. This review process provides the way to go through study well and to know the fact in depth Li Yuan et al. [16], in this paper, the researcher used a wavelet for multiscale atomization of the source and the conclusive images to be fused to acquire mutually high and low-frequency image sets. A profound CNN is utilized to experience the dictate aligning between the low and high-frequency pictures of the initial and final data to get a more informative and clearer fused image. The researcher uses high and low-frequency pictures to point out both convolutional networks. According to the researcher, the experiment shows that the proposed method can procure a conclusive fused image, which is superior and better on the basis of different evaluation criteria carried out. Bosse et al. [17] demonstrate an approach based on deep neural networks to image quality assessment with its new adapting features like IQA settings, use in both no-reference and full-reference, and its blended learning of local quality and local weights. The experiment was evaluated on LIVE, CISQ, and TID 2013.

R. Riyahia et al. [18] Stated that the image fused by PCA has the best spectral reliability as the PCA approach perpetuates (conserves) spatial and spectral data of the attributes of the original image. It is better than the results generated by either the Brovey or Ehlers methods. Deepak Kumar Sahu et al. [19] reviewed numerous image fusion methods like primitive fusion, DWT-based fusion, PCA-based fusion, etc. The Results prepared by the researcher state that spatial domain image fusion offers high spatial resolution along with the image unfocusing problem. The DWT is a suitable method for image fusion by providing polished spectral content. But according to the author, both qualities must be possessed by a good final image, so the fusion of the DWT & spatial domain fusion method elevates the process performance rather than using DWT and PCA algorithm individually. Nahvi, N et al. [20] discuss the application of the DWT method on numerous images to produce a final, conclusive idea with nearly all the relevant information content. Many other methods like PCA, HIS, and HPF are also discussed. The researcher proposed a new calculative calculation using distinctive fusion methods like pixel averaging, min-max, and max-min methods for clinical image fusion.

Bedi S. et al. [3] discuss three image fusion algorithms; primary fusion, DWT, and pyramid fusion. It gives finding about the past related works at a portion of the current fusion techniques like primitive, DWT, PCA-based fusion, etc. The paper expounds on various calculations and their similar investigation together. Li, S. et al. [21] depict a counterfeit neural system using the picture set with this multi-focus image fusion issue. Exploratory outcomes show that the proposed strategy outflanks the discrete wavelet transform methodology, especially once the development in the objects or mis-registration of the input images is developed. All things consider, starter tests propose that the fusion result isn't precisely delicate to the image block size, given that its worth isn't excessively outrageous. Huang W. et al. [2] introduced a new multi-center image fusion method utilizing PCNN. The vitality of picture Laplacian is used to evaluate the lucidity of the image array. The Proposed technique performs based on PCNN that needs no training. The researcher uses a recursive method to acquire appropriate worth. The researcher uses an iterative approach to obtain good results. Test results show that the proposed technique beats the DWT, Li, and Miao methods. Several researchers in various fields have become interested in fuzzy logic technique [22].

Dhivya R. et al. [23] created a technique based on fuzzy logic, which is implied for disclosing the edges of the image that is deprived of ruling the threshold value. This method begins with articulating an image by a channel of unpredictable 3 x 3 matrices. The conclusive result of the proposed method provides an undeviating effect in the lines'. This adaptive filtering technique is used in charge to remove the noise and gives better results for PSNR and MSE than the existing ones. Satellite images have become the most desirable framework for the fusion process due to its scalability and broader application for human welfare and contribution to the scientific world [24]. Zhijun Wang et al. [25] compared various image fusion techniques in their paper. They proposed a framework called the GIF method, which preserves the ratio between the respective bands, upholds the radiometric unification of the data, increases the spatial resolution, and gives more emphasis to dainty signature variations.

3. Proposed Methodology

Image fusion is intended to exploit corresponding information in multiresolution images to construct a specific composite image with the enhanced or amplified order content. The methodology is completely elaborated in Figure 1 given below

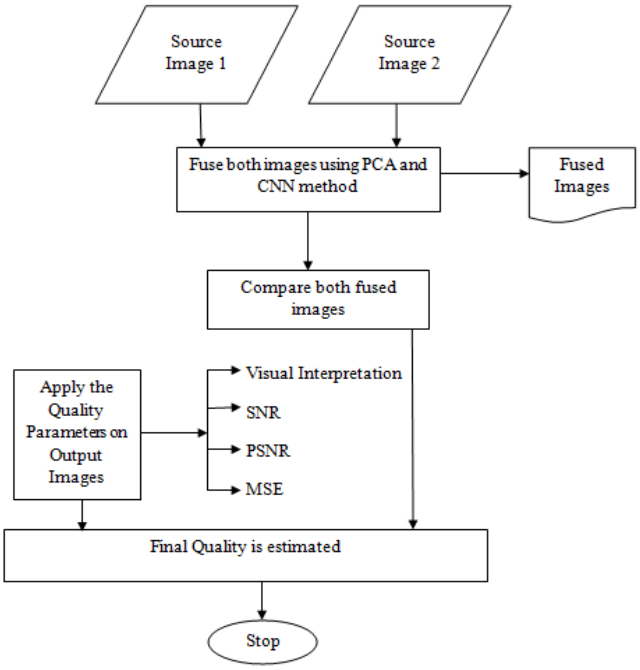


Figure 1. Applied Methodology

Source images taken for fusion must be the same size; PCA and CNN algorithm is then applied separately to the source image as Figure 2. PCA develops an uncorrelated feature space for images that can be used for further analysis instead of the original multispectral one. PCA algorithm is as follows:

I. Settle the data into a column vector of 2*n.
II. Pick all the possible principal components of the multispectral band.
III. Calculate the covariance of the vector matrix.
IV. Calculate Eigenvector and Eigenvalues.
V. Choose the highest covariance, and the first principal component is taken along with it.
VI. All the following principal components lie in the subspace of the prior PC.

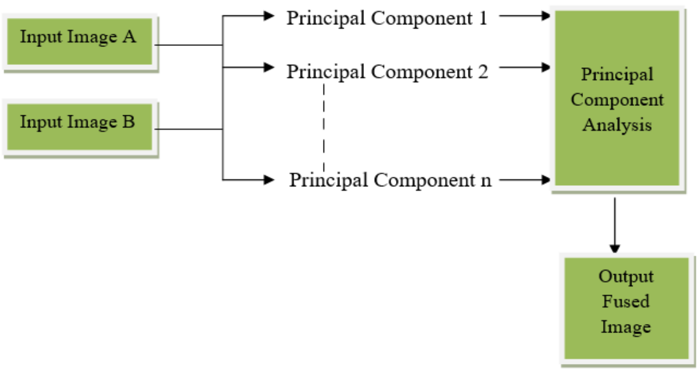


Figure 2. PCA Image Fusion

As Figure 3, CNN is a method used for decision-making in critical image-related issues by focusing and extracting high-level feature and frequency details of the source images and then providing a concluded, final picture giving a decision about the information used in embodiments [26, 27].

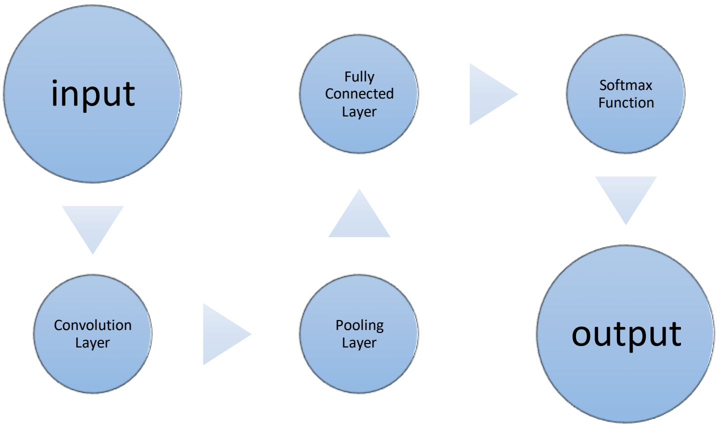


Figure 3. Layered Architecture of Convolution Neural Network

The algorithm for CNN image fusion is:

I. The chosen source image must be identical in size.
II. Input images are then fed into convolution layer: Convolution layer contains filters for the initial filtering of source images.
III. Convolution output is then fed into the pooling layer: the pooling layer is then reduced to the dimension size produced by the convolution layer and used to calculate position parameters using pooling functions such as max pooling, min pooling, and mean pooling.
IV. The output of the pooling layer is the forward to Fully Connected layer: This layer then classifies the images into digital vision features, which are then analyzed separately.
V. At last, the output is processed by the Softmax function: This function, also known as the normalization function, takes input from the fully connected layer and normalizes the input components into probability components within the interval (0,1).

4. Quality Assessment

The process of gathering and analysing data in order to determine how closely something adheres to predetermined standards and criteria is known as quality assessment. This paper discusses four classes of image quality assessment (IQA) algorithms: Visual Interpretation, SNR, Peak-SNR, and MSE. Each technique approaches the IQA problem from a different frame of reference and uses other initial principles. Subjective and objective both methods are used for the quality assessment in the fusion process in different applicable scenarios. Objective method is best suited for the dark images and infrared images or the images in the night mode [28, 29].

4.1. Visual Interpretation

As per prior assessment rules, individual judgment, or even evaluation, realize how to ensure the eminence of the image. The interpreter investigates the final fused image's differentiations, immersion, sharpness, and texture. It is effortless to infer however depends enormously upon spectator information and can not be represented by a mathematical model.

4.2. SNR-Signal to Noise Ratio

SNR is the amount of sign quality compared with added noise. It is generally estimated in decibels (dB). If the noise and signal strength are in microvolts is Vs. and Vn, respectively, then the SNR is calculated using Equation (1):

$$S/N = 20 \log 10(Vs/Vn) \tag{1}$$

There are three conditions to calculate SNR:

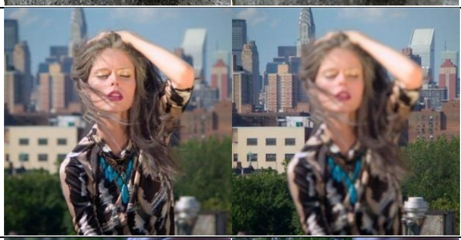
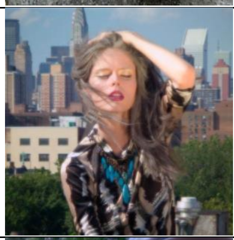
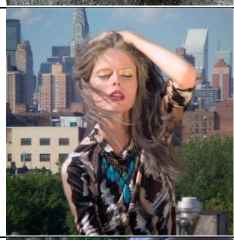
1) $If\ \ Vs = Vn\ then\ SNR = 0$

2) It is the ideal condition where,

$$if\ Vs > Vn, high\ SNR\ is\ positive$$

3) $If\ \ Vs < Vn\ , low\ SNR\ is\ negative.$

Table 1. Visual representation of results

Fig.	Input mage	Output of PCA	Output of CNN
1			
2			
3			
4			
5			

4.3. PSNR-Peak Signal to Noise Ratio

PSNR is used to compute the peak SNR between two images; it is also calculated in decibels. The quality of the final image is directly proportional to the PSNR value; if the PSNR value is high, the quality of the final image also hikes or vice versa. To calculate the PSNR value, first, we have to calculate MSE from Equation (2):

$$MSE= \Sigma M,N\ [I1\ (m,n) - I2\ (m,n)]\ 2\ /\ (M*N) \tag{2}$$

Here in the equation: M and N are the numbers of rows and columns in the source images.
PSNR is then calculated by Equation (3):

$$PSNR=10log10(R2MSE) \tag{3}$$

Here, R is the highest deviation in the source image

4.4. MSE-Mean Square Error

MSE is the comparison parameter to determine the quality of an image. It represents the aggregate squared error between the two images. MSE plays a vital role in the calculation of PSNR value. MSE value is directly proportional to the error presented in the image Equation (4). The less the MSE, the less the error. A lower MSE value represents a high-quality image.

$$MSE = \Sigma M,N\ [I1\ (m,n) - I2\ (m,n)]\ 2 / (M*N) \quad (4)$$

5. Experimental Work and Analysis

An experiment is performed on five image blocks of different sizes. In all five blocks of images, two images of the same block have other blur parts. For comparison, evaluation is performed on the fused images with the help of the quality mentioned above assessment parameters.

Table 1 shows the fusion of input images into a single fused output image with all the informative parts that are not visually blurred.

Table 2. Statistical Calculation of SNR by PCA and CNN

Fig.	SNR (PCA)	SNR (CNN)
Fig 1.	11.8182	11.5172
Fig 2.	11.6429	11.0904
Fig 3.	14.2511	13.6077
Fig 4.	12.6876	12.4355
Fig 5.	12.7549	12.4799

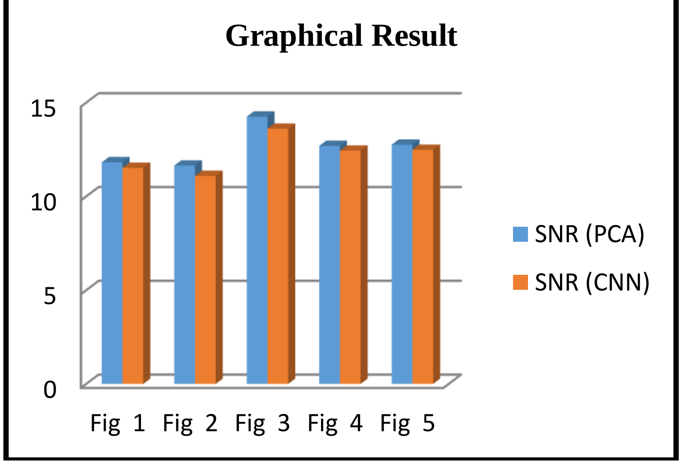


Table 2 exhibits the SNR values for all five image blocks using both PCA and CNN.

Table 3. Statistical Calculation of PSNR by PCA and CNN

Fig.	MSE (PCA)	MSE (CNN)
Fig 1.	65.4080	67.7141
Fig 2.	62.0299	66.1038
Fig 3.	49.4294	53.2296
Fig 4.	59.1780	60.9204
Fig 5.	58.7210	60.6102

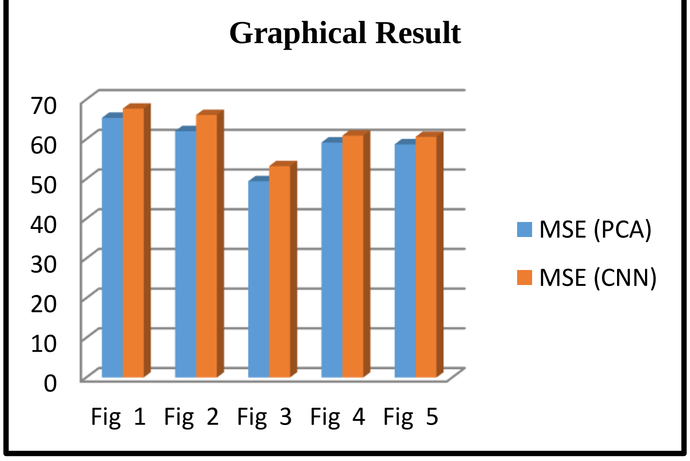


Table 3 shows the PSNR values for all five image blocks using both PCA and CNN.

Table 4. Statistical Calculation of MSE by PCA and CNN

Fig.	PSNR (PCA)	PSNR (CNN)
Fig 1.	18.5003	18.7636
Fig 2.	17.5594	17.6043
Fig 3.	17.5279	17.5695
Fig 4.	17.6046	17.6011
Fig 5.	17.6171	17.6241

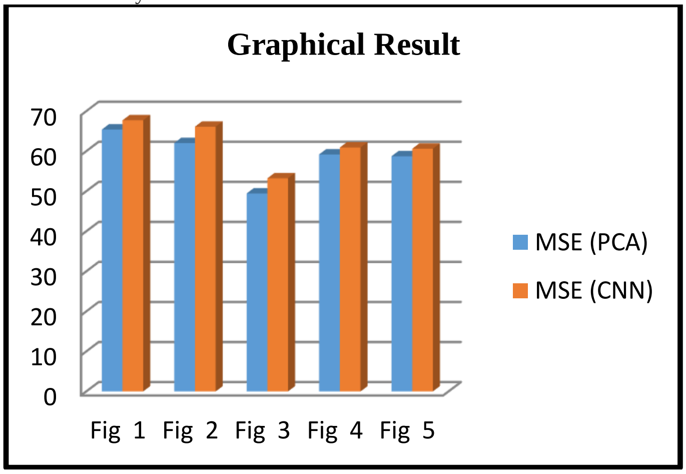


Table 4 shows the MSE values for the entire five image block using both PCA and CNN. Fusion results on PCA and CNN are shown in Table 1. After examining both the source and fused images, an apparent effect can be seen. Besides visual observation, all the statistical parameters like SNR, PSNR, and MSE show different mathematical values for all the five image blocks.

6. Conclusion

Image fusion is trending nowadays in every domain in our day-to-day life. After experimenting with the PCA and CNN fusion Techniques with different images, it demonstrates that the result produced by CNN is more informative and smoother than PCA, but the time consumption and cost are very high; also, there are some issues on the boundaries of fused images of CNN there present some unrecognized boundary pixels that need to be fixed in the future along with some time managing algorithms. PCA is simple and less time-consuming and has some blurring problems; hence, must be some quality enhancement technique for PCA in future work.

References

1. Sarah, S., Singh, V., Gourisaria, M.K., and Singh, P.K. Retinal Disease Detection using CNN through Optical Coherence Tomography Images. In *2021 5th International Conference on Information Systems and Computer Networks (ISCON)*, IEEE, pp. 1-7, 2021.
2. Huang, W. and Jing, Z. Multi-Focus Image Fusion using Pulse Coupled Neural Network. *Pattern Recognition Letters*, vol. 28, no. 9, pp.1123-1132, 2007.
3. Bedi, S.S. and Khandelwal, R. Comprehensive and Comparative Study of Image Fusion Techniques. *International Journal of Soft Computing and Engineering (IJSCE) ISSN*, vol. 3, no. 1, pp. 2231-2307, 2013.
4. Looney, D. and Mandic, D.P. Multiscale Image Fusion using Complex Extensions of EMD. *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp.1626-1630, 2009.
5. Gopmandal, F. and Pal, S. Color Image Fusion using Fuzzy Logic. *International Research Journal of Engineering and Technology (IRJET)*, vol. 2, 2015.
6. Zhang, K., Zuo, W., Chen, Y., Meng, D., and Zhang, L. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE transactions on image processing*, vol. 26, no. 7, pp. 3142-3155, 2017.
7. Jin, K.H., McCann, M.T., Froustey, E., and Unser, M. Deep Convolutional Neural Network for Inverse Problems in Imaging. *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509-4522, 2017.
8. Li, Y., Liu, D., Li, H., Li, L., Li, Z., and Wu, F. Learning a Convolutional Neural Network for Image Compact-Resolution. *IEEE Transactions on Image Processing*, vol. 28, no. 3, pp. 1092-1107, 2018.
9. KoteswaraRao, K. and Swamy, K.V. Multimodal Medical Image Fusion using NSCT and DWT Fusion Frame Work. *Int. J. Innov. Technol. Explor. Eng*, vol. 9, no. 2, pp. 3643-3648, 2019.
10. Lyu, K., Li, Y., and Zhang, Z. Attention-Aware Multi-Task Convolutional Neural Networks. *IEEE Transactions on Image Processing*, vol. 29, pp. 1867-1878, 2019.
11. Yang, Y., Wu, J., Huang, S., Fang, Y., Lin, P., and Que, Y. Multimodal Medical Image Fusion based on Fuzzy Discrimination with Structural Patch Decomposition. *IEEE journal of biomedical and health informatics*, vol. 23, no. 4, pp.1647-1660, 2018.
12. HR, R.U. and Sujatha, B.K. Fine Grained Medical Image Fusion using Type-2 Fuzzy Logic. *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 14, no. 2, 2019.
13. Pandey, B.K., Pandey, D., and Agarwal, A. Encrypted Information Transmission by Enhanced Steganography and Image Transformation. *International Journal of Distributed Artificial Intelligence (IJDAI)*, vol. 14, no. 1, pp. 1-14, 2022.
14. Sharma, P.K., Srivastava, A., and Perti, A. Novel Idea for Real-Time Health Monitoring using Wearable Devices. *International Journal of Mechanical Engineering and Technology (IJMET)*, vol. 9, no. 13, pp. 213-216, 2018.

15. Nagar, S. and Jain, M. Neural Network Techniques in Medical Image Processing. *Advances in Computational Intelligence and Communication Technology*, pp. 469-478, 2021.
16. Li, J., Yuan, G., and Fan, H. Multifocus Image Fusion using Wavelet-Domain-Based Deep CNN. *Computational intelligence and neuroscience*, 2019.
17. Bosse, S., Maniry, D., Müller, K.R., Wiegand, T., and Samek, W. Deep Neural Networks for No-Reference and Full-Reference Image Quality Assessment. *IEEE Transactions on image processing*, vol. 27, no. 1, pp. 206-219, 2017.
18. Riyahi, R., Kleinn, C., and Fuchs, H. Comparison of Different Image Fusion Techniques for Individual Tree Crown Identification using Quick Bird Images, 2009.
19. Sahu, D.K. and Parsai, M.P. Different Image Fusion Techniques–A Critical Review. *International Journal of Modern Engineering Research (IJMER)*, vol. 2, no. 5, pp. 4298-4301, 2012.
20. Nahvi, N. and Sharma, O.C. Comparative Analysis of Various Image Fusion Techniques for Biomedical Images: A Review. *Engineering Research and Applications*, vol. 4, no. 5, pp. 81-86, 2014.
21. Li, S., Kwok, J.T., and Wang, Y. Multifocus Image Fusion using Artificial Neural Networks. *Pattern Recognition Letters*, *23*(8), pp.985-997, 2002.
22. Bajpai, A. and Kushwah, V.S. Importance of Fuzzy Logic and Application Areas in Engineering Research. *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 7, pp. 1467-1471, 2019.
23. Dhivya, R. and Prakash, R. Edge Detection of Satellite Image using Fuzzy Logic. *Cluster Computing*, vol. 22, no. 5, pp. 11891-11898, 2019.
24. Jain, M. and Sinha, A. A Framework to Classify the Satellite Images. Int. J. Comp. Sci. Inf. Techn., vol. 7, no. 1, pp. 71-73, 2016.
25. Wang, Z., Ziou, D., Armenakis, C., Li, D., and Li, Q. A Comparative Analysis of Image Fusion Methods. *IEEE transactions on geoscience and remote sensing*, vol. 43, no. 6, pp. 1391-1402, 2005.
26. Rastogi, R., Upadhyay, H., Rastogi, A.R., Sharma, D., Bishnoi, P., Kumar, A., and Tyagi, A. Knowledge Extraction in Digit Recognition using MNIST Dataset: Evolution in Handwriting Analysis. *International Journal of Knowledge Management (IJKM)*, vol. 17, no. 4, pp. 52-75, 2021.
27. Dubey, G. Segmentation and Detection of Road Region in Aerial Images using Hybrid CNN-Random Field Algorithm. In *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, IEEE, pp. 502-506, 2020.
28. Liu, Z., Blasch, E., Xue, Z., Zhao, J., Laganiere, R., and Wu, W. Objective Assessment of Multiresolution Image Fusion Algorithms for Context Enhancement in Night Vision: A Comparative Study. *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 1, pp. 94-109, 2011.
29. Gupta, M. A Fusion of Visible and Infrared Images for Victim Detection. In *High Performance Vision Intelligence*, Springer, Singapore, pp. 171-183, 2020.