DOI: http://dx.doi.org/10.21111/fij.v9i1.12166
Diterima: 24 Mei 2024 Revisi: - Terbit: 13 Juli 2024

Enhancing Covid-19 Diagnosis: Glrlm Texture Analysis And Kelm For Lung X-Ray Classification

Dian Candra Rini Novitasari [1)] *, Alvin Nuralif Ramadanti [2)], Dina Zatusiva Haq [3)]

Mathematics Department, UIN Sunan Ampel [1)], Mathematics Department, UIN Sunan Ampel [2)], Mathematics Department, UIN Sunan Ampel [3)]
septian@gmail.com [1)] *, nuralifalvin@gmail.com [2)], zatusivad@gmail.com [3)]

Abstrak

This study aims to diagnose COVID-19 using GLRLM feature extraction, known for its high accuracy, and optimize Kernel Extreme Learning Machine (KELM) with Genetic Algorithm (GA) for improved computational efficiency, along with Principal Component Analysis (PCA) for data reduction. The gamma values in KELM are optimized using GA, yielding the best solution function. Results reveal that at angles of 0°, 45°, and 135°, the optimal gamma value with KELM is 1, while at 90°, GA determines it to be 1.35. This adjustment demonstrates the critical role of gamma values in achieving optimal performance. Performance analysis of various classification methods demonstrates that GLRLM-PCA-Optimized KELM outperforms others, achieving an accuracy exceeding 97%, particularly notable at 90° angles. This study shows that the importance of hyperparameter optimization in enhancing classification accuracy, revealing a significant improvement of over 1% compared to non-optimized models.

Kata kunci*: COVID-19, GLRLM, KELM, Feature Reduction, PCA*

Abstract

Penelitian ini bertujuan untuk mendiagnosis COVID-19 menggunakan ekstraksi fitur GLRLM yang dikenal dengan akurasi tinggi, dan mengoptimalkan Kernel Extreme Learning Machine (KELM) dengan Algoritma Genetika (GA) untuk meningkatkan efisiensi komputasi, bersama dengan Principal Component Analysis (PCA) untuk reduksi data. Nilai gamma dalam KELM dioptimalkan menggunakan GA, menghasilkan fungsi solusi terbaik. Hasil penelitian menunjukkan bahwa pada sudut 0°, 45°, dan 135°, nilai gamma optimal dengan KELM adalah 1, sedangkan pada 90°, GA menentukan nilainya menjadi 1,35. Penyesuaian ini menunjukkan peran penting nilai gamma dalam mencapai kinerja optimal. Analisis kinerja berbagai metode klasifikasi menunjukkan bahwa GLRLM-PCA-KELM yang Dioptimalkan mengungguli yang lain, mencapai akurasi lebih dari 97%, terutama mencolok pada sudut 90°. Studi ini menyoroti pentingnya optimasi hyperparameter dalam meningkatkan akurasi klasifikasi, mengungkapkan peningkatan signifikan lebih dari 1% dibandingkan dengan model KELM konvesional.

Keywords*: COVID-19, GLRLM, KELM, Feature Reduction, PCA*

1. INTRODUCTION

Covid-19, or Corona Virus Disease 2019, is a disease discovered in December 2019 in Wuhan, China, and is caused by Severe Acute Respiratory Syndrome Corona Virus 2 (SARS-CoV-2) [1]. Corona Virus classify as a zoonosis with an incubation period of 14 days with initial symptoms of cough, fever, and shortness of breath [2]. Zoonoses are viruses that can evolve to infect animals and humans. It would be hazardous if infecting humans because it can cause infections in the respiratory tract [3]. Bacteria or fungi can also cause respiratory tract infections. One of the respiratory tract infections caused by bacteria or fungi is pneumonia which occurs in the alveoli and bronchioles [4]. By 2020, WHO estimates more than 4 million deaths by pneumonia or diseases related to air pollution [5]. In addition, pneumonia is one of the highest causes of death globally, whereas many as 16% of deaths are among children under five [6].

Pneumonia occurs not only in toddlers but in all age groups, as is the case with Covid-19, where babies aged 12 months can also be infected. Covid-19 can be examined by an antigen swab test, antibody, and PCR (Polymerase Chain Reaction) test [7]. Research by [8] shows that the antigen test has a sensitivity of 30-80% and the antibody test has a sensitivity of 88.66%, with the advantages of being fast and cheap, but the accuracy value is less. Meanwhile, the PCR test has a sensitivity of 100% and is expensive, and not all hospitals have a PCR test.

Artificial Intelligence (AI) makes breakthroughs in identifying a disease by using X-ray images in the era of technological advances. X-ray is a technique of slight electromagnetic radiation that can penetrate

various objects in the human body to record digital images on film [9]. An X-ray is a tool for conducting examinations because almost all hospitals have the equipment. In addition, X-ray has the advantage that the price for the examination is much lower, and the diagnosis is also more accurate [10]. X-ray image analysis utilizes Computer-Aided Diagnosis (CAD), where CAD has four stages, namely preprocessing, feature extraction, classification, and model evaluation [11].

In the preprocessing stage, Contrast Limited Adaptive Histogram Equalization (CLAHE) improves the image's contrast. Various studies have used this method, such as that carried out by [12] which resulted in an MSE of 89.90%. In addition, research conducted by [13] compares HE, AHE, and CLAHE. HE produced MSE 148.1662, RMSE 12.1448, and PNSR 60.9338. AHE produced MSE 148.1459, RMSE 12.1439, and PSNR 60.9343. Meanwhile, CLAHE produced MSE 131.5763, RMSE 11,4360, and PSNR 62.1521. The next stage is to perform feature extraction to obtain the characteristic value of an image [14]. Various methods to perform feature extraction, such as Gray Level Co-Occurrence Matrix (GLCM), Gray Level Difference Matrix (GLDM), and Gray Level Run Length Matrix (GLRLM). The GLRLM method distinguishes between fine and coarse images [15]. Research [16] comparing GLCM, GLDM, and GLRLM obtained the best results on feature extraction in GLRLM with an accuracy of 93.97%. In addition, research conducted by [17] obtained a high accuracy value equal to 98.04%.

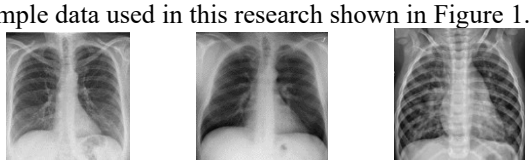
The result of the feature extraction process will be input into the classification process. Before the classification stage, a method is needed to reduce the dimensions of the data without losing its authenticity value [18]. Data reduction is essential because it can increase the accuracy of an image and reduce computation time [19]. One of the methods is Principal Component Analysis (PCA). the PCA method yielded an accuracy rate of 96.62% for cancer classification, which is higher than the classification accuracy without PCA, which was 87.89%. Classification is the stage used to diagnose whether the X-ray image is Covid-19, healty lung, or Pneumonia. It uses the Kernel Extreme Learning Machine (KELM) method at this stage.

KELM is a development of the Extreme Learning Machine (ELM) method and has the advantage of faster computation time [20]. Based on previous research, KELM has the advantage of faster computation and good performance. Parameter selection plays a pivotal role in determining the effectiveness of the KELM model. KELM attempts to discover the best model parameters through grid search, often resulting in overfitting, sluggish learning rates, and a decline in generalization efficiency [21]. To solve this problem, in this research using optimization method to discover the best model parameters. Genetic Algorithm (GA) is an optimization method inspired by the natural selection process based on biological selection. GA algorithm is a powerful tool for facing complex problems and dealing with parameters crucial in decision-making [22].

Based on the explanation from the background above, this study aims to identify Covid-19 by using the GLRLM feature extraction, which has the best accuracy value, and Optimizes Kernel Extreme Learning Machine (KELM) with GA, which has advantages in computing time with PCA data reduction.

2. RESEARCH AND METHOD

The research methodology comprises several critical stages, each meticulously designed to ensure the precision and effectiveness of our COVID-19 classification model based on lung X-ray data. The sample data used in this research shown in Figure 1.



(a) (b) (c)

Figure 1. (a) Example of Covid-19 Image Data; (b) Example of Healty lung Image Data; (c) Example of Pneumonia Image Data

The research flowchart shown in Fig 2. The first step is image enhancement using the CLAHE method, setting the stage for subsequent analysis. Feature extraction follows, extracting information from the X-ray images, that represent essential textures and patterns crucial for distinguishing COVID-19 cases from other respiratory conditions. To simplify the process and improve model efficiency, this research implements feature reduction techniques, selecting the most informative and discriminating features while reducing dimensionality.

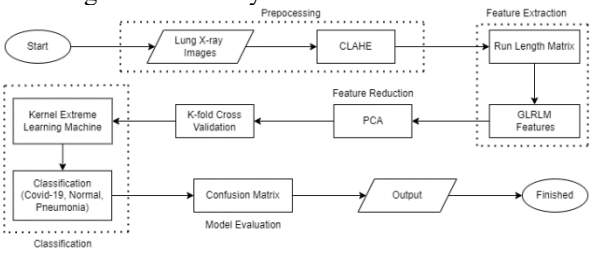


Figure 2. Graphical abstract COVID-19 Classification

In this research uses the Kernel Extreme Learning Machine (KELM) known for its efficiency in handling complex data patterns and evaluates the model's performance using various metrics. The kernel function in the KELM method is the most crucial part. This research uses the Radial Basis Function (RBF) type, with the variable γ as a hyperparameter. The γ value will be optimized using the Genetics Algorithm (GA) method to produce an identification system with maximum performance. The sequential nature of our research methodology is illustrated in Fig 2, providing a visual representation of how each stage contributes to the development of a robust and accurate COVID-19 classification model.

A. Feature Extraction using Gray Level Run Length Matrix (GLRLM)

Gray Level Run Length Matrix (GLRLM) produces a matrix with 2D features, where each element gives a total gray level in the given direction [23]. The orientation angle of the pixel search direction in this method is 0°,45°,90°, and 135°. This study uses the GLRLM method, using 11 features of GLRLM, which are in Table 1 [24].

Table 1. GLRLM Features and formula

No	Features	Formula	Usefulness
1	Short Run Emphasis (SRE)	$SRE = \frac{1}{r}\sum_{x=1}^{M}\sum_{y=1}^{N}\frac{P(x,y)}{y^2}$	Measures the distribution of short runs
2	Long Run Emphasis (LRE)	$LRE = \frac{1}{r}\sum_{x=1}^{M}\sum_{y=1}^{N}P(x,y)y^2$	Measures the distribution of long runs
3	Gray Level Non-uniformity (GLN)	$GLN = \frac{1}{r}\sum_{x=1}^{M}\left(\sum_{y=1}^{N}P(x,y)\right)^2$	Shows the similarity of pixel values in all images, and the level of gray in all images depends on the GLN value
4	Run-Length Non-uniformity (RLN)	$RLN = \frac{1}{r}\sum_{x=1}^{M}\left(\sum_{y=1}^{N}P(x,y)\,y\right)$	Shows the similarity of path length in all images
5	Run Percentage (RP)	$RP = \frac{r}{\sum_{x=1}^{M}\sum_{y=1}^{N}yP(x,y)}$	Measuring the similarity and division of space in an image
6	Low Gray Level Run Emphasis (LGRE)	$LGRE = \frac{1}{r}\sum_{x=1}^{M}\sum_{y=1}^{N}\frac{P(x,y)}{x^2}$	Shows the value of the relative distribution of low gray level values
7	High Gray Level Run Emphasis (HGRE)	$HGRE = \frac{1}{r}\sum_{x=1}^{M}\sum_{y=1}^{N}x^2P(x,y)$	Shows the value of the relative distribution of high gray level values
8	Short Run Low Gray Level Emphasis (SRLGE)	$SRLGE = \frac{1}{r}\sum_{x=1}^{M}\sum_{y=1}^{N}\frac{P(x,y)}{x^2y^2}$	Shows the relative distribution of short runs and low gray levels
9	Short Run High Gray Level Emphasis (SRHGE)	$SRHGE = \frac{1}{r}\sum_{x=1}^{M}\sum_{y=1}^{N}\frac{P(x,y)x^2}{y^2}$	Shows the relative distribution of short runs and high gray levels
10	Long Run Low Gray Level Emphasis (LRLGE)	$LRLGE = \frac{1}{r}\sum_{x=1}^{M}\sum_{y=1}^{N}\frac{P(x,y)y^2}{x^2}$	Shows the relative distribution of long runs and low gray levels
11	Long Run High Gray Level Emphasis (LRHGE)	$LRHGE = \frac{1}{r}\sum_{x=1}^{M}\sum_{y=1}^{N}P(x,y)x^2y^2$	Shows the value of the relative distribution of long runs and high gray levels

The variables used in the GLRLM are r is number of pixel values run length matrix, M is the gray pixel value of an image, N is the number of consecutive pixels in an image, and $P(x,y)$ is run length matrix value in row x and column y.

B. Feature Reduction using Principal Component Analysis (PCA)

Feature reduction methods play a pivotal role in data-driven research and analysis, it aims to extract essential information while reduce the challenges posed by high-dimensionality datasets [25]. Principal Component Analysis (PCA) stands as a prominent technique in this domain. PCA, when applied to data, identifies and retains the most influential components, effectively reducing dimensionality while maintaining the core information [26]. This method calculates the mean and variance of the image, then computes the covariance matrix shown in Equation (1).

$$C = \hat{X}^T \times \hat{X} \quad (1)$$

where C is the product of covariance, and $\hat{X}^T$ is the matrix transpose of matrix $\hat{X}$. After that, it is continued by calculating the eigenvalues and eigenvectors shown in Equation (2). If the eigenvalue is below 0.5, it will be reduced. Then, the remaining variables are calculated in their eigenvector values and transformed.

$$Cv = \lambda v \quad (2)$$

Where C is a matrix of size $n \times n$, the vector v represents the eigenvector and the scalar λ represents the eigenvalue.

C. Kernel Extreme Learning Machine (KELM)

Kernel Extreme Learning Machine (KELM) is an Extreme Learning Machine (ELM) method developed by adding a kernel function where the number of layers in this method is the same as the ELM method. The idea of developing this method is from the Support Vector Machine (SVM) method, where SVM utilizes the kernel function to increase the dimensions of the data so that the data can be separated optimally [27]. This method does not need the mapping of features on the hidden layer and the selection of hidden neurons because the kernel function is the primary key of this method. Based on the SVM method's classification results, excellent results are obtained if there is a kernel function. The KELM method architecture is shown in Figure 3.

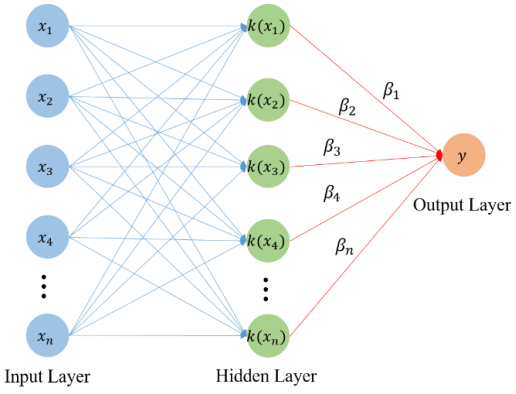


Figure 3. KELM Architecture

In this method, the kernel function calculates the inner product directly on a feature space with high dimensions without performing a mapping function first. The kernel function in the hidden layer can smooth data and simplify classification problems better. The Radial Basis Function (RBF) is a kernel function with only one hyperparameter, relieving model configuration and training costs [28]. The RBF kernel function is suitable for use on low and high dimensional data, so it is considered an ideal kernel function [29].

$$K(x, y) = e^{-\gamma \|x-y\|^2}$$

with x and y represents sample data, while γ it is a unique hyperparameter [30].

D. Genetic Algorithm

Optimization using Genetic Algorithms (GA) is an approach that adopts the principles of biological evolution to find optimal or near-optimal solutions to complex problems [31]. In GA optimization, an initial population of solutions is initialized randomly, and they are then evaluated using an objective function that measures the quality of the solution based on the goal to be achieved. Better solutions are judged based on how well they solve the problem and are given a chance to "survive" and "reproduce," while less good solutions will tend to be eliminated from the population [32]. These selected solutions reproduce and produce a new generation of solutions through selection, crossover, and mutation. This process continues iteratively through several generations, with the expectation that the overall quality of the solution will improve over time. By setting appropriate parameters such as population size, crossover probability, mutation probability, and iteration-stopping criteria, GA can be directed to find optimal or near-optimal solutions to various complex problems [33].

E. Confusion Matrix

The confusion matrix is a technique to determine the accuracy of actual data and predictive data in classifying [34]. In the confusion matrix, there are several terms to represent the results of the classification process. The results of the confusion matrix are shown in Figure 4.

In the confusion matrix, there are four terms, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TP represents data with a positive value on actual and predictive data, and TN represents data that has a negative value on actual and predictive data. FP represents positive data in the actual data and negative in the predicted data. Meanwhile, FN represents negative data in actual data and positive in predictive data.

In three classes classification, Covid-19, healthy lung, and Pneumonia class, the TP value is the healthy lung class data predicted to be true in the healthy lung class. The TN value is data for the Covid-19 and Pneumonia classes predicted to be correct in the Covid-19 and Pneumonia classes. The FP value is data in the Covid-19, and Pneumonia classes are predicted to be wrong in the healthy lung class. Meanwhile, the FN value is data in the healthy lung class, which was predicted wrong in the Covid-19 and Pneumonia classes.

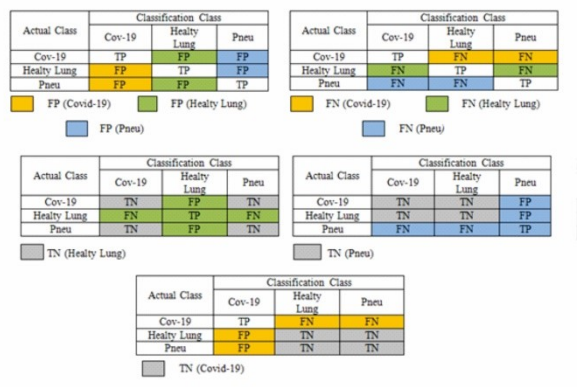


Figure 4. Multiclass Confusion Matrix (3 Classes)

In addition, there are also three parameters in measuring the level of truth of the classification process, namely accuracy, sensitivity, and specificity, which can be calculated by the formula shown by Equations (3), (4), and (5) [35].

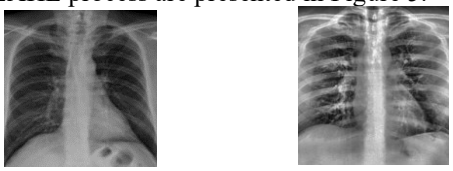
$$Accuracy = \frac{TP_{all}}{n_{all}} \times 100\% \quad (3)$$

$$Sensitivitas = \frac{\sum \left(\frac{TP}{TP + FN}\right)}{n} \times 100\% \quad (4)$$

$$Spesifisitas = \frac{\sum \left(\frac{TN}{TN + FP}\right)}{n} \times 100\% \quad (5)$$

3. RESULTS AND DISCUSSION

Classification using 900 X-ray image data consisting of 300 Covid-19 data, 300 healthy lung data, and 300 Pneumonia data. The first process is to increase the contrast of the image using CLAHE. The results of the CLAHE process are presented in Figure 5.



(a) (b)

Figure 5. (a) Example of Covid-19 Image Data; (b) CLAHE Process Results

After the process of increasing the contrast in the image, the next step is feature extraction using GLRLM with 11 features, namely SRE, LRE, GLN, RP, RLN, LGRE, HGRE, SRLGE, LRLGE, and LRHGE with degrees of 0°,45°,90°, and 135°. The results of feature extraction using are shown in Table 2.

Table 2. GLRLM Feature Extraction Result

Data	SRE	LRE	GLN	RLN	RP	LGRE	HGRE	SRLGE	SRHGE	LRLGE	LRHGE	Class
1	0.11	18.75	788.52	2000.28	4.20	0.21	386.43	0.03	18.46	3.62	12322.10	Covid-19
2	0.08	24.21	892.72	1967.45	4.93	0.19	132.25	0.02	9.01	4.93	3566.05	Healty lung
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
900	0.07	23.49	1938.74	2982.89	6.94	0.29	63.62	0.02	4.69	7.50	1426.70	*Pneumonia*

Based on the feature extraction results, the SRE feature tends to be high in the Covid-19 image, which means that the Covid-19 image has a smoother texture than the healthy lung and pneumonia image. The RLE feature shows a healthy lung image, and pneumonia has a coarser texture because the RLE value is high. The GLN, RLN, RP, HGRE, and SRHGE features show that the pneumonia image has a higher value than the Covid-19 image and is a healthy lung. The LGRE, SRLGE, and LRHGE features show healthy lung images with higher values. Meanwhile, the LRLGE, SRLGE, and LRHGE features show a healthy lung image with a higher value. Meanwhile, the LRLGE feature shows that the Covid-19 image is higher than healthy lung and pneumonia images. After feature extraction, the next step is feature reduction using PCA. The results of feature reduction using PCA with a threshold of 0.5 show that each angle of GLRLM feature extraction has different selected features. Features with a loading factor value of more than 0.5 are shown in Table 3.

Table 3. Selected Feature of PCA Algorithm

	0	45	90	135
SRE	0	0	0	0
LRE	0	1	1	1
GLN	1	0	0	0
RLN	1	1	1	1
RP	0	1	1	1
LGRE	1	1	1	1
HGRE	1	0	0	0
SRLGE	0	0	0	0
SRHGE	0	0	0	1
LRLGE	1	0	1	0
LRHGE	0	1	1	1

Based on Table 3, number 1 indicates the utilization of this feature in PCA results for a specific angle, represented by a gray column, with a loading factor exceeding the specified threshold. Conversely, 0 signifies either the non-utilization of the feature or a loading factor value below 0.5. The table illustrates the selection of features across different angles. The LRE feature is chosen at 45°, 90°, and 135°, while RLN is selected for all angles. GLN appears solely at 0°, whereas RP emerges at 90° and 135°. The LGRE feature is opted for across all angles except 0°, while HGRE exclusively appears at 0°. The SRHGE feature is designated for 135°, whereas LRLGE is allocated to 45° and 90°. LRHGE is designated for all angles except 0°. Notably, SRE and SRLGE were not selected from any angle. The dataset undergoes division into training and testing sets utilizing k-fold cross-validation with k=5. Following the feature selection, classification proceeds employing various ELM development methodologies, including ELM, KELM, and GA-KELM, each experiment presented in Table 4.

Table 4. Evaluation System of Each Experiment

		Angle	Accuracy (%)	Sensitivity (%)	Spesificity (%)
Without PCA	GLRLM-ELM	0	78.89	80.50	78.89
		45	77.22	83.70	77.22
		90	74.44	79.99	74.44
		135	78.89	83.75	78.89
	GLRLM-KELM	0	79.33	81.14	77.29
		45	80.45	84.89	76.71
		90	78.21	82.78	74.83
		135	79.33	81.83	76.57
With PCA	GLRLM-PCA-ELM	0	90.00	90.91	90.00
		45	91.11	91.14	91.11
		90	91.67	91.93	91.67
		135	90.00	90.60	90.00
	GLRLM-PCA-KELM	0	95.00	95.39	95.00
		45	95.56	95.68	95.56
		90	95.00	95.39	95.00
		135	95.00	95.39	95.00
	GLRLM-PCA-Optimized KELM	0	96.11	96.11	96.11
		45	97.22	97.23	97.22
		90	**97.78**	**97.78**	**97.78**
		135	96.67	96.70	96.67

At each corner trial, the gamma value in KELM is optimized to get the best solution function. At angles of 0, 45, and 135, the optimal gamma value using the KELM method is 1, while at an angle of 90, the optimal gamma value based on the results of the GA method is equal to 1.35. This illustrates the importance of adjusting the gamma value to certain angles in achieving optimal performance when using the KELM method. The gamma value in the KELM method, which is not optimized, is determined to be equal to 3 in each KELM trial. The table presents the performance results of several classification methods (GLRLM-ELM, GLRLM-KELM, GLRLM-PCA-ELM, GLRLM-PCA-KELM, and GLRLM-PCA-Optimized KELM) in classifying data at various angles (0°, 45°, 90°, and 135°).

Without using PCA, the GLRLM-ELM and GLRLM-KELM models show a relatively low level of accuracy, with an average accuracy of about 77.36% and 79.33%, respectively. However, when PCA was applied, model performance was significantly increased, especially in GLRLM-PCA-Optimized KELM, which achieved the highest accuracy at 96.94%. The accuracy result shows that the use of PCA in feature processing has produced a more informative feature representation and can improve the classification model's performance. In this context, GLRLM-PCA-Optimized KELM stands out as the most effective model in classifying data, especially at 90° angles, with an accuracy of more than 97%. It shows that hyperparameter optimization in the kernel function can significantly increase accuracy with an accuracy difference of more than 1%. Visualization of the average accuracy in Table 4 is shown in Figure 5.

Fig 6 illustrates the performance of various classification methods, namely GLRLM-ELM, GLRLM-KELM, GLRLM-PCA-ELM, GLRLM-PCA-KELM, and GLRLM-PCA-Optimized KELM, in analyzing data using various texture features. There was a consistent improvement in accuracy, sensitivity, and specificity from GLRLM-ELM to GLRLM-PCA-Optimized KELM, where the latter model showed the highest performance with average accuracy, sensitivity, and specificity values of 97%. It indicates that using principal component analysis (PCA) and model optimization using the KELM method simultaneously can improve classification performance, showing significant potential in texture feature analysis for data classification applications. Table 5 compares this research's evaluation results with previous X-ray (CXR) Image classification research.

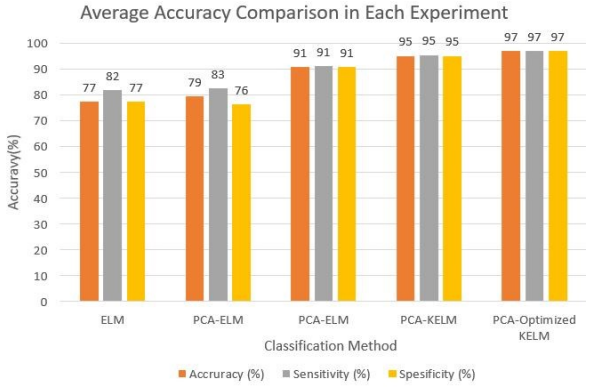


Figure 6. Average accuracy, sensitivity, and spesificity in each experiment

Table 5. Comparison Evaluation

Method	Accruracy (%)	Sensitivity (%)	Spesificity (%)
Texture Analysis Algorithm for Digital Test of COVID-19 Patients [36]	94.43	95.00	93.86
Texture Features with Ensemble Classification Method [37]	93.51	92.92	96.43
Hybrid artificial intelligence models (CNN-Softmax) [38]	95.20	93.30	100.00
Propose Method	**97.78**	**97.78**	**97.78**

Table 5 presents the performance evaluation results of several methods in digital testing for COVID-19 patients based on accuracy, sensitivity, and specificity. Research [36] achieved 94.43% accuracy, 95.00% sensitivity, and 93.86% specificity. Meanwhile, research [37] achieved 93.51% accuracy, 92.92% sensitivity, and 96.43% specificity. Research [38] achieved 95.20% accuracy, 93.30% sensitivity, and 100.00% specificity. Meanwhile, the proposed method achieved the highest performance with accuracy, sensitivity, and specificity of 97.78%, respectively. Analysis of this table shows that the proposed method performs better than other methods in testing COVID-19 in patients. In future research, accuracy can be increased by exploring several development methods in the identification process. Further research can be carried out using CNN feature learning as feature extraction in images, as research [39] shows that CNN performs better in feature extraction compared to conventional methods. Apart from that, other ELM methods can also be developed to obtain good accuracy. Research [40], shows that the development of the ELM method, namely the DELM method, performs better than the KELM method.

4. CONCLUSION

The results of this research show that the use of GLRLM texture analysis and Kernel Extreme Learning Machine (KELM) optimization with genetic algorithms (GA), as well as feature reduction using principal component analysis (PCA), can simultaneously improve data classification performance, especially in identifying Covid-19 from images X-ray of the lungs. Feature selection and model optimization found that GLRLM-PCA-Optimized KELM stands out as the most effective model with an accuracy of over 97%, especially at 90° angles. These results show that hyperparameter optimization in the kernel function can significantly improve classification accuracy, with an accuracy difference of more than 1%. Thus, using the PCA method in feature processing and KELM model optimization shows important potential in texture feature analysis for data classification applications, especially in COVID-19 diagnosis using Chest X-rays.

5. REFERENCES

[1] C. C. Lai, T. P. Shih, W. C. Ko, H. J. Tang, and P. R. Hsueh, "Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges," *Int. J. Antimicrob. Agents*, vol. 55, no. 3, p. 105924, 2020, doi: 10.1016/j.ijantimicag.2020.105924.

[2] S. A. Hassan, F. N. Sheikh, S. Jamal, J. K. Ezeh, and A. Akhtar, "Coronavirus (COVID-19): a review of clinical features, diagnosis, and treatment," *Cureus*, vol. 12, no. 3, 2020.

[3] P.-I. Lee and P.-R. Hsueh, "Emerging threats from zoonotic coronaviruses-from SARS and MERS to 2019-nCoV," *J. Microbiol. Immunol. Infect.*, vol. 53, no. 3, p. 365, 2020.

[4] A. Calderaro, M. Buttrini, B. Farina, S. Montecchini, F. De Conto, and C. Chezzi, "Respiratory tract infections and laboratory diagnostic methods: A review with a focus on syndromic panel-based assays," *Microorganisms*, vol. 10, no. 9, p. 1856, 2022.

[5] W. Khan, N. Zaki, and L. Ali, "Intelligent Pneumonia Identification from Chest X-Rays:

A Systematic Literature Review," *IEEE Access*, vol. 9, pp. 51747–51771, 2021, doi: 10.1109/ACCESS.2021.3069937.

[6] N. Akand, P. K. Sarkar, M. J. Alam, M. Kamruzzaman, M. M. Hossain, and M. A. Islam, "Mothers Knowledge Related To Preventive Measure of Pneumonia in Hospitalized Children Under 5 Years Age: A Tertiary Care Center Experience," *IOSR J. Nurs. Heal. Sci.*, vol. 9, no. 2, p. 7, 2020.

[7] M. Sobolewska-Pilarczyk *et al.*, "COVID-19 infections in infants," *Sci. Rep.*, vol. 12, no. 1, p. 7765, 2022.

[8] A. Kovács, P. Palásti, D. Veréb, B. Bozsik, A. Palkó, and Z. T. Kincses, "The sensitivity and specificity of chest CT in the diagnosis of COVID-19," *Eur. Radiol.*, vol. 31, pp. 2819–2824, 2021.

[9] T. Tang, M. Zhang, and A. S. Mujumdar, "Intelligent detection for fresh-cut fruit and vegetable processing: Imaging technology," *Compr. Rev. Food Sci. Food Saf.*, vol. 21, no. 6, pp. 5171–5198, 2022.

[10] H.-G. Kim, K. M. Lee, E. J. Kim, and J. San Lee, "Improvement diagnostic accuracy of sinusitis recognition in paranasal sinus X-ray using multiple deep learning models," *Quant. Imaging Med. Surg.*, vol. 9, no. 6, p. 942, 2019.

[11] M. Shehata *et al.*, "A New Computer-Aided Diagnostic (CAD) System for Precise Identification of Renal Tumors," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 2021, pp. 1378–1381.

[12] W.-N. Mohd-Isa, J. Joseph, N. Hashim, and N. Salih, "Enhancement of digitized X-ray films using Contrast-Limited Adaptive Histogram Equalization (CLAHE)," *F1000Research*, vol. 10, no. 1051, p. 1051, 2021.

[13] R. K. Hapsari, M. I. Utoyo, R. Rulaningtyas, and H. Suprajitno, "Comparison of Histogram Based Image Enhancement Methods on Iris Images," *J. Phys. Conf. Ser.*, vol. 1569, no. 2, 2020, doi: 10.1088/1742-6596/1569/2/022002.

[14] P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," *Pattern Recognit. Lett.*, vol. 141, pp. 61–67, 2021.

[15] A. Moawad *et al.*, "A Real-Time Energy and Cost Efficient Vehicle Route Assignment Neural Recommender System," *arXiv*, no. Figure 1, pp. 1–14, 2021.

[16] D. C. R. Novitasari, A. Lubab, A. Sawiji, and A. H. Asyhar, "Application of feature extraction for breast cancer using one order statistic, GLCM, GLRLM, and GLDM," *Adv. Sci. Technol. Eng. Syst. J.*, vol. 4, no. 4, pp. 115–120, 2019.

[17] S. Harnale and D. D. Maktedar, "Oral Cancer Detection: Feature Extraction & SVM Classification," *Int. J. Adv. Netw. Appl.*, vol. 11, no. 03, pp. 4294–4297, 2019, doi: 10.35444/ijana.2019.11036.

[18] E. Teye, C. L. Y. Amuah, T. McGrath, and C. Elliott, "Innovative and rapid analysis for rice authenticity using hand-held NIR spectrometry and chemometrics," *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.*, vol. 217, pp. 147–154, 2019, doi: https://doi.org/10.1016/j.saa.2019.03.085.

[19] A. K. Idrees and A. K. M. Al-Qurabat, "Energy-efficient data transmission and aggregation protocol in periodic sensor networks based fog computing," *J. Netw. Syst. Manag.*, vol. 29, no. 1, pp. 1–24, 2021.

[20] L. Xiao, W. Shao, F. Jin, and Z. Wu, "A self-adaptive kernel extreme learning machine for short-term wind speed forecasting," *Appl. Soft Comput.*, vol. 99, p. 106917, 2021.

[21] W. Chai, Y. Zheng, L. Tian, J. Qin, and T. Zhou, "GA-KELM: Genetic-algorithm-improved kernel extreme learning machine for traffic flow forecasting," *Mathematics*, vol. 11, no. 16, p. 3574, 2023.

[22] M. Daviran, M. Shamekhi, R. Ghezelbash, and A. Maghsoudi, "Landslide susceptibility prediction using artificial neural networks, SVMs and random forest: hyperparameters tuning by genetic optimization algorithm," *Int. J. Environ. Sci. Technol.*, vol. 20, no. 1, pp. 259–276, 2023.

[23] V. Mishra and S. K. Rath, "Detection of breast cancer tumours based on feature reduction and classification of thermograms," *Quant. Infrared Thermogr. J.*, vol. 18, no. 5, pp. 300–313, 2021.

[24] H. Zhang, C. L. Hung, G. Min, J. P. Guo, M. Liu, and X. Hu, "GPU-Accelerated GLRLM Algorithm for Feature Extraction of MRI," *Sci. Rep.*, vol. 9, no. 1, pp. 1–13, 2019, doi: 10.1038/s41598-019-46622-w.

[25] L. Zhang and J. Wen, "A systematic feature selection procedure for short-term data-driven building energy forecasting model development," *Energy Build.*, vol. 183, pp. 428–442, 2019.

[26] S. Bhattacharya *et al.*, "A novel PCA-firefly based XGBoost classification model for intrusion detection in networks using GPU," *Electronics*, vol. 9, no. 2, p. 219, 2020.

[27] N. Pratiwi and Y. Setyawan, "Analisis Akurasi Dari Perbedaan Fungsi Kernel Dan Cost Pada Support Vector Machine Studi Kasus Klasifikasi Curah Hujan Di Jakarta," *J. Fundam. Math. Appl.*, vol. 4, no. 2, pp. 203–212, 2021, doi: 10.14710/jfma.v4i2.11691.

[28] G. Tang, L. Yang, S. Ren, L. Meng, F. Yang, and H. Wang, "An automatic source code vulnerability detection approach based on KELM," *Secur. Commun. Networks*, vol. 2021,

2021.

[29] H. Shi, H. Xiao, J. Zhou, N. Li, and H. Zhou, "Radial basis function kernel parameter optimization algorithm in support vector machine based on segmented dichotomy," in *2018 5th International Conference on Systems and Informatics (ICSAI)*, 2018, pp. 383–388.

[30] D. Jahed Armaghani, P. G. Asteris, B. Askarian, M. Hasanipanah, R. Tarinejad, and V. Van Huynh, "Examining hybrid and single SVM models with different kernels to predict rock brittleness," *Sustainability*, vol. 12, no. 6, p. 2229, 2020.

[31] A. J. Fofanah, S. Koroma, and H. I. Bangura, "Experimental Exploration of Evolutionary Algorithms and their Applications in Complex Problems: Genetic Algorithm and Particle Swarm Optimization Algorithm," *Br. J. Healthc. Med. Res.*, vol. 10, no. 2, 2023.

[32] S. Yacoubi, G. Manita, A. Chhabra, O. Korbaa, and S. Mirjalili, "A multi-objective chaos game optimization algorithm based on decomposition and random learning mechanisms for numerical optimization," *Appl. Soft Comput.*, p. 110525, 2023.

[33] H. M. Asih, R. A. C. Leuveano, and D. A. Dharmawan, "Optimizing lot sizing model for perishable bread products using genetic algorithm," *J. Sist. dan Manaj. Ind.*, vol. 7, no. 2, pp. 139–154, 2023.

[34] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix," *Pattern Recognit.*, vol. 91, pp. 216–231, 2019, doi: 10.1016/j.patcog.2019.02.023.

[35] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. W. Quinn, and M. A. Moni, "Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison," *Comput. Biol. Med.*, vol. 136, p. 104672, 2021.

[36] D. Al-Karawi, S. Al-Zaidi, N. Polus, and S. Jassim, "Ai based chest x-ray (cxr) scan texture analysis algorithm for digital test of covid-19 patients," *medRxiv*, pp. 2005–2020, 2020.

[37] H. Alquran, M. Alsleti, R. Alsharif, I. A. Qasmieh, A. M. Alqudah, and N. H. B. Harun, "Employing texture features of chest x-ray images and machine learning in covid-19 detection and classification," in *Mendel*, 2021, vol. 27, no. 1, pp. 9–17.

[38] A. M. Alqudah, S. Qazan, H. Alquran, I. A. Qasmieh, and A. Alqudah, "COVID-19 detection from x-ray images using different artificial intelligence hybrid models," *Jordan J. Electr. Eng.*, vol. 6, no. 2, pp. 168–178, 2020.

[39] S. Ali, J. Li, Y. Pei, M. S. Aslam, Z. Shaukat, and M. Azeem, "An effective and improved cnn-elm classifier for handwritten digits recognition and classification," *Symmetry (Basel).*, vol. 12, no. 10, pp. 1–15, 2020, doi: 10.3390/sym12101742.

[40] D. C. R. Novitasari *et al.*, "Image Fundus Classification System for Diabetic Retinopathy Stage Detection Using Hybrid CNN-DELM," *Big Data Cogn. Comput.*, vol. 6, no. 4, p. 146, 2022.