BAREKENG: Journal of Mathematics and Its Applications
December 2024 Volume 18 Issue 4 Page 2387–2396
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol18iss4pp2387-2396

EFFICIENCY AND ACCURACY OF CONVOLUTIONAL AND FOURIER TRANSFORM LAYERS IN NEURAL NETWORKS FOR MEDICAL IMAGE CLASSIFICATION

Fauzi Nafi'udin [1*], Hasih Pratiwi[2*], Etik Zukhronah[3]

[1,2,3] *Statistics Study Program, Faculty of Mathematics and Natural Science, Universitas Sebelas Maret*
Jalan Ir. Sutami 36A, Surakarta, 57126, Indonesia

*Corresponding author's e-mail: * **hpratiwi@mipa.uns.ac.id***

Article History:

Received: 22nd, March 2024
Revised: 28th, May 2024
Accepted: 24rd, July 2024
Published:14th, October 2024

Keywords:

Neural Network;
Convolution Layer;
Fourier Transform Layer;
Image Classification.

ABSTRACT

In an era where information flow is moving at a rapid pace, image data processing is becoming increasingly important as technology advances, including in healthcare. Convolutional Neural Network (CNN) has been a common approach in image classification, but the larger the volume of data and the complexity of the task, the more expensive the computational cost of CNN. With the rapid growth in the amount of image data, efficiency in data processing is becoming increasingly important. In this study, the performance of neural network models using the convolution layer and Fourier transform layer in medical image data classification was compared. The results show that models with a Fourier transform layer tend to provide higher accuracy and better Area Under Curve (AUC) compared to models using a convolution layer. In addition, the model with the Fourier transform layer also shows faster execution time per epoch, which indicates efficiency in data processing. However, the convolution layer has an advantage in terms of model size, although it is not significantly different from the Fourier transform layer. In conclusion, the Fourier transform layer has an advantage in the classification of medical image data.

How to cite this article:

F. Nafi'udin, H. Pratiwi and E. Zukhronah., "EFFICIENCY AND ACCURACY OF CONVOLUTIONAL AND FOURIER TRANSFORM LAYERS IN NEURAL NETWORKS FOR MEDICAL IMAGE CLASSIFICATION," *BAREKENG: J. Math. & App.*, vol. 18, iss. 4, pp. 2387-2396, December, 2024.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

The world is experiencing rapid development in various sectors, including the digital world, which is full of fast-moving information flows. Information that can be interpreted as structured data is spread through various mediums, such as spoken language, newspapers, and videos [1]. Over the past few decades, technological limitations have only allowed tabular data processing. With the advancement of existing technologies, managing and analyzing complex data, including image or photo data, has been made possible. Image processing has evolved, and the first digital photograph was created in 1957 by Russell Kirsch [2].

The growing importance of image data processing can be seen because almost everyone worldwide can now generate data in photographs. There are various methods to process images, one of which is through classification, an attempt to group images based on certain criteria. The use of neural networks has been a rapidly growing method in handling this task, as seen in the research by Ehtisham et al. [3], who used CNN to identify defects in wood structures. Although CNNs are effective in extracting features from images, the increase in global data volume makes processing using CNNs computationally expensive [4].

One solution to this challenge is using the Fourier transform, which continues to evolve today. The Fourier transform is a mathematical method that changes the representation of a function or object from the spatial or time domain to the frequency domain. This process is widely used in various fields, including digital image processing [5]. In this context, the method in question is the FFT (Fast Fourier Transform), a development of the DFT (Discrete Fourier Transform), often used in signal and image processing. In image processing, the Fourier transform helps extract important information from an image by transforming it into the frequency domain, which can then be restored to its original form.

For example, the Fourier Transform, specifically the Fast Fourier Transform (FFT), was initially applied by Minami et al. [6] to distinguish potentially life-threatening heartbeat rhythms through electrocardiogram (ECG) signals. This method proved more effective than previous approaches due to its lighter computation. In addition, other studies use different approaches. Tile-based Fast Fourier Transform (tFFT), a neural network model, improves the efficiency of FFT-based convolution in the CNN layer, which shows a significant improvement in convolution operations [7]. In addition, Vasilache et al. [8] assessed the effectiveness of CNN training using the latest NVIDIA GPU (Graphics Processing Unit). They introduced a new approach in implementing Convolution Fast Fourier Transform (cuFFT), which provides 1.5 to 5 times faster speed.

In the context of image classification, a recent approach by Zak et al. [9] introduced a Fourier transform layer as an alternative to the convolution layer in CNNs. This component aims to speed up the image classification process without sacrificing accuracy, offering a promising solution to reduce the reliance on powerful GPUs during training. With encouraging results, the model using the Fourier transform layer achieved almost the same accuracy as the conventional model while reducing the training time by at least 27% when using a standard Central Processing Unit (CPU).

Based on the things mentioned above, this study compares the performance efficiency of neural network models with the convolution and Fourier transform layers, as discussed in the journal by Zak et al. [9] in processing several image data. The data used in this study is image data in the health sector. The reason for selecting this data is the importance of efficiency in image data processing, where processing speed also plays an important role in addition to classification accuracy. The selection of this data is considered relevant in society, and the comparison results are expected to help determine which method is more effective in overcoming the complexity of image data processing in classification tasks.

2. RESEARCH METHODS

The research method used in this study is divided into 5 sections. Section 2.1 explains the theoretical basis of the analysis performed. Section 2.2 describes in detail the dataset used. Furthermore, the model configuration will be explained in detail in section 2.3. Section 2.4 describes the various parameters used in the training process. The last section, section 2.5, contains the approach used in comparing the convolution layer and Fourier transform layer.

2.1 Theoretical Analysis

Convolutional Neural Network (CNN) is a type of deep learning model used to process data that has grid patterns, such as images [10]. In research written by Hubel and Wiesel [11] , the CNN model was inspired by the organization of the animal visual cortex. In addition, this model is designed to automatically and adaptively learn spatial features, ranging from low-level to high-level patterns. The primary objective of CNN is to develop a method that decreases the overall parameter count while constructing a neural network that is more intricate with fewer parameters [12]. The convolution layer is the primary component of the CNN architecture in charge of feature extraction, which usually consists of a combination of linear and non-linear operations, namely convolution operations and activation functions [13]. Figure 5 shows one example of a CNN taken from Alzubaidi's research, which illustrates the general structure of CNN architecture, consisting of three fundamental layers: convolution layer, pooling layer, and fully connected layer [4].

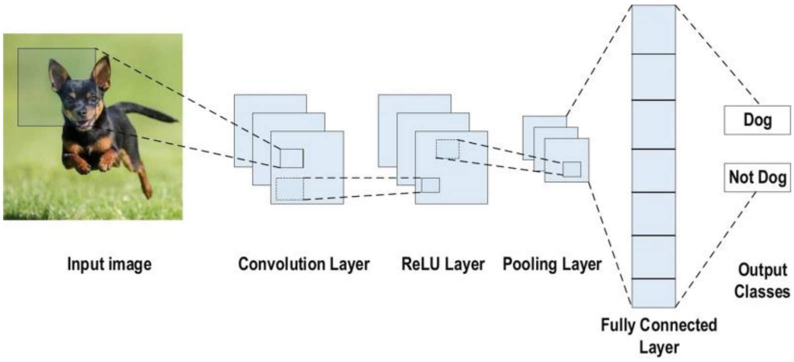

Figure 1. Example of CNN Architecture

Nwankpa et al. stated that the activation function is utilized in neural networks to compute the weighted sum of input and bias, which decides whether a neuron is activated [14]. This function analyzes the given data and generates an output for the neural network, including the parameters found in the data. Nair and Hinton introduced the Rectified Linear Unit (ReLU) as an activation function, gaining widespread popularity in deep learning [15]. LeCun et al. reported that ReLU speeds up model learning in their study [16]. Furthermore, according to a journal authored by Ramachandran et al., ReLU has emerged as a highly effective and widely used activation function [17]. ReLU is a nearly linear function, so it retains the characteristics of linear models, which allows it to be easily optimized using the gradient descent technique [18]. Equation (1) defines the ReLU activation function.

$$f(x) = max(0;\ x) = \begin{cases} x, & if\ x \geq 0 \\ 0, & if\ x\ < 0 \end{cases} \tag{1}$$

Through computation, the ReLU activation function results in a value of 0 when given a negative input and generates a positive value matching the input when given a positive input. Es-Sabery et al. stated that the ReLU function has various benefits, including its capacity to address the issue of instability in gradient value changes and quicker convergence because of its straightforward formula [12].

Fast Fourier Transform (FFT) is an algorithm invented by James W. Cooley and John W. Tukey [19]. This algorithm performs Discrete Fourier Transform (DFT) operations more quickly and efficiently. DFT itself is a method to analyze signals in the frequency domain. Time complexity refers to how efficient or how much time an algorithm takes to complete its task, depending on the size of its input. Compared to the DFT, which has a time complexity of $O(n2)$, which means that the time required by the algorithm increases quadratically as the input size increases, the FFT update has a time complexity of $O(n\ log(n))$. This means that the FFT algorithm is more efficient as its execution time increases logarithmically as the input size increases, making it more suitable for large data [20]. In mathematics, the Fourier transform and inverse Fourier transform can be defined as Equation (2) and Equation (3), respectively [21].

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} \tag{2}$$

and

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} \tag{3}$$

where $F(\omega)$ represents the function in the frequency domain while $f(t)$ represents the function in the time domain. The variable ω denotes frequency in radians per second and t denotes the time variable.

The Fourier transform layer is an adaptation of the Fourier transform technique into a neural network layer that utilizes the Fast Fourier Transform (FFT) on the input data [9]. After the FFT process, the real part of the transform is multiplied by the weight of the layer, which is a multiplication kernel for the entire real part of the Fourier transform. Thus, only two operations, FFT and multiplication, are required in the training process. FFT is known as an optimally efficient calculation tool.

2.2 Data Description

This study used three publicly available health datasets on the kaggle.com website: Breast Cancer, Lymphoma, and SiPaKMeD. The Breast Cancer data consists of images of breast cancer cells with dimensions (700 × 460 × 3) taken from 82 patients [22]. The image data type is Portable Network Graphics (PNG). This data includes two classes representing different breast tumor types, such as benign and malignant. The total number of images in this data is 7909. The second dataset, Lymphoma Data, was created by Orlov et al. [23] and contains images of malignant lymph node cancer cells. It includes three classes: chronic lymphocytic leukemia, follicular lymphoma, and mantle cell lymphoma. This data consists of 374 images, and the image data type is Tagged Image File Format (TIF). Meanwhile, the last data is SiPaKMeD data, created by Plissiti et al. [24], focusing on images of cervical cancer disease cells. Consisting of five classes, this data includes cell types such as superficial-intermediate, parabasal, koilocytotic, dyskeratotic, and metaplastic, with 4,049 images with bitmap data type (BMP). All data was placed in local storage, and during the processing process, there was no change in the type of image that was directly resized and converted to numbers.

The data was originally divided into training and testing only, or training, validation, and testing will be combined in the data preparation stage. Furthermore, the dataset will be partitioned again with a predetermined proportion of 80% for training data, 10% for validation data, and 10% for testing data. This approach is applied to overcome variations in the initial partitioning of data that may only consist of two classes: training data and testing data. The main objective is to ensure that the entire dataset follows the same configuration, creating a balanced and objective basis for the model training and testing process.

2.3 Model Configuration

The method used in this research's model configuration adopts the neural network architecture configuration described in the journal Zak et al. [9]. The development of the neural network model focuses on two main layers, namely the convolution layer and Fourier transform layer, which is then further divided into Convolution Small Kernel (CSK), Convolution Large Kernel (CLK), Fourier Model (FM), and Inverse Fourier Model (iFM). CSK uses a convolution layer with the same parameters as the Fourier transform layer. CLK is a variation that uses an overall number of parameters equivalent to the Fourier Model (FM). FM is a benchmark with the Fourier transform layer, while iFM introduces inverse properties to the Fourier transform layer. The overall architecture involves input layers, hidden layers, and fully connected layers, with hidden layers being either convolution layers or Fourier transform layers and fully connected layers involving flattened and dense layers.

2.4 Model Training

The architecture model is designed and organized based on several predefined rules to perform the training process. Training is carried out up to 100 epochs by applying callbacks based on the validation accuracy value and a patience level of 4. The batch size used is 8, and the loss function used is categorical cross-entropy, which is optimized using the Adam optimizer. In the Fourier transform layer and the convolution layer, the activation function is ReLU, while in the last output layer, the activation function used is softmax. Evaluation of the training results includes monitoring the loss value, accuracy, validation loss, and validation accuracy, as well as measuring the time taken by the model during the training process. It should be noted that the neural network model training process used in this research uses GPU.

After the training stage, the model parameters are stored, and then the stored parameters are called to complete the prediction process on the test data. In this study, repetition was performed 20 times for each model variation. The computing resources used in this study involve a 12th-generation Intel(R) Core(TM) i9-12900K processor with 24 CPUs, operating at a speed of approximately 3.2 GHz. These resources are supported by 64 GB of RAM and an RTX 3060 Ti GPU. This research also uses Python version 3.12 and Tensorflow version 2.14.0 and operates on the Windows 11 platform.

2.5 Model Evaluation

Several metrics are used to compare and determine the better model in the model evaluation process. Grandini et al. explain that the confusion matrix is a table used to describe the performance of a classification model by comparing the actual (correct) classification and the classification predicted by the model [25]. Table 1 is an example of a multiclass confusion matrix.

Table 1. Confusion Matrix Multiclass

		Prediction		
		A	B	C
Actual	A	TP_A	E_{AB}	E_{AC}
	B	E_{BA}	TP_B	E_{BC}
	C	E_{CA}	E_{CB}	TP_C

According to Tharwat [26], in Table 1, the grey diagonal represents correct predictions, while the red diagonal represents incorrect predictions. True Positive (TP) is where the actual and predicted values must be identical. False Negative (FN) is the sum of values from the same row except the TP value. False Positive (FP) is the sum of values from the same column except the TP value. True Negative (TN) is the sum of values from all columns and rows except the class's calculated value. In addition, error (E) is a symbol that describes the number of prediction errors. Accuracy is one of the key metrics that measure the extent to which a model can provide correct predictions. The higher the accuracy value, the better the performance of the model. According to Tharwat [26], accuracy measures the ratio between the number of samples successfully classified correctly and the overall number of samples, providing an overall picture of the model's success in predicting classifications across data. The formula for calculating accuracy in multi-class classification is Equation (4):

$$Accuracy = \frac{TP_A + TP_B + TP_C}{Total\ Data} \times 100\% \quad (4)$$

where TP_A, TP_B, and TP_C are observations where the prediction is equal to the actual class.

In addition to accuracy, Area Under Curve (AUC) is also a concern, as it reflects how well the model can distinguish between positive and negative classes. The higher the AUC value, the better the model's ability in terms of sensitivity and specificity. As stated by Zak et al., the Area Under the Curve (AUC) is a measure that utilizes probabilities derived from the model to assess its effectiveness in classifying samples not present in the dataset. The closer the AUC gets to 1, the higher the model's performance [9]. Furthermore, the training time of the model is also an important metric, where evaluation is done to determine how fast the model can complete the training process. A shorter training time can indicate the model's efficiency in learning the data. Model size is also considered one of the important factors in model efficiency and flexibility in practical problems. The use of these metrics provides a complete picture of the quality and performance of the developed model.

After obtaining the data, statistical testing will determine the best model for each dataset. This process involves a series of statistical tests. The initial stage involves the Shapiro-Wilk test to evaluate the distribution of the data [27]. Next, ANOVA or Kruskal-Wallis tests will be performed, depending on the identified data distribution [28]. Afterward, post hoc tests, such as the Tukey test [29] or Mann-Whitney U test [30] will be applied for further analysis. It is important to note that a significance level of 5% (alpha = 0.05) was used during this examination.

3. RESULTS AND DISCUSSION

The results and discussion of this research are compiled using the methods previously described. The following is the analysis process carried out by the research methods used. The three datasets were initially divided into 80% for training, 10% for validation, and 10% for testing, as shown in Table 2. Table 2 illustrates the balanced data variation, with a limited amount of data in the Lymphoma data, a moderate amount in the SiPaKMeD data, and a large amount in the Breast Cancer data. This variation shows a fairly diverse representation of the three datasets. Before neural network processing, it is imperative to note that all image data was uniformly resized to (150 × 150 × 3) and rescaled.

Table 2. Data Distribution

Dataset Name	Number of Data	Train	Validation	Testing
Breast Cancer	7783	6226	780	777
Lymphoma	374	298	40	36
SiPaKMeD	4049	3237	409	403

Table 2 illustrates the balanced data variation, with a limited amount of data in the Lymphoma data, a moderate amount in the SiPaKMeD data, and a large amount in the Breast Cancer data. This variation shows a fairly diverse representation of the three datasets. Before neural network processing, it is imperative to note that all image data was uniformly resized to (150 × 150 × 3) and rescaled.

The method used in the model configuration stage is the research written by Zak et al. [9] to design the model architecture. Four models were formed by varying two different layers. The aspects set here include the size of the image to be processed, the dimensions of the kernel used, and the number of filters or neurons applied to each layer. The notation used to represent these layers is as follows: C for the convolution layer, F for the flattened layer, D for the dense layer, and FT for the Fourier transform layer. More detailed information on the model characteristics can be found in Table 3.

Table 3. Model Configuration

Layer		Filter	Kernel Size	Parameter	Layer		Filter	Kernel Size	Parameter
CSK					CLK				
1	C	900	5 × 5	68400	1	C	10	32 × 32	30730
2	F	-	-	0	2	F	-	-	0
3	D	3	-	57553203	3	D	3	-	424833
Total Parameter				57621603	Total Parameter				455563
FM					iFM				
1	FT	-	-	67500	1	FT	-	-	67500
2	F	-	-	0	2	F	-	-	0
3	D	3	-	405003	3	D	3	-	405003
Total Parameter				472503	Total Parameter				472503

Table 3 is written using only the Lymphoma dataset with three classes. The results show that in the first layer of the Convolution Small Kernel (CSK), the number of parameters formed is proportional to the Fourier Model. However, in the convolution large kernel (CLK), the accumulated parameters seem more or less comparable to the Fourier Model. Although there may be slight variations when applying to other datasets due to the different number of classes, the concept of comparison remains consistent across the evaluated datasets.

The evaluation process involves assessing the training results across four datasets through various statistical tests. The Shapiro-Wilk test is first used to determine data distribution, followed by the ANOVA test for normal data or the Kruskal-Wallis test for non-normal data. Post hoc tests (Tukey or Mann-Whitney U) are then applied based on the data distribution. The analysis focuses on accuracy, AUC, and training time, with bolded values indicating optimal results and underlined values showing no significant difference from the optimal result. Table 4 summarizes the average accuracy results.

Table 4. Accuracy Results for Each Model

Dataset Name	CSK	CLK	FM	iFM
Breast Cancer	68,21 ± 0	68,21 ± 0	77,19 ± 1,48	**83,94 ± 1,4**

Dataset Name	CSK	CLK	FM	iFM
Lymphoma	36,25 ± 0,62	36,11 ± 0	41,67 ± 6,11	**52,08 ± 9,4**
SiPaKMeD	26,92 ± 9,5	20,58 ± 0,06	71,94 ± 2,69	**79,69 ± 2,09**

Table 4 shows that the model with the Fourier transform layer dominates in terms of accuracy. This can be seen from the high consistency of results for both Fourier models on each of the evaluated datasets. The iFM model using the Fourier transform layer achieved the highest accuracy, while the FM model ranked second for all tested datasets. Furthermore, the average AUC results are presented in **Table 5**.

Table 5. Area Under the Curve (AUC) Results for Each Model

Dataset Name	CSK	CLK	FM	iFM
Breast Cancer	50,45 ± 4,68	50 ± 0	74,02 ± 3,97	**86,55 ± 2,19**
Lymphoma	54,28 ± 6,64	49,73 ± 0,92	59,94 ± 5,08	**69,14 ± 8,16**
SiPaKMeD	56,5 ± 9,4	50 ± 0	88,93 ± 1,89	**95,54 ± 0,69**

Table 5 shows that iFM is superior in all these comparisons. Meanwhile, the results for CSK and CLK are not satisfactory over a wide range of data. Although the FM model is inferior to iFM, it is good enough to outperform the model using the convolution layer. Furthermore, the average training time in seconds for each model is presented in **Table 6**.

Table 6. Results of Average Time Per Epoch in Seconds for Each Model

Dataset Name	CSK	CLK	FM	iFM
Breast Cancer	<u>71,51 ± 13</u>	**68,12 ± 13,8**	<u>76,37 ± 13,7</u>	<u>75,4 ± 13,76</u>
Lymphoma	4,49 ± 1	4,12 ± 1,23	3,95 ± 1,01	**3,83 ± 1,17**
SiPaKMeD	22,04 ± 0,79	8,91 ± 2,78	9,36 ± 2,53	**8,12 ± 2,45**

Table 6 shows that iFM appears to have a shorter training time than the other models in most trials. However, it should be noted that there is an exception in the Breast Cancer dataset, where CLK outperforms the other models in terms of training time efficiency. Meanwhile, FM showed variations in training time depending on the dataset. On the other hand, CSK tends to have a longer training time, especially on the SiPaKMeD dataset, and requires special attention in terms of training time efficiency. Furthermore, the total training time (average time of one epoch × number of epochs) in seconds for each model is presented in **Table 7**.

Table 7. Results of Total Training Time in Seconds for Each Model

Dataset Name	CSK	CLK	FM	iFM
Breast Cancer	357,5 ± 64,8	**340,6 ± 69,1**	645,8 ± 338	725,8 ± 233
Lymphoma	23,72 ± 6,47	**20,92 ± 6,72**	29,52 ± 7,6	45,17 ± 28,6
SiPaKMeD	155,5 ± 40,9	**48,68 ± 19,3**	75,13 ± 31,8	102,1 ± 49,4

Table 7 shows that CLK appears to have a shorter training time than the other models in all trials. On the other hand, iFM tends to have a longer training time, especially on the Breast Cancer dataset, and requires special attention regarding training time efficiency. Furthermore, the model size in kilobytes for each model is presented in **Table 8**.

Table 8. Model Size for Each Dataset

Dataset Name	CSK	CLK	FM	iFM
Breast Cancer	450463 KB	**3708 KB**	3980 KB	3980 KB
Lymphoma	675280 KB	**5367 KB**	5562 KB	5562 KB
SiPaKMeD	1124915 KB	**8685 KB**	8726 KB	8726 KB

Table 8 shows variation among the four models, but it is notable in the CSK model that the size of the model is too far compared to the other models. When viewed from the smallest size, the CLK model outperforms the other models for the entire dataset. However, it should be noted that the size of FM and iFM do not vary significantly from CLK for the entire dataset. In terms of layer comparison, it can be said that the model with the Fourier transform layer has a more stable and similar size for all tested datasets compared to

the convolution layer. This will certainly be useful and important, especially in terms of the efficiency and flexibility of the model in practical problems.

From the above results, it can be observed that there are conditions where the convolution layer stagnates or does not show any performance improvement for accuracy and AUC from various experiments. This factor is due to the difficulty of the model in learning the data. Although the training time of each epoch of the convolution layer may vary, the total training time is relatively short as some epochs do not learn the data well, thus ending the learning process sooner.

On the other hand, the Fourier transform layer consistently performed better in terms of accuracy and AUC than the convolution layer. Furthermore, the training process of the Fourier model takes longer than that of the convolution layer. The training time per epoch for the Fourier model tends to be shorter, especially for the inverse Fourier model. However, the total training time tends to be longer because more epochs are required to achieve the optimal level of performance in the model training process.

4. CONCLUSIONS

From the results and discussion described above, some important conclusions can be drawn that illustrate the efficiency and effectiveness of neural network models with convolutional and Fourier transform layers in image data classification:

1. Models with the Fourier transform layer consistently demonstrated superior performance in terms of accuracy and AUC compared to those with convolutional layers, with the iFM model achieving the highest scores across various datasets.
2. The iFM model generally had a faster execution time per epoch, except on the Breast Cancer dataset where the CLK model was faster. Despite this, the overall training time was the shortest for the CLK model and the longest for the iFM model, especially for the Breast Cancer dataset.
3. Model size analysis indicated that the CLK model was the smallest, while the FM and iFM models were similar in size, with the Fourier transform layer showing better stability.
4. The Fourier transform layer is more effective and efficient in terms of accuracy, execution time, and model size stability, underscoring its practical application potential in image data classification.
5. Limitations of this study include the use of datasets that were not processed on the CPU, simple data preprocessing steps, lack of validation methods like cross-validation, and insufficient experimentation with hyperparameter tuning. These factors could impact the results and suggest areas for future research.

Combining Fourier transform and convolutional layers offers a promising direction for developing more complex and efficient models for image classification tasks.

REFERENCES

[1] S. Ati, Nurdien, H. Kistanto, and A. Taufik, *Pengantar Konsep Informasi, Data, dan Pengetahuan*. Universitas Terbuka, 2014.

[2] J. Stockham, "Hard Copy," Jul. 2015. doi: 10.14236/ewic/eva2015.25.

[3] R. Ehtisham *et al.*, "Classification of defects in wooden structures using pre-trained models of convolutional neural network," *Case Studies in Construction Materials*, vol. 19, Dec. 2023, doi: 10.1016/j.cscm.2023.e02530.

[4] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00444-8.

[5] N. Chittora and D. Babel, "A BRIEF STUDY ON FOURIER TRANSFORM AND ITS APPLICATIONS," *International Research Journal of Engineering and Technology*, p. 1127, 2008, [Online]. Available: www.irjet.net

[6] K. I. Minami, H. Nakajima, and T. Toyoshima, "Real-time discrimination of ventricular tachyarrhythmia with fourier-transform neural network," *IEEE Trans Biomed Eng*, vol. 46, no. 2, pp. 179–185, 1999, doi: 10.1109/10.740880.

[7] J. Lin and Y. Yao, "A Fast Algorithm for Convolutional Neural Networks Using Tile-based Fast Fourier Transforms," *Neural Process Lett*, vol. 50, no. 2, pp. 1951–1967, Oct. 2019, doi: 10.1007/s11063-019-09981-z.

[8] N. Vasilache, J. Johnson, M. Mathieu, S. Chintala, S. Piantino, and Y. LeCun, "Fast Convolutional Nets With fbfft: A GPU Performance Evaluation," Dec. 2014, [Online]. Available: http://arxiv.org/abs/1412.7580

[9] J. Zak, A. Korzynska, A. Pater, and L. Roszkowiak, “Fourier Transform Layer: A proof of work in different training scenarios,” *Appl Soft Comput*, vol. 145, p. 110607, Sep. 2023, doi: 10.1016/J.ASOC.2023.110607.
[10] K. Fukushima, “Biological Cybernetics Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position,” 1980.
[11] D. H. Hubel and T. N. Wiesel, “RECEPTIVE FIELDS AND FUNCTIONAL ARCHITECTURE OF MONKEY STRIATE CORTEX,” 1968.
[12] F. Es-Sabery, A. Hair, J. Qadir, B. Sainz-De-Abajo, B. Garcia-Zapirain, and I. Torre-Diez, “Sentence-Level Classification Using Parallel Fuzzy Deep Learning Classifier,” *IEEE Access*, vol. 9, pp. 17943–17985, 2021, doi: 10.1109/ACCESS.2021.3053917.
[13] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology,” Aug. 01, 2018, *Springer Verlag*. doi: 10.1007/s13244-018-0639-9.
[14] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, “Activation Functions: Comparison of trends in Practice and Research for Deep Learning,” Nov. 2018, [Online]. Available: http://arxiv.org/abs/1811.03378
[15] V. Nair and G. E. Hinton, “Rectified Linear Units Improve Restricted Boltzmann Machines,” 2010.
[16] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
[17] P. Ramachandran, B. Zoph, and Q. V. Le, “Searching for Activation Functions,” Oct. 2017, [Online]. Available: http://arxiv.org/abs/1710.05941
[18] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
[19] J. W. Cooley and J. W. Tukey, “An Algorithm for the Machine Calculation of Complex Fourier Series,” 1964.
[20] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering*. Cambridge University Press, 2019. doi: 10.1017/9781108380690.
[21] M. S. Reis, P. M. Saraiva, and B. R. Bakshi, “Denoising and Signal-to-Noise Ratio Enhancement: Wavelet Transform and Fourier Transform,” *Comprehensive Chemometrics*, vol. 2, pp. 25–55, Jan. 2009, doi: 10.1016/B978-044452701-1.00099-5.
[22] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, “A Dataset for Breast Cancer Histopathological Image Classification,” *IEEE Trans Biomed Eng*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016, doi: 10.1109/TBME.2015.2496264.
[23] N. V. Orlov *et al.*, “Automatic classification of lymphoma images with transform-based global features,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 4, pp. 1003–1013, Jul. 2010, doi: 10.1109/TITB.2010.2050695.
[24] M. E. Plissiti, P. Dimitrakopoulos, G. Sfikas, C. Nikou, O. Krikoni, and A. Charchanti, “Sipakmed: A New Dataset for Feature and Image Based Classification of Normal and Pathological Cervical Cells in Pap Smear Images,” in *Proceedings - International Conference on Image Processing, ICIP*, IEEE Computer Society, Aug. 2018, pp. 3144–3148. doi: 10.1109/ICIP.2018.8451588.
[25] M. Grandini, E. Bagli, and G. Visani, “Metrics for Multi-Class Classification: an Overview,” Aug. 2020, [Online]. Available: http://arxiv.org/abs/2008.05756
[26] A. Tharwat, “Classification assessment methods,” *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168–192, Jan. 2021, doi: 10.1016/j.aci.2018.08.003.
[27] B. W. Yap and C. H. Sim, “Comparisons of various types of normality tests,” *J Stat Comput Simul*, vol. 81, no. 12, pp. 2141–2155, Dec. 2011, doi: 10.1080/00949655.2010.520163.
[28] S. F. Sawyer, “Analysis of Variance: The Fundamental Concepts,” *Journal of Manual & Manipulative Therapy*, vol. 17, no. 2, pp. 27E-38E, Apr. 2009, doi: 10.1179/jmt.2009.17.2.27e.
[29] A. Nanda, Dr. B. B. Mohapatra, A. P. K. Mahapatra, A. P. K. Mahapatra, and A. P. K. Mahapatra, “Multiple comparison test by Tukey’s honestly significant difference (HSD): Do the confident level control type I error,” *International Journal of Statistics and Applied Mathematics*, vol. 6, no. 1, pp. 59–65, Jan. 2021, doi: 10.22271/maths.2021.v6.i1a.636.
[30] N. Nachar, “The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from the Same Distribution,” 2008.