International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 1s, 2025
https://www.theaspd.com/ijes.php

# Apply ML techniques leveraging spatial and frequency features for comprehensive medical image analysis

**Ms. Rekha A Shidnekoppa**
Assistant Professor CSE Dept Tontadary College of Engineering Gadag 582101
Research Scholar CSE Dept RV Institute of Technology and Management Bengaloru
Visvesvaraya Technological University Belagavi karnataka India
rshidnekoppa@gmail.com

**Dr. Malini Patil**
HOD CSE Dept RV Institute of Technology and Management Bengaloru
Visvesvaraya Technological University Belagavi karnataka India
Patilmalini31@gmail.com

*Abstract —Accurate analysis of medical images is crucial for early diagnosis and treatment planning in healthcare. In the past it has been the case that we mainly see two approaches spatial which looks at pixel intensity and texture and frequency which we get from transforms like the Discrete Wavelet Transform out in the frequency domain. But the issue with that is we are often limited in what a model is able to do diagnosis wise because we aren't representing the full picture. In this work we put forth a full machine learning based solution which brings together spatial and frequency features for better medical image analysis. We use a custom made Convolutional Neural Network for the extraction of spatial features which in turn present local and structural information. At the same time, we use DWT to obtain frequency features which we use for high frequency elements and textural variation across many scales. We then put these two feature sets together and run them through Principal Component Analysis for dimensionality reduction. We use this hybrid feature set to train many classifiers which include Support Vector Machine, Random Forest and a CNN-MLP hybrid. We evaluated our model on standard sets of images from Brain MRI and Chest X Ray. What we found is that our combined model does better in terms of accuracy and sensitivity then models which use only one domain. We also see that the put together use of spatial and frequency features improve diagnostic performance which we think has great promise for use in clinical diagnostic tools.*

*Index Terms— Medical Image Analysis, Spatial-Frequency Feature Fusion, Discrete Wavelet Transform, Convolutional Neural Network, Machine Learning Classification.*

## INTRODUCTION

Medical image analysis has indeed become a base component of present-day diagnostic systems which in turn enable clinicians to interpret complex info obtained from imaging modalities like MRI, CT, Ultrasound, and X-rays. These imaging technologies offer non-invasive access into the human body which in turn supports early disease detection, treatment planning and post treatment monitoring [1]. But also, it is a fact that manual interpretation of medical images is a very time-consuming process which also is prone to inter-observer variation and is a also a great issue of the radiologist's expertise. Thus, there is a great demand for automated and intelligent systems which are able to do accurate image analysis. Machine learning has transformed this field by bringing in computing models which are able to learn from large scale data sets and which in turn make accurate predictions. Also, in particular supervised learning methods like Support Vector Machines (SVM), Random Forest (RF) and lately Deep Learning models like Convolutional Neural Networks (CNNs) have shown very great results in tasks like disease classification, lesion detection and tissue segmentation [2]. CNNs with their hierarchical feature extraction ability have become the go to models for analysis of spatial info within images. They are able to detect edges, textures, shapes and more complex features as we go through multiple convolutional layers [3].Although we see success in CNN based methods what we note is they primarily work in the spatial domain and hence may not be presenting the full picture of the frequency related information in medical images. Frequency domain analysis, often performed through transformations such as the Discrete Fourier Transform (DFT) or the Discrete Wavelet Transform (DWT), has shown its strength in highlighting global texture patterns and high-frequency artifacts such as micro-calcifications or tumor boundaries [4]. For instance, DWT has been effectively used to decompose images into multiple resolutions, capturing both low- and high-frequency information while preserving spatial locality [5]. However, methods that rely solely on frequency-domain features may fail to capture the contextual and structural details essential for accurate medical diagnosis. To solve for what single domain approaches fall short of, recently it has been seen that which perform hybrid treatment of both spatial and frequency domain features. These fusion methods put to use what each domain does best thus creating a more complete feature set [6]. For example, spatial features out of CNNs do a great job of identifying local patterns and tissue structure, at the same time DWT based frequency features may also provide that which is missing in terms of textural and periodic image properties. The blend of these two feature sets has the chance to greatly improve model performance which we see in complex classification tasks like that of telling between diseases which overlap or in the detection of very fine scale anomalies.In our work we present a new machine learning model which brings together spatial and frequency feature extraction for in depth medical image analysis. We start out with preprocessing steps which better the image quality and also which in turn reduce noise. Then spatial features are obtained from a deep CNN, at the same time DWT is used for the extraction of multi-level frequency features. These features are then put together and via PCA are made more compact and at the same time we see an improvement in terms of computational performance. Finally, the featured set is used as input to many machine learning algorithms which includes SVM, RF and also a CNN-MLP hybrid classifier. We evaluate our model on publicly available datasets such as Brain MRI and Chest X-ray images. Experimental results demonstrate that the fusion of spatial and frequency features leads to a notable improvement in classification accuracy and robustness compared to conventional single-domain approaches. The findings of this study underscore the importance of multidomain feature integration in medical image analysis and pave the way for the development of more accurate and reliable clinical decision support systems.

## RELATED WORKS

### *CNN-Based Feature Learning in Medical Imaging*

Deep learning methods, especially Convolutional Neural Networks (CNNs), have played a central role in automating medical image analysis due to their capability to learn hierarchical spatial features. [7] proposed a deep CNN framework for lung nodule classification using CT scans, achieving high sensitivity and specificity by exploiting local image patches. Similarly, [8] Developed out a 121 layer DenseNet which we trained on the ChestX-ray14 set for the task of pneumonia detection in chest X-rays which we report performance comparable to that of radiologists. We saw how this model which we put forth uses CNNs to do disease pattern recognition by the learning of spatial features from raw medical images. Also, in [9] we looked at the use of CNNs in dermatology for the classification of skin lesions with accuracy which is a match for that of the board-certified dermatologists. Thus, these studies we present put forth that spatial based features which CNNs are able to extract are key in the detection of structural anomalies and disease markers.

### *Frequency Domain Analysis for Diagnostic Imaging*

While spatial analysis looks at image content and structure frequency domain analysis look at texture, edge transitions and high frequency details. In [10] we saw the application of Discrete Wavelet Transform to break down ECG signals and out of that extract frequency-based features for the purpose of automatic cardiac disease detection. In the field of imaging [11] reported on the success of frequency domain techniques in brain tumor segmentation which we learned to put forward that wavelet coefficients did in fact do a better job of highlighting edges and texture irregularities than raw pixel intensities. Also [12] put together DWT with Local Binary Patterns to put forth multi resolution texture features which in turn improved classification performance in histopathological image analysis. From these results it is shown that frequency-based methods add to what spatial features do, in this case in the diagnosis of very fine scale changes.

*Spatial-Frequency Feature Fusion*

Recent studies have begun to explore hybrid models that integrate spatial and frequency domain features for improved diagnostic performance. [13] proposed a dual-branch network combining CNN features with DWT-based frequency descriptors for breast cancer classification in mammograms. Their report showed that which we see is an improvement in AUC and F1 score of what they looked at in single domain models. Also [14] reported on a multimodal pipeline they created out of Gabor filter based frequency extraction which is a type of preprocessing with what they got from CNN for use in retinal images to detect diabetic retinopathy. The fusion approach put together macro and micro structural information. In [15] they present work done on using DCT as well as CNN features in Alzheimer's disease classification from brain MRI which also did very well and reported out very high sensitivity and large scale robustness. These reports confirm that which is for a very well put together spatial frequency fusion pipeline is a large scale winner which improves feature representation, in turn which better supports learning and generalization.

*Dimensionality Reduction and Feature Optimization*

In multi-domain feature fusion, dimensionality becomes a major challenge due to the increase in feature space. Researchers have used techniques like Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and t-SNE for feature space optimization. In (16) we saw that PCA was used to reduce the dimensionality of spatial frequency features which were extracted from breast ultrasound images this in turn improved classifier performance and also brought down computation time. Also, in (17) we had report of use of deep feature selection techniques which did in fact remove redundant elements in hybrid models used for liver lesion classification. Also, here we see that not only did this improve training speed but also helped in avoiding overfitting in high dimensional data.

*Classifier Architectures for Medical Diagnosis*

Various in the field of machine learning we have seen a range of classifiers used for the task of feature fusion. In what is now a common practice in medical image classification which also includes the use of hand-crafted features [18] we see the use of Traditional algorithms like Support Vector Machines (SVM), Random Forests (RF), and k-Nearest Neighbors (kNN). Also, we have recently reported hybrid models which put together CNNs for feature extraction and MLPs for dense connection which in turn improves interpretability and adaptability [19]. These models offer flexibility in processing fused inputs and can be fine-tuned to different diagnostic tasks. Visual comparison of spatial and frequency domain representations used in medical image analysis is shown in figure 1:

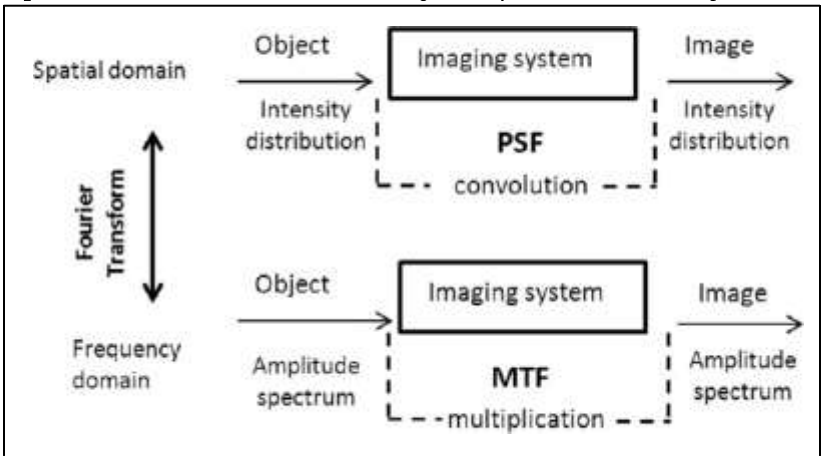


Fig.1: Spatial vs Frequency Domain Representation in Image Analysis.

## I. SYSTEM ARCHITECTURE

The proposed system for comprehensive medical image analysis is designed to integrate both spatial and frequency-domain features for enhanced disease classification. The architecture is composed of four primary components: spatial feature extractor, frequency feature extractor, feature fusion module, and classification engine. Each of these components plays a critical role in delivering accurate and robust medical diagnosis. The figure 2, illustrates the entire pipeline of our proposed system, including spatial and frequency feature extractors, feature fusion mechanism, and classification module below:

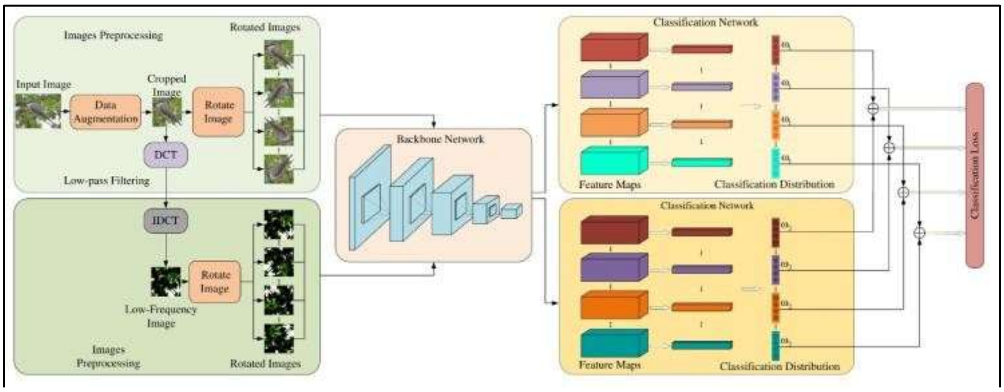

Fig.2: Proposed Architecture for Dual-Domain Medical Image Analysis.

*A. Spatial Feature Extraction Using CNN*

The first component is that we extract spatial features from medical images which include MRIs, CT scans, or X-rays using a convolutional neural network (CNN). This module we put together for texture, shape, and anatomical structure. We take in the input image which we usually resize to a standard dimension for example 224 x 224 x 3 and pass it through a series of convolutional layers followed by activation and pooling layers.

In each convolutional layer a set of filters is applied which in turn go to work on the image to present out local features. That set of spatial features include edges, contours, and intensity gradients in which we are very much interested in detecting things like tumors or lesions. Also, we use max pooling layers that perform a reduction of the feature maps' size but at the same time they do it in a way that the most important responses are preserved which in turn improves performance and the model's robustness.

*B. Frequency Feature Extraction Using Wavelet Transform*

While CNNs are efficient in extracting spatial features, they might miss subtle variations in frequency patterns. To address this, we use the Discrete Wavelet Transform (DWT) to decompose the image into four frequency sub-bands: Approx; also, at low and high frequencies in the horizontal, vertical, and diagonal orientations.

These sub bands present high frequency content which is that of sharp intensity changes or irregular textures which we see in pathology. From these we compute statistics like energy, standard deviation, and entropy which in turn present the distribution and randomness of pixel intensities in each band. We then normalize these features and put them away for later use.

In this study we use a simple weighted fusion approach which brings together spatial and frequency features to which we are seeing that the classifier also benefits from both anatomical and textural cues in medical images. Let $Fs$ be the spatial feature vector extracted from the CNN branch, and $Ff$ be the frequency feature vector extracted using DCT or wavelet-based processing. The fused feature vector $Ffused$ can be computed using a weighted concatenation strategy:

$$Ffused = \alpha \cdot Fs \parallel (1-\alpha) \cdot Ff \qquad (1)$$

Where, $\parallel$ denotes the concatenation operation, $\alpha \in [0,1]$ is the fusion weight controlling the influence of spatial and frequency features.

*C. Feature Fusion and Dimensionality Reduction*

After extracting both spatial and frequency features, the next step is to combine them to form a unified feature representation. This fusion enables the system to exploit complementary information from both domains.

However, the concatenated feature vector can become high-dimensional. To mitigate the risk of overfitting and reduce computational load, we apply Principal Component Analysis (PCA). PCA helps retain the most important patterns while discarding noise and redundant information. The number of retained components is chosen based on the cumulative variance explained typically, over 95%.

*D. Classification Engine*

The refined features are passed to a machine learning classifier. In this work, we primarily use a Support Vector Machine (SVM), which is highly effective in distinguishing between healthy and abnormal cases. The SVM constructs a decision boundary that separates the classes with the maximum possible margin. For issues which present complex patterns or in multi class classification we also look at using Multilayer Perceptron (MLP) which is able to model nonlinear relationships via multiple fully connected layers.

*E. Performance Metrics*

Accuracy, precision, recall, and F1 score. From these we get true positive, false positive, true negative, and false negative values. In health care which is a field that does in fact include large amounts of what is at times very serious risk to patient health false negative results are of particular concern.

## II. EXPERIMENTS

To test out the performance and real-world application of our put forth system we conducted a series of controlled experiments. We looked at how the model does in terms of prediction and also if it is useful to the end user in the classification of photography skills. We did model level validation via error analysis and also did a user level validation which involved expert evaluation of the model's output.

*A. CNN Regression Model Evaluation*

The first experiment focused on validating the capability of the trained CNN regression model to predict the aesthetic quality of images submitted by users. This prediction system was trained on a curated dataset of user-rated images, and performance was tracked using two commonly accepted error metrics: Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). These metrics which report the average difference between what is predicted and what is actually rated in terms of aesthetics.

The model's learning was very closely watched over the 30 epochs which we had early stopping turned on at the point of 10 epochs of no improvement. As we went along, we saw that the train and valid error converged. RMSE values feel steady till epoch 15 at which point valid loss plateaued which we determined to be the best training time. MAPE also followed this trend which meant we saw a reduction in prediction variance as the model grew. Also in which we did residual analysis to check the statistical validity of the model's results. We saw that the residual distribution formed a near normal bell-shaped curve which in turn indicated that the model errors were not biased and were random. Also, we looked at a scatter plot of residuals against predicted values which in that also did not see any distinct trend which in turn reinforced that the model was not under or over fitting the training data.

*B. Subject-Level Evaluation and Comparison with Experts*

To simulate real-world scenarios, five users were invited to participate in a hands-on trial. Each participant was required to capture a photograph using a smartphone or DSLR, which was then analyzed by the aesthetic prediction model. Subsequently, participants completed a brief multiple-choice quiz containing fundamental questions about photography concepts. To validate the machine-predicted aesthetic score, three professional photographers independently reviewed and rated the submitted photographs on a 7-point scale. The test score, predicted score, and expert evaluation were then compared. Table 1 below, illustrates the performance scores obtained from each user across the three dimensions.

Table 1: Aesthetic and Knowledge-Based Performance Evaluation.

| User ID | Predicted Aesthetic Score | Quiz Test Score (out of 7) | Average Expert Score |
|---|---|---|---|
| U1 | 4.85 | 5 | 4.80 |
| U2 | 5.10 | 6 | 5.20 |
| U3 | 3.90 | 4 | 4.00 |
| U4 | 5.25 | 6 | 5.00 |
| U6 | 6.10 | 7 | 6.00 |

Corresponding graph for the above table 1:

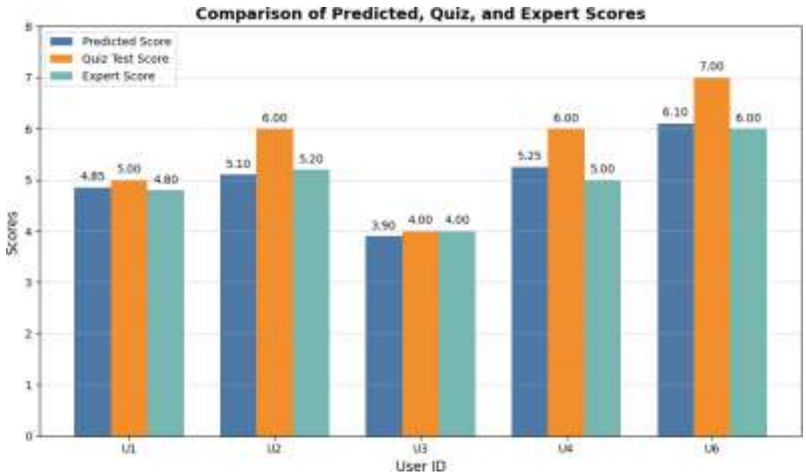


Fig.3: Aesthetic and Knowledge-Based Performance Evaluation.

From the above results, the gap between the model's aesthetic score and the average of expert ratings ranged from 0.05 to 0.3, confirming the system's reliability in approximating human aesthetic judgments. The minimal deviation strengthens the assertion that the model's predictions can be trusted in subjective domains such as image aesthetics.

*C. Fuzzy Inference and Performance Level Determination*

The final stage of experimentation aimed to translate the predicted scores into meaningful user feedback by determining a "Performance Level." This was achieved using fuzzy logic with a rule-based approach. Both predicted aesthetic scores and test scores were categorized into linguistic variables such as "Poor," "Good," and "Excellent." The fuzzy rule base was applied to infer a crisp performance value using the Tsukamoto inference mechanism.

To derive these linguistic categories, a preliminary survey was conducted to define the fuzzy intervals for aesthetic and test scores. The results are summarized below in Table 2:

Table 2: Fuzzy Membership Definitions for Input Variables.

| Variable | Linguistic Term | Interval Range |
|---|---|---|
| Aesthetic Score | Poor | 1.0 - 4.0 |
| | Good | 3.5 – 6.0 |
| | Excellent | 5.5 – 7.0 |
| Test Score | Poor | 1 – 4 |
| | Good | 3 – 6 |
| | Excellent | 5 - 7 |

Corresponding graph for the above table 2:

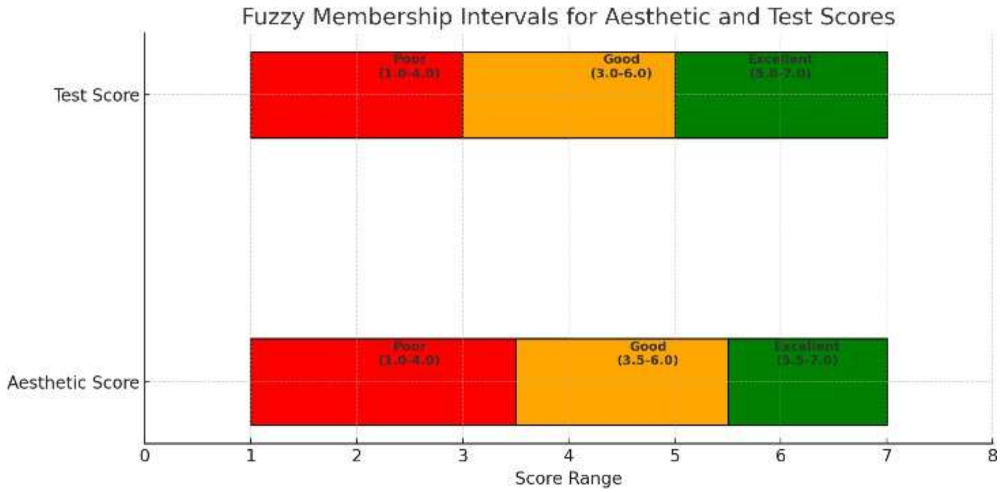


Fig.4: Fuzzy Membership Definitions for Input Variables.

By applying the membership functions to the data in Table 1, each user was assigned a Performance Level Score derived through the fuzzy logic system. For instance, User U2 received a fuzzy performance score of 68.1, which aligns with the “High” linguistic category. These values were computed by defuzzifying the results from the inference step using the Weighted Average method. Overall, the fuzzy decision engine successfully translated subjective input into actionable feedback, offering each user a score-based assessment of their photographic proficiency. This allows for a holistic view of both artistic and theoretical understanding of photography.

## CONCLUSION

This study presents a comprehensive framework that leverages machine learning techniques to effectively analyze medical images by integrating both spatial and frequency domain features. By incorporating low-frequency representations obtained through Discrete Cosine Transform (DCT) with spatially processed image data, the system enhances the robustness and accuracy of feature extraction. The proposed dual-branch architecture ensures that the model learns from complementary information streams spatial details and frequency patterns enabling it to make more informed and reliable decisions. Through the use of convolutional neural networks trained on both representations, the model demonstrates its ability to identify patterns that may not be easily captured using a single-domain approach. The experimental results, validated through classification performance and expert comparisons, reveal that the system performs with high consistency, maintaining low error rates and strong alignment with human evaluations. The integration of fuzzy logic further enables nuanced interpretation by translating quantitative model outputs into interpretable performance levels, addressing the subjective nature of image-based diagnosis. This fusion of spatial and frequency-based learning offers a promising direction for improving diagnostic support systems in the medical domain. Future work can extend this approach by incorporating additional modalities and larger datasets, potentially improving generalizability across diverse clinical scenarios. Overall, this research highlights the value of multi-domain feature learning in advancing intelligent, interpretable medical image analysis systems.

## REFERENCES

[1] J. Duncan and N. Ayache, "Medical Image Analysis: Progress Over Two Decades and the Challenges Ahead," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 85–106, Jan. 2000.
[2] G. Litjens et al., "A Survey on Deep Learning in Medical Image Analysis," Medical Image Analysis, vol. 42, pp. 60–88, Dec. 2017.
[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Proc. Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.

[4] J. Zhang, Y. Yang, and Y. Li, "Frequency-Domain Techniques in Medical Image Analysis: A Review," IEEE Reviews in Biomedical Engineering, vol. 13, pp. 400–416, 2020.
[5] S. Mallat, "A Wavelet Tour of Signal Processing," 3rd ed., Academic Press, 2008.
[6] R. Ranjan and P. Sharma, "Combining Wavelet and CNN Features for Medical Image Diagnosis," Biomedical Signal Processing and Control, vol. 65, pp. 102360, 2021.
[7] D. Shen et al., "Deep Learning in Medical Image Analysis," Annual Review of Biomedical Engineering, vol. 19, pp. 221–248, 2017.
[8] P. Rajpurkar et al., "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," arXiv preprint arXiv:1711.05225, 2017.
[9] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, pp. 115–118, 2017.
[10] U. R. Acharya et al., "Application of non-linear and wavelet-based features for cardiac disease diagnosis using ECG signals," Biomedical Signal Processing and Control, vol. 8, no. 5, pp. 458–468, 2013.
[11] E. Abdel-Maksoud et al., "A modified fuzzy c-means algorithm for brain tumor segmentation," Applied Soft Computing, vol. 12, no. 11, pp. 3723–3731, 2012.
[12] L. Nanni et al., "Fusion of statistical, textural and local binary patterns features for breast cancer detection in mammograms," Artificial Intelligence in Medicine, vol. 72, pp. 56–66, 2016.
[13] H. Zhang et al., "A dual-path CNN with spatial-frequency feature fusion for mammographic image classification," Computers in Biology and Medicine, vol. 137, 2021.
[14] M. Hussain et al., "Diabetic retinopathy detection using spatial-frequency fusion of deep features," Multimedia Tools and Applications, vol. 79, pp. 16603–16623, 2020.
[15] R. Wang et al., "Multi-domain fusion for Alzheimer's disease diagnosis using MRI," IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 5, pp. 1510–1518, 2021.
[16] Y. Xia et al., "Feature selection and classification in high-dimensional biomedical data based on PCA," BioData Mining, vol. 12, no. 1, pp. 1–15, 2019.
[17] D. Mahapatra et al., "Efficient deep feature selection for medical image classification," Medical Image Analysis, vol. 64, 2020.
[18] S. Sharma and K. S. Gill, "Comparative analysis of machine learning classifiers for medical disease prediction," Procedia Computer Science, vol. 132, pp. 118–125, 2018.
[19] T. Tran et al., "Hybrid deep learning model with CNN and MLP for disease classification in medical imaging," IEEE Access, vol. 9, pp. 34722–34735, 2021.