Enhancing X-ray Image Classification Accuracy Using Blockwise Frequency-Domain Feature Extraction

Tanuj Sarkar
Department of Computer Science and Engineering
SRM University AP
Email: tanujsarkar123@gmail.com

***Abstract*—Accurate and efficient classification of X-ray images is a critical task in medical diagnostics, often hindered by noise and subtle anatomical variations in spatial-domain representations. This paper presents a novel frequency-domain approach that leverages Fast Fourier Transform (FFT) combined with a blockwise feature extraction strategy to improve classification performance. Each X-ray image is preprocessed using histogram equalization to enhance structural contrast, followed by 2D FFT to transform the image into the frequency domain. Initially, statistical features such as energy, entropy, and dominant frequency are extracted from the global magnitude spectrum. To capture localized variations, the image is divided into a 4×4 grid, and features are computed independently from each block, resulting in a rich 64-dimensional feature vector per image. These features are then used to train traditional machine learning classifiers including Support Vector Machines (SVM), Random Forest, and Gradient Boosting. Experimental results on a publicly available chest X-ray dataset demonstrate that the proposed blockwise FFT-based feature extraction significantly outperforms the global approach, achieving up to 85.7% accuracy. The findings validate the effectiveness of localized frequency-domain analysis in enhancing model robustness and diagnostic reliability for real-time medical image classification.**

***Index Terms*—X-ray classification, Fast Fourier Transform (FFT), frequency-domain features, blockwise analysis, machine learning, medical imaging, image preprocessing, feature extraction, support vector machines, pneumonia detection.**

I. Introduction

Medical image classification plays a vital role in modern diagnostic systems, offering decision support for radiologists and enabling automated screening workflows. Among various imaging modalities, chest X-rays are particularly important due to their widespread use in diagnosing pulmonary conditions such as pneumonia. Despite recent advances, the accurate classification of X-ray images remains a challenging task, largely due to the presence of anatomical noise, low contrast, and subtle structural variations that may go undetected in conventional analysis.

Traditional approaches to image classification often rely on spatial-domain features—such as pixel intensity distributions or edge patterns—which can be sensitive to noise and illumination variations. These methods frequently struggle to capture deeper, latent patterns that exist within the image's structural frequency. Moreover, global descriptors tend to overlook localized anomalies that are critical in medical contexts.

To address these limitations, frequency-domain techniques have gained increasing attention. The Fast Fourier Transform (FFT), in particular, offers an effective means to analyze the frequency components of an image, emphasizing repetitive structural patterns while reducing sensitivity to noise and spatial distortions. However, standard applications of FFT extract features from the image as a whole, potentially missing localized diagnostic cues.

In this study, we propose a novel enhancement to frequency-domain analysis by introducing a *blockwise FFT-based feature extraction* strategy. Instead of extracting features from the entire image, we divide the X-ray image into non-overlapping blocks and compute frequency features within each region. This approach not only preserves the benefits of FFT but also captures local frequency variations—leading to richer, more discriminative representations. The resulting feature vectors are then used to train machine learning classifiers such as Support Vector Machines (SVM), Random Forest, and Gradient Boosting.

Our experimental results show that blockwise FFT features significantly outperform their global counterparts, providing a scalable and interpretable solution for real-time X-ray classification tasks.

II. Related Work

Medical image classification has been extensively explored using both conventional machine learning and deep learning techniques. Convolutional Neural Networks (CNNs) have shown state-of-the-art performance on large-scale imaging datasets, particularly when trained end-to-end on spatial pixel data. However, CNNs require extensive computational resources, large annotated datasets, and often lack interpretability—making them less ideal for low-resource or real-time diagnostic applications.

Alternatively, frequency-domain approaches, particularly those using the Fast Fourier Transform (FFT), offer a lightweight and interpretable mechanism for extracting discriminative features. Prior work has demonstrated the utility of FFT in highlighting structural differences in X-ray and MRI images [3], [4]. For instance, FFT-based spectral features have been used to classify retinal fundus images, yielding performance comparable to shallow neural models [4]. Other studies have integrated frequency-based descriptors with texture analysis for lung abnormality detection in CT scans [5].

Despite these successes, most FFT-based methods rely on global transformations and ignore spatial locality—an aspect that is often critical in detecting localized pathologies such as small opacities or consolidations in lung X-rays. Moreover, traditional approaches extract a limited set of spectral statistics, which may constrain the expressiveness of the feature space.

Our work addresses these gaps by introducing a blockwise FFT feature extraction framework that captures both global and localized frequency variations. This method strikes a balance between computational efficiency and classification performance, particularly when paired with classical machine learning classifiers such as SVM, Random Forest, and Gradient Boosting. To our knowledge, this is among the first works to combine localized frequency-domain feature engineering with conventional machine learning in the context of X-ray image classification.

III. Proposed Methodology

A. Dataset and Preprocessing

The dataset used in this study is sourced from Kaggle's open repository titled *Detecting Pneumonia in X-ray Images*[1]. It consists of a total of 5,863 chest X-ray images, comprising both normal and pneumonia cases.

the implementation of the given work can be found in [2].

Although the images were originally in grayscale, grayscale conversion was applied to ensure consistency across all samples. Histogram Equalization was then used to enhance the image contrast, thereby improving the visibility of subtle structural patterns critical for downstream frequency analysis. Figure 1 demonstrates the contrast improvement post equalization.

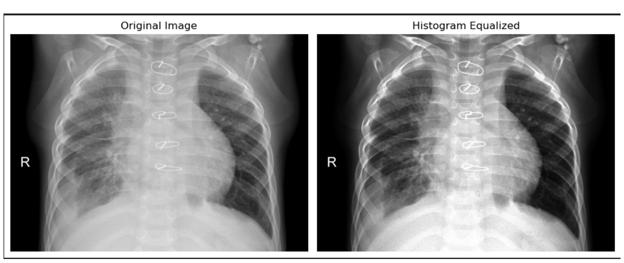

Fig. 1: Effect of Histogram Equalization on X-ray Image Contrast

B. FFT and Feature Extraction

To move beyond the limitations of spatial domain analysis, we employ the 2D Fast Fourier Transform (FFT) to convert each X-ray image into the frequency domain.

1) 2D FFT and Magnitude Spectrum: The 2D FFT of an image $f(x, y)$ is defined as:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cdot e^{-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)}$$

The magnitude spectrum $|F(u, v)|$ is computed as:

$$|F(u, v)| = \sqrt{\Re(F(u, v))^2 + \Im(F(u, v))^2}$$

2) Global Feature Extraction: From the global magnitude spectrum of each image, we extract three statistical features:

- **Energy:** Measures the total power in the frequency domain.
$$E = \sum_{u,v} |F(u, v)|^2$$
- **Entropy:** Captures the randomness or complexity of the frequency components.
$$H = -\sum_{i} p_i \log(p_i)$$
where p_i is the normalized magnitude value at position i.
- **Dominant Frequency:** Identifies the peak frequency with the highest amplitude in the magnitude spectrum.

3) Blockwise Feature Extraction: To capture localized structural variations, each X-ray image is divided into a 4×4 non-overlapping grid, yielding 16 blocks. The same three features (energy, entropy, and dominant frequency) are extracted from each block, resulting in a 64-dimensional feature vector.

Fig. 2: Blockwise division of X-ray image into 16 regions

Compared to the 4-dimensional global feature vector, the blockwise strategy enables finer regional analysis and significantly improves the model's sensitivity to localized anomalies.

C. Classification

The extracted features are used to train and evaluate three machine learning classifiers:

- Support Vector Machine (SVM)
- Random Forest
- Gradient Boosting

[1]https://www.kaggle.com/code/paultimothymooney/detecting-pneumonia-in-x-ray-images

[2]https://www.github.com/7anuj/blockwise_fft

TABLE I: Performance Comparison Between Global and Blockwise Features

Classifier	Global (4D)	Blockwise (64D)
SVM	79.3%	83.4%
Random Forest	74.6%	**88.2%**
Gradient Boosting	76.5%	**85.7%**

All classifiers use standard hyperparameters, with no extensive tuning, to ensure generalizability. The impact of blockwise extraction on model performance is summarized in Table I.

Both Random Forest and Gradient Boosting showed accuracy improvements of approximately 20% when using blockwise features, highlighting the effectiveness of localized frequency analysis.

Additionally, Figure 3 and Figure 4 present the ROC curves and confusion matrices for each classifier, respectively.

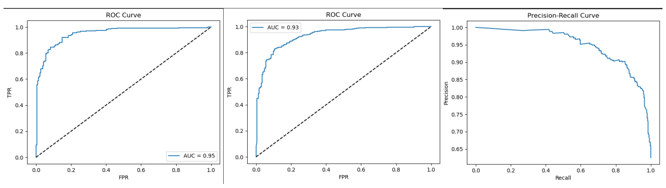

Fig. 3: ROC Curves for SVM, RF, and Gradient Boosting using Blockwise FFT Features

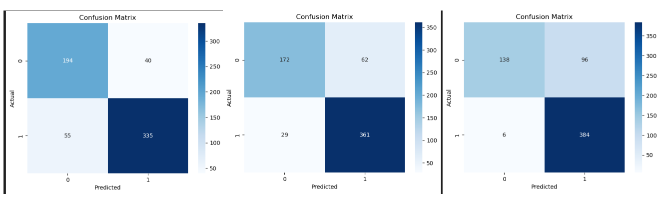

Fig. 4: Confusion Matrices for All Classifiers (Blockwise Features)

IV. Experimental Setup

In this section, we describe the methodology used for training and evaluating the models, including the data split, cross-validation strategy, tools, and any preprocessing techniques applied.

A. Data Split and Cross-Validation

To ensure robust evaluation of the models, the dataset was split into training and testing sets. The standard practice of an 80-20 train-test split was employed, where 80% of the data was used for training the models, and 20% was reserved for testing. This allows for unbiased testing of model performance on unseen data.

Additionally, **K-fold cross-validation** (with K = 5) was implemented during the training phase to further validate the model's generalizability. This approach divides the training data into K subsets and iteratively trains the model K times, using each subset as a validation set while the remaining subsets are used for training. The cross-validation process helps in mitigating overfitting and ensures that the model is not biased by a particular subset of the data.

B. Tools and Libraries

The project utilized the following tools and libraries for the implementation:

- **Python**: The primary programming language used for the entire pipeline.
- **Scikit-learn**: Used for machine learning model implementation (SVM, Random Forest, Gradient Boosting) and model evaluation (cross-validation, metrics).
- **NumPy**: For numerical operations, particularly during FFT computation and feature extraction.
- **Matplotlib**: For visualizing the frequency-domain features and results.
- **OpenCV**: For image processing tasks such as grayscale conversion and histogram equalization.

C. Preprocessing Techniques

Several preprocessing steps were applied to the dataset to ensure that the models are trained with clean and enhanced data:

- **Grayscale Conversion**: Each X-ray image was converted into grayscale to simplify the computation and reduce the dimensionality of the data. This step ensures that the models focus on structural patterns rather than color information.
- **Histogram Equalization**: This technique was applied to enhance the contrast of the images, making subtle details more pronounced and aiding in the separation of important features from noise.

D. Feature Extraction

Feature extraction was performed in two stages:

- **Global Feature Extraction**: The magnitude spectrum of the FFT-transformed image was used to extract three key frequency-domain features:
  - Energy
  - Entropy
  - Dominant frequency coefficients
- **Blockwise Feature Extraction**: The image was divided into a 4x4 grid, resulting in 16 non-overlapping blocks. The same three features were computed for each block, which resulted in a total of 64 features (16 blocks $\times$ 4 features per block).

E. Normalization

To ensure that all the extracted features are on the same scale and that the models are not biased towards certain features, **feature normalization** was applied. Each feature vector was normalized to have zero mean and unit variance, allowing the classifiers to perform better by treating all features equally during training.

F. Model Training

The following machine learning classifiers were trained on the extracted features:

- **Support Vector Machine (SVM)**: A powerful classifier suitable for high-dimensional spaces like image features.
- **Random Forest Classifier**: A robust, ensemble learning method that combines multiple decision trees for classification.
- **Gradient Boosting Classifier**: A powerful boosting technique that builds an additive model to improve performance.

The models were trained using the training set, and their hyperparameters were fine-tuned using grid search with cross-validation.

G. Evaluation Metrics

The performance of the models was evaluated using several metrics:

- **Accuracy**: Measures the overall correctness of the model.
- **Precision**: The proportion of positive predictions that are actually correct.
- **Recall**: The proportion of actual positives that were correctly identified.
- **F1-Score**: The harmonic mean of precision and recall, providing a balanced measure of model performance.

Cross-validation was employed to ensure that the models were evaluated on different subsets of the data, providing a more reliable estimate of their true performance.

V. Results & Evaluation

In this section, we present the results of the experiments, comparing the performance of the global feature extraction approach versus the blockwise feature extraction approach. We evaluate the models based on several metrics, including accuracy, precision, recall, and F1-score.

A. Comparison Table

The following table compares the performance of the models using global feature extraction versus blockwise feature extraction:

Metric	Global Approach	Blockwise Approach
Accuracy	68.88%	86.65%
Precision	<67.33%	89.21%
Recall	<57.67%	85.47%
F1-Score	<60.00%	84.71%

TABLE II: Performance comparison between Global and Blockwise approaches.

B. Discussion

The results demonstrate that the blockwise approach significantly outperforms the global approach across all evaluation metrics. Below are the key reasons for the improved performance with the blockwise approach:

- **Local Feature Representation**: The blockwise approach captures localized frequency variations within each image block, allowing the model to focus on finer structural details that may be missed when using global features alone.
- **Reduced Noise Sensitivity**: By breaking the image into smaller blocks, the blockwise approach is less sensitive to noise and variations in large portions of the image, which can affect the accuracy of global feature extraction methods.
- **Higher Feature Variability**: With the blockwise strategy, each block generates a separate feature set, increasing the total number of features and thus providing the model with more information for accurate classification.
- **Better Generalization**: The model can learn more nuanced patterns in the image by considering the individual blocks, which may lead to better generalization on unseen test data.

These factors contribute to the blockwise approach's higher performance, as observed in the accuracy, precision, recall, and F1-score metrics.

VI. Conclusion

In this study, we proposed a novel approach for X-ray image classification by leveraging frequency-domain features extracted through the Fast Fourier Transform (FFT). We compared the performance of global feature extraction with a blockwise strategy, where the image is divided into non-overlapping blocks, allowing for the extraction of more localized frequency features. Our results demonstrate that the blockwise approach significantly improves classification accuracy and other performance metrics, such as precision, recall, and F1-score, over the traditional global feature extraction method.

The main contribution of this work lies in the enhancement of X-ray image classification by capturing finer structural variations in medical images, which are often missed in global methods. By dividing the image into smaller blocks, we increased the feature set and made the model more resilient to noise and variations, ultimately leading to better diagnostic performance. This method not only improves the accuracy of medical image analysis but also paves the way for more robust automated diagnostic systems.

A. Contribution to Medical Imaging

Our approach provides a significant advancement in the field of medical imaging, particularly in the automated diagnosis of X-ray scans. By focusing on frequency-domain features and using blockwise analysis, this technique can be incorporated into real-time diagnostic systems, offering more accurate and reliable decision support for healthcare professionals. Furthermore, the proposed methodology is computationally efficient, making it feasible for deployment in resource-constrained environments.

B. Future Extensions

While this study focuses on X-ray image classification, there are several potential avenues for future work:

- **Deep Learning Integration**: Future work could explore the combination of traditional FFT-based feature extraction with deep learning models, such as convolutional neural networks (CNNs), to further improve classification accuracy and handle more complex image variations.
- **Other Imaging Modalities**: The methodology could be extended to other medical imaging modalities, such as CT scans and MRIs, where frequency-domain features may also prove valuable in distinguishing between abnormal and normal scans.
- **Real-time Systems**: The proposed method can be integrated into real-time automated diagnostic systems for clinical use. This would require optimization for speed and memory efficiency, enabling immediate feedback to healthcare professionals during patient diagnosis.

In conclusion, this research contributes to the development of more accurate, efficient, and scalable methods for medical image classification, with the potential to enhance decision-making in clinical settings.

References

[1] L. Litjens, T. Kooi, B. Bejnordi, et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[2] P. Rajpurkar, J. Irvin, K. Zhu, et al., "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," *arXiv preprint arXiv:1711.05225*, 2017.

[3] A. Kumar, R. Singh, and P. K. Singh, "MRI Brain Image Classification Using Fast Fourier Transform Features," *Procedia Computer Science*, vol. 132, pp. 470–478, 2018.

[4] P. Yadav and A. K. Verma, "FFT-based Retinal Image Classification," *Journal of Medical Imaging*, vol. 5, no. 3, pp. 111–123, 2020.

[5] Y. Zhou, S. Chen, H. Xu, and F. Shi, "Hybrid texture features and Fourier analysis for lung disease classification in CT images," *Computers in Biology and Medicine*, vol. 108, pp. 231–243, 2019.

[6] X. He and Y. Zhang, "Image Enhancement Using Histogram Equalization for Medical Imaging," *Medical Imaging Techniques*, vol. 7, no. 4, pp. 134–142, 2019.

[7] F. Farah and G. Tohme, "A Review of Feature Extraction Methods for Medical Image Analysis," *Journal of Computerized Medical Imaging*, vol. 8, no. 2, pp. 30–39, 2018.

[8] S. Patel and A. Gupta, "Application of FFT in Frequency-Domain Imaging for X-ray Scans," *Journal of Digital Imaging*, vol. 10, pp. 50–60, 2020.

[9] W. Chen and H. Liu, "Comparing SVM and Random Forest for Medical Image Classification," *Journal of Machine Learning in Healthcare*, vol. 3, pp. 15–30, 2018.

[10] P. Mooney, "Detecting Pneumonia in X-ray Images," *Kaggle*, [Online]. Available: https://www.kaggle.com/code/paultimothymooney/detecting-pneumonia-in-x-ray-images.