Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics e-ISSN: 2656-8624
Homepage: https://ijeeemi.org/; Vol. 7, No. 2, pp. 447-460, May 2025

RESEARCH PAPER

OPEN ACCES

Hybrid Features to Classify Lung Tumor Using Machine Learning

Rizki Dwi Rahmawan, Umi Salamah, Ery Permana Yudha
Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia

ABSTRACT

A lung tumor is an abnormal mass of cells inside a body. As a benign tumor is unproblematic, but a malignant tumor is cancerous because it can travel across the body and interfere with its surrounding tissue. Detecting these cancerous cells in the lung is important because delayed detection may hamper effective treatment options, leading to a lower survival rate. However, classifying tumor malignancy is highly dependent on the knowledge and experience of the radiologist. This study combines texture-based features extracted from lung Computed Tomography Scan (CT Scan) images such as Gray Level Co-occurrence Matrix (GLCM), Gray Level Run-length Matrix (GLRLM), Gray Level Size-zone Matrix (GLSZM), and Haralick Features aims to create a lung tumor classification system. This research contributes by creating an efficient and reliable system through Relief-F feature selection that uses features with the highest weight in rank that are able to differentiate classes of tumor malignancy and help medical professionals diagnose tumors more early in the treatment. As a comparison, several conventional machine learning classifiers, including SVM RBF, KNN, RF, DT, and XGBoost, were utilized to evaluate classifier performance. The result showed that the accuracy of the proposed hybrid features with a random forest classifier was the most performing approach with an evaluation score of accuracy of 99.55%, precision of 99.55%, recall of 99.55%, and F1-Score of 99.54%. Furthermore, accuracy among other classifiers was also higher than 90%. Proofing the selected features retain essential class information, demonstrating the study's applicability in developing automated lung tumor classification systems from CT scans.

PAPER HISTORY

Received March 10, 2025
Accepted May 29, 2025
Published May 30, 2025

KEYWORDS

Lung Cancer Tumor,
Machine Learning,
Classification,
Texture Feature Extraction,
Hybrid features,
Feature Selection

AUTHOR EMAIL

rizkidwirahmawan@student.uns.ac.id
umisalamah@staff.uns.ac.id
erypermana@staff.uns.ac.id

I. INTRODUCTION

A tumor is a group of abnormal cells in a body. Cells are considered tumors when they divide more than usual and do not die as they should. A Benign tumor stays in place without attacking other areas in the body, inversely, a malignant tumor is cancerous because it grows uncontrollably and spreads to other areas through the bloodstream or lymphatic system [1]. Lung cancer itself is a type of cancer with the highest mortality and incidence than other type of cancer in the world. In 2022, there were 2,480,675 lung cancer cases, with a total death of 1,817,469 [2]. In Indonesia in the same year, lung cancer became the most common type of cancer in men with 29,107 new cases meeting 15.4% of all types of cancer in Indonesia, In addition, lung cancer became the second most common type of cancer after breast cancer with 38,904 new cases [3]. One obstacle in lung cancer care is the difficulty of early diagnosis. These malignant cancerous tumors in the lungs need to be detected quickly to avoid spreading [1]. A delayed diagnosis may limit treatment possibilities, thus, reducing survival chances [4]. Currently, chest radiographs are the primary radiological examination used to diagnose suspected lung cancer, and CT Scan is a more effective radiological examination for the diagnosis of lung cancer [5]. However, behind its effectiveness, radiologists' expertise, experience, and analytical skills are crucial for tumor malignancy classification. Thus, an automated process is needed that can support radiologists to classify tumors with high accuracy and through non-invasive procedures [6].

Many researchers have attempted to create a model for the classification and segmentation of lung cancer, including deep learning. A study by [7] used Ebola Optimization Search Algorithm (EOSA)-Convolutional Neural Network (CNN) architecture, resulting in an accuracy of 93,21% and an F1-Score of 92,72%. [8] utilized other architecture, namely fine-tuned VGG-19 with Principal Component Analysis (PCA) and Adaptive Neuro-Fuzzy Inference System (ANFIS), resulting in an accuracy of 92,73% and an F1-Score of 91,59%. Another approach is taken by [9], using an autoencoder to extract features and SVM for the classification. This method achieved an accuracy of 97% and F1-Score of 96%. Deep learning provides valuable insights, but the complexity of many models restricts their use in real-time settings.

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

Designing lightweight models optimized for edge computing holds the potential to reshape lung cancer diagnostics. Ensuring model development considers both optimal efficiency and adaptability for real-world medical use. [10].

The technique that can be applied to solve this problem is to perform feature extraction and feature selection. The texture feature extraction method is widely used because it is relatively easier to implement [11]. Moreover, the texture feature is important because it can interpret the micro and macro-level features in the image [12]. Whereas, a good feature selection can maintain the physical perception of the original feature by retaining some key features and can reduce computing costs and improve classification performance [13]. In [14], features used are GLCM and Gabor Filter, the research achieved accuracy of 89,89% and 97,94% of F1-Score. [15] take a different approach combining GLCM features and Local Binary Pattern (LBP) getting accuracy of 93% and 92,50% F1-Score. Another manual texture extraction is in research [16], using Binary Count Ratio (BCR) feature and Euclidean distance as the classifier, as a result, this approach achieved an accuracy of 98.32% and an F1-score of 97.05%.

Although many types of features are already utilized for lung tumor classification, there is no clear result of what features are important for the classification process. GLCM, GLRLM, GLSZM and Haralick features has shown a promising result in [17] where from previous research, the study able to increase accuracy in classifying non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) up to 9.49% showing its capability in differentiate lung cancer type. Thus, by using feature selection, this research aims to fill this gap by combining texture-based features such as GLCM, GLRLM, GLSZM, and Haralick Features and provide information on important features in lung tumor classification and its effectiveness in classifying lung tumor from CT Scan images using conventional machine learning classifiers, including support vector machine radial base function (SVM RBF), KNN, RF, XGBoost, and DT. The contributions of this study are 1) Combining texture-based features as a feature vector for classification, 2) Feature efficiency in classification through Relief-F feature selection, which minimizes the size of the feature vector, 3) Better understanding of features that are useful for lung tumor classification, 4) a lightweight method of classifying lung cancer, therefore increasing accessibility in places with limited resources.

The study is divided into distinct sections. In Section II, the dataset, methods, and classification approaches are reviewed. Section III outlines the results. Section IV offers an analysis, comparing findings with existing literature and discussing limitations. Section V concludes the study, highlighting primary insights and potential future work.

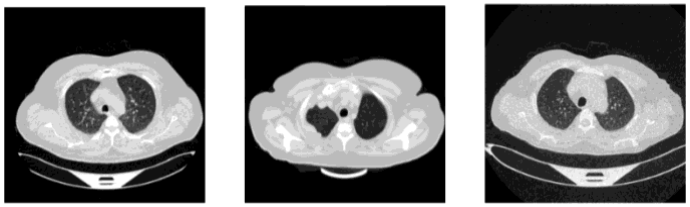
II. Materials and Method

I. Dataset Pre-Processing

The IQ-OTH/NCCD lung cancer dataset [18] used in this research is from Kaggle and publicly available to download at https://www.kaggle.com/hamdallak/the-iqothnccd-lung-cancer-dataset. Previous research and study mentioned earlier also use the same dataset. This dataset contains three label of lung CT Scan: malignant, benign, and normal. The dataset was acquired over three months during fall 2019 at The Iraq-Oncology Teaching Hospital/National Center for Cancer Disease, comprising 1097 images from 110 cases marked by an oncologist and a radiologist. All images are grayscaled without any weighting. Distribution of each classes are shown in Table 1. Fig. 1 displays dataset's sample image of each class.

Table 1. Dataset Class Distribution

Class	Number of Samples
Malignant	561
Normal	416
Benign	120



(a) (b) (c)

Fig. 1. Dataset's image sample of each class, (a) Benign, (b) Malignant, (c) Normal.

II. Feature Extraction

Rather than depending on a single type of feature, this research utilized a hybrid approach that incorporates a diverse type of texture-based features.

1. Gray Level Co-occurrence Matrix (GLCM)

GLCM applies statistical methodologies to extract meaningful features. GLCM interprets pixel spatial relationships using a gray-level dependency matrix [19]. This co-occurrence matrix is formed in four specific adjacencies (0°, 45°, 90°, and 135°). Then, with the process of averaging each of those orientations, the final GLCM matrix is generated [20]. 24 GLCM features used in this research are shown in Table 2.

2. Gray Level Run-length Matrix (GLRLM)

GLRLM uses a computational approach for the amount of gray level runs at varied lengths. A gray level runs is a collection of linearly adjacent points or pixels of images with the same grayscale value. The GLRLM matrix has a size $a \times b$ where a depicts the highest intensity value in the grayscale, and b is the maximum run length [21]. 16 GLRLM features used in this study are shown in Table 3.

3. Gray Level Size-zone Matrix (GLSZM)

GLSZM calculates the number of groups or zones of neighboring pixels that are interconnected with the same degree of grayness to form the basis of a matrix. The more homogeneous the texture, the wider and flattering the matrix produced [22]. A size-zone is defined as an

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

interconnected voxel with a distance of 1. 15 GLSZM features used in this research are shown in Table 4.

Table 2. List of GLCM Features

No.	Feature Name
1	Informational Measure of Correlation (IMC) 1
2	Inverse Difference (ID) / Homogeneity 1
3	Inverse Difference Normalized (IDN)
4	Sum Squares
5	Cluster Prominence
6	Informational Measure of Correlation (IMC) 2
7	Correlation
8	Cluster Shade
9	Maximal Correlation Coefficient (MCC)
10	Cluster Tendency
11	Joint Energy
12	Inverse Difference Moment Normalized (IDMN)
13	Autocorrelation
14	Joint Average
15	Inverse Difference Moment (IDM) / Homogeneity 2
16	Sum Average
17	Inverse Variance
18	Contrast
19	Difference Average
20	Joint Entropy
21	Maximum Probability
22	Difference Entropy
23	Sum Entropy
24	Difference Variance

4. Haralick Features

The haralick texture feature is a statistical feature used to describe the overall texture of an image [23]. The haralick feature is obtained from GLCM, which examines the spatial interactions among adjacent pixels. Rotational invariance is an important criterion for features extracted from images due to their resistance to the rotated image. By computing the mean value at each orientation (0°, 45°, 90°, and 135°), the Haralick feature becomes invariant through this averaging process. This procedure guarantees that features remain stable and dependable nonetheless of image orientation, thereby fortifying reliability of the analysis [24]. 13 Haralick features used in this research are shown in Table 5.

A. Feature Selection

Relief-F is utilized in this research as shown in schematic diagram in Fig. 2. Relief-F operates by randomly choosing samples and identifying the nearest neighbors within the same class. Then, for the selected sample, its value is compared against the hits and misses of the k nearest neighbors, and then the relevance value of each feature is updated. [25]. In Relief-F, feature weight is an important metric used to measure the relationship between the distance value between the original feature set and feature A [26].

$$w[A] = w[A] - \frac{\sum_{j=1}^{k} diff(A,R,H_j)}{mk} + D(C) \quad (1)$$

$$D(C) = \sum_{C \neq class(R)} [\frac{P(C)}{1-P(class(R))} \sum_{j=1}^{k} diff(A,R,M_j(C))]/(mk) \quad (2)$$

The equation for Relief-F is shown in Eq. (1) to Eq. (3) [26] where w[A] is feature weight and D(C) is the weighting equation for the misses instances. Let R be a random sample drawn from the dataset, m be the total number of sample randomly selected, and k be number of nearest neighbors considered. The term class(R) identifies the class to which sample R belongs. Furthermore, P(C) represents the fraction of the number of samples' selected features to the overall total of features across all samples' classes.

Table 3. List of GLRLM Features

No.	Feature Name
1	Short Run High Gray Level Emphasis (SRHGLE)
2	Gray Level Non-Uniformity Normalized (GLNN)
3	Long Run High Gray Level Emphasis (LRHGLE)
4	Run Length Non-Uniformity Normalized (RLNN)
5	Run Entropy (RE)
6	Short Run Emphasis (SRE)
7	Run Percentage (RP)
8	Long Run Emphasis (LRE)
9	Low Gray Level Run Emphasis (LGLRE)
10	Gray Level Variance (GLV)
11	Long Run Low Gray Level Emphasis (LRLGLE)
12	Gray Level Non-Uniformity (GLN)
13	Run Length Non-Uniformity (RLN)
14	. Short Run Low Gray Level Emphasis (SRLGLE)
15	Run Variance (RV)
16	High Gray Level Run Emphasis (HGLRE)

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

Table 4. List of GLSZM Features

No.	Feature Name
1	High Gray Level Zone Emphasis (HGLZE)
2	Small Area High Gray Level Emphasis (SAHGLE)
3	Large Area High Gray Level Emphasis (LAHGLE)
4	Gray Level Non-Uniformity (GLN)
5	Large Area Low Gray Level Emphasis (LALGLE)
6	Gray Level Variance (GLV)
7	Gray Level Non-Uniformity Normalized (GLNN)
8	Size-Zone Non-Uniformity Normalized (SZNN)
9	Low Gray Level Zone Emphasis (LGLZE)
10	Zone Percentage (ZP)
11	Size-Zone Non-Uniformity (SZN)
12	Large Area Emphasis (LAE)
13	Small Area Emphasis (SAE)
14	Zone Entropy (ZE)
15	Small Area Low Gray Level Emphasis (SALGLE)

Table 5. List of Haralick Features

No.	Feature Name
1	Sum Variance
2	Sum of Squares : Variance
3	Angular Second Moment
4	Difference Variance
5	Informational Measure of Correlation 2
6	Inverse Difference Moment
7	Correlation
8	Entropy
9	Sum Average
10	Informational Measure of Correlation 1
11	Sum Entropy
12	Contrast
13	Difference Entropy

In this framework, H denotes the k-nearest neighbor sample that comes from the same class as R, while M refers to the k-nearest neighbor drawn from a different class than R. Moreover, the expression $diff(\mathrm{A}, \mathrm{R}_1, \mathrm{R}_2)$ quantifies the difference, for feature A, between R_1 and R_2. Finally, $M_j(C)$ specifies the jth k-nearest neighbor within class C. The formula for $diff(\mathrm{A}, \mathrm{R}_1, \mathrm{R}_2)$ is provided as Eq. (3). [26]

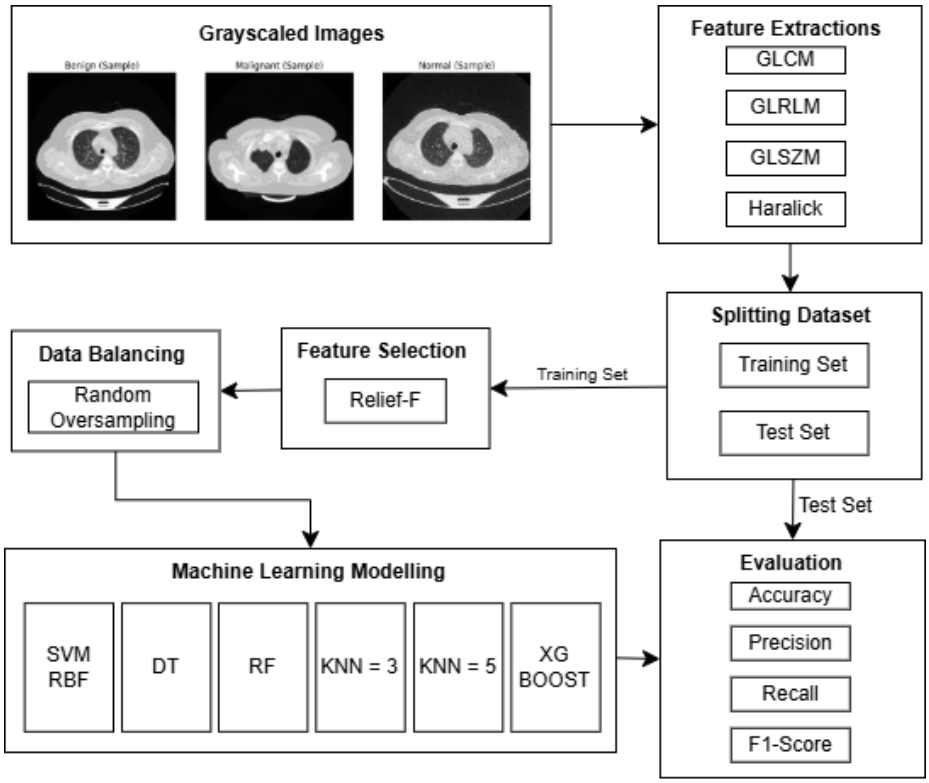


Fig. 2. Schematic Diagram of Proposed Method

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

$$diff(A,R_1,R_2)=\begin{cases}\frac{|R_1[A]-R_2[A]|}{\max(A)-\min(A)} & continuous\\ 0 & discrete, R_1[A]\neq R_2[A]\\ 1 & discrete, R_1[A]=R_2[A]\end{cases} \quad (3)$$

B. Data Balancing

Random Oversampling (ROS) is implemented to balance the dataset's class distribution. A naïve resampling method that does not use any heuristic and data knowledge and by randomly duplicating existing minority instances or synthesizing new ones, the method balances the dataset until the target class distribution is reached. [27]. The increased number of minority sample create a balance between classes [28]. Because it duplicates multiple instances, using ROS may lead to overfitting [29].

C. Classification

After the texture features of the CT Scan are extracted and selected, those features are inserted into machine learning models to be trained. This research uses five types of machine learning: SVM RBF, DT, RF, KNN, and XGBoost. These methods are also shown in Fig. 2.

1. SVM Radial Base Function (SVM RBF)

SVM is one of supervised machine learning algorithms. The researchers used SVM because its hyperplanes can differentiate classes. SVM employs the principle of maximum margin classification and, thanks to its deterministic nature, SVM avoids local minima [30]. The dual form of SVM mini, SVM is calculates as a minimization problem shown in Eq. (4) [31]. While, The RBF kernel is expressed in Eq (5) [31]. Fig. 3 displays visualization of SVM.

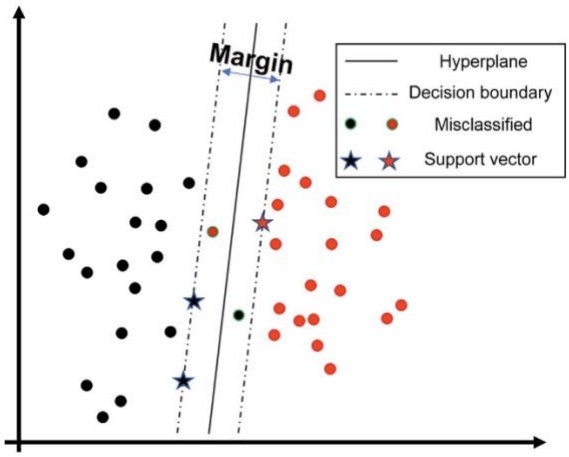


Fig. 3. SVM visualization without kernel function [32]

$$\min\frac{1}{2}\sum_{ij}y_i\alpha_i y_j\alpha_j K(x_i x_j)-\sum_i\alpha_i$$
$$subject\ to\ \textstyle\sum_i y_i\alpha_i=0$$
$$with\ 0\le\alpha_i\le C \quad (4)$$
$$K(x_i x_j)=\exp\left(-\gamma\|x_i-x_j\|^2\right) \quad (5)$$

Let x represents input feature or data point and y is class label associated with training features. The parameter γ controls the rate at which a point's influence decreases as the distance increases. With the increase in γ, The decision surface grows curvier, fitting tightly to the training data. A smaller γ value will create a flatter decision surface, and a simpler model [31].

2. Decision Tree (DT)

DT is a classification method that does not require predefined parameters, making it an effective choice for handling multiclass classification. DT constructs a tree-shaped decision framework using training data, beginning with root and leaf nodes referred to as internal nodes. At each node, the algorithm selects the most relevant feature along with its associated threshold value [32]. Fig. 4 displays visualization of DT.

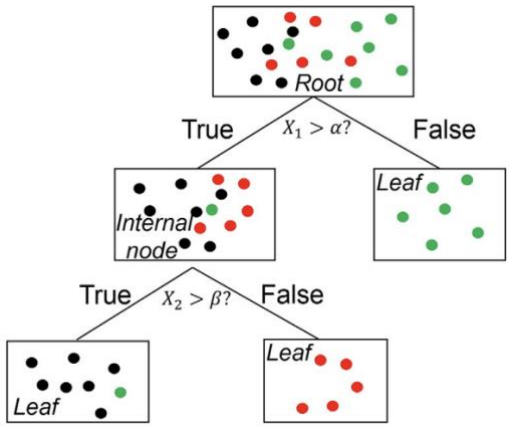


Fig. 4. DT visualization with two features [32]

3. Random Forest (RF)

RF is one of ensemble learning method. RF constructs more than one decision trees in parallel using bootstrapping and then aggregates their predictions, a process known as bagging. Bootstrapping involves training multiple individual DTS simultaneously on diverse different sets of the training data, utilizing different sets of features [32]. Fig. 5 displays visualization of random forest classification.

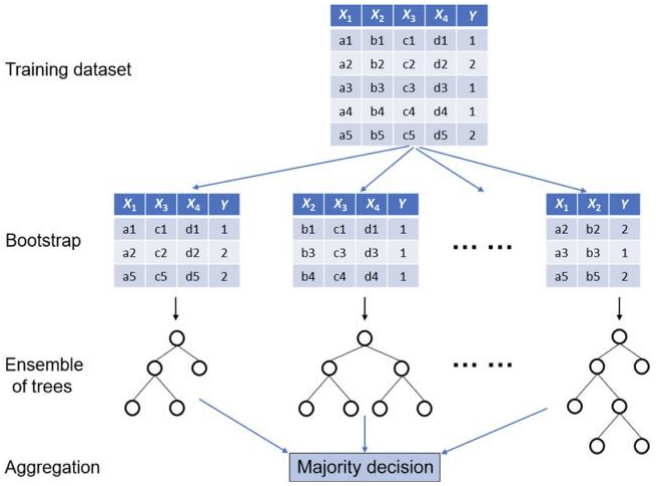


Fig. 5. RF visualization with four features (X_1,X_2,X_3,X_4) and two classes (Y=1,2) [32]

4. K-Nearest Neighbor (KNN)

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

KNN is an algorithm that is widely used for supervised learning technique. KNN makes predictions by analyzing the real-time data structure as soon as it is introduced to new data, without the need for an explicit training phase beforehand. KNN works on the principle of probability similarity, where data points sharing similar characteristics naturally cluster together, which means that predictions are driven by how closely new data aligns with the clusters formed by the training set. [33].

5. XGBoost

XGBoost is a machine learning ensemble that leverages gradient boosting. The process begins with making many bootstrap samples, then, learning is carried out using DT, where the residue from a tree is calculated and used as a weight for the successor tree. This update is repeated until all learning is done taking into account the residue, resulting in lower errors as the update occurs along the direction of the negative gradient. XGBoost leverages gradient-based optimization to fine-tune its loss function while incorporating regularization techniques that curb overfitting and manage the model's complexity. [34].

D. Evaluation

Evaluation metrics used in this research are shown in Eq. (6) to Eq. (9) [35]. In evaluating the performance of loss functions in the problem of data scarcity, several metrics that can be considered include sensitivity, accuracy, F1-Score, precision, and area under the curve [36].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

$$F1_{score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

A True Positive (TP) counts the instances that actually belong to a certain positive class and are correctly identified as positive. Conversely, a True Negative (TN) is the number of instances outside that positive class which are accurately classified as negative. A False Positive (FP) occurs when an instance that does not belong to the positive class is mistakenly predicted as positive, while a False Negative (FN) happens when an instance that should be considered positive is incorrectly labeled as negative.

Accuracy represents the ratio of correctly classified images to the total number of images evaluated. Precision indicates the proportion of true positive predictions among all images identified as positive. Sensitivity (or recall) reveals the percentage of actual positive cases that were correctly detected. The F1-Score is computed by taking the harmonic mean of precision and recall, which delivers a balanced assessment of both metrics.

III. RESULTS

A. Feature Extraction

Feature extraction is carried out after data acquisition and data preprocessing in the form of image grayscaling to retrieve texture information contained in the image. This research uses a combination of GLCM, GLRLM, GLSZM, and Haralick features. Extraction is done on the entire image without a segmentation process and uses the whole section as a region of interest (ROI). The extraction process produces data in the form of comma-separated values containing each feature's values for each image.

1. GLCM Features

The GLCM Texture Features in this study amounted to 24. The texture result was obtained by setting the pixel distance to 1 and the mean or average for each angle of the GLCM neighbourhood, namely 0°, 45°, 90°, and 135°. Every extracted feature's name is shown in Table 2.

2. GLRLM Features

The GLRLM Texture Features in this study amounted to 16. The texture results were obtained with the mean or average for each angle of the GLRLM neighborhood, namely 0°, 45°, 90°, and 135°. Every extracted feature's name is shown in Table 3.

3. GLSZM Features

The GLSZM Texture Features in this study amounted to 15. The texture result is obtained on the condition that a connected voxel has a distance of 1. Every extracted feature's name is shown in Table 4.

4. Haralick Features

The features of Haralick texture in this study are 13. The acquisition of features was done using the Mahotas library. Every extracted feature's name is shown in Table 5.

B. Feature Selection

Feature selection is carried out to determine features with good weight in distinguishing the classification label of an image and its features. The feature selection used in this study is Relief-F and only use the training set as shown in Fig. 2 to prevent data leakage. Relief-F is implemented with 10 number of neighbors and works by calculating the weight of each feature, the features is then sorted from the most weighted to the least weighted features. In Table 6, the most weighted features are shown in first and second rankings, namely the gray texture feature of the GLRLM and GLSZM level variance with a weight of 0.1567 and the less weighted texture feature is MCC GLCM ranking the last with a value of 0.0253.

Fig. 6 shows the plot for an experiment that tested the use of the number of features by the method used. The plot process was carried out using the 5-Fold Cross Validation technique on the training set. This stage is carried out to produce a better evaluation of limited data. In Fig. 6, the F1-Score results began to stagnate or even decrease starting from 9 and 10 number of features. Thus, the 10 most highly weighted features used in Relief-F in this study. Thereafter, the training set undergoes random oversampling before being used to train machine learning models. Features used are shown in the table rows given

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

color in Table 6. The distribution of selected features can be seen in Table 7 showing the most widely used haralick features of 4 followed by GLSZM, GLRLM, and GLCM with the fewest features, namely 1.

Table 6. Relief-F Ranked Features

No.	Feature Name	Weight
1	glrlm_GrayLevelVariance	0.1567
2	glszm_GrayLevelVariance	0.1567
3	glszm_GrayLevelNonUniformityNormalized	0.1287
4	glrlm_GrayLevelNonUniformityNormalized	0.1287
5	H_InfoMeasureOfCorrelation1	0.1236
6	glcm_ClusterProminence	0.1220
7	H_InverseDifferenceMoment	0.1164
8	H_DifferenceVariance	0.1118
9	H_AngularSecondMoment	0.1101
10	glszm_LowGrayLevelZoneEmphasis	0.1096
11	glrlm_LowGrayLevelRunEmphasis	0.1096
12	glrlm_RunVariance	0.0998
13	glrlm_LongRunHighGrayLevelEmphasis	0.0981
14	glszm_LargeAreaHighGrayLevelEmphasis	0.0981
15	glrlm_LongRunEmphasis	0.0976
16	glszm_LargeAreaEmphasis	0.0976
17	glszm_LargeAreaLowGrayLevelEmphasis	0.0908
18	glrlm_LongRunLowGrayLevelEmphasis	0.0908
19	glcm_MaximumProbability	0.0903
20	H_DifferenceEntropy	0.0900
21	glcm_InverseVariance	0.0896
22	glcm_Imc1	0.0858
23	glrlm_RunEntropy	0.0839
24	glszm_ZoneEntropy	0.0839
25	H_Entropy	0.0799
26	glrlm_GrayLevelNonUniformity	0.0795
27	glszm_GrayLevelNonUniformity	0.0795
28	glcm_ClusterShade	0.0794
29	H_SumEntropy	0.0735
30	glrlm_RunPercentage	0.0731
31	glszm_ZonePercentage	0.0731
32	glcm_ClusterTendency	0.0727
33	glcm_JointEnergy	0.0721
34	glcm_SumSquares	0.0711
35	glcm_SumAverage	0.0706
36	glcm_JointAverage	0.0706
37	H_SumVariance	0.0663
38	glcm_DifferenceEntropy	0.0661
39	H_SumAverage	0.0651
40	H_SumOfSquares	0.0645
41	glcm_Id	0.0643
42	glcm_SumEntropy	0.0633
43	glrlm_ShortRunHighGrayLevelEmphasis	0.0609
44	glszm_SmallAreaHighGrayLevelEmphasis	0.0609
45	glcm_Idm	0.0602
46	glszm_SmallAreaLowGrayLevelEmphasis	0.0601
47	glrlm_ShortRunLowGrayLevelEmphasis	0.0601
48	glcm_Autocorrelation	0.0593
49	original_glcm_JointEntropy	0.0548
50	glszm_SmallAreaEmphasis	0.0544
51	glrlm_ShortRunEmphasis	0.0544
52	glrlm_HighGrayLevelRunEmphasis	0.0535
53	glszm_HighGrayLevelZoneEmphasis	0.0535
54	original_glcm_Idn	0.0518
55	glszm_SizeZoneNonUniformityNormalized	0.0494
56	glrlm_RunLengthNonUniformityNormalized	0.0494
57	glcm_DifferenceAverage	0.0470
58	glcm_Imc2	0.0461
59	H_InfoMeasureOfCorrelation2	0.0461
60	glszm_SizeZoneNonUniformity	0.0446
61	glrlm_RunLengthNonUniformity	0.0446
62	H_Correlation	0.0364
63	H_Contrast	0.0344
64	glcm_DifferenceVariance	0.0336
65	glcm_Idmn	0.0331
66	glcm_Correlation	0.0331
67	glcm_Contrast	0.0309
68	glcm_MCC	0.0253

Table 7. Selected Features Distribution

No.	Feature Name	Count
1	GLCM	1
2	GLRLM	2
3	GLSZM	3
4	Haralick	4

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

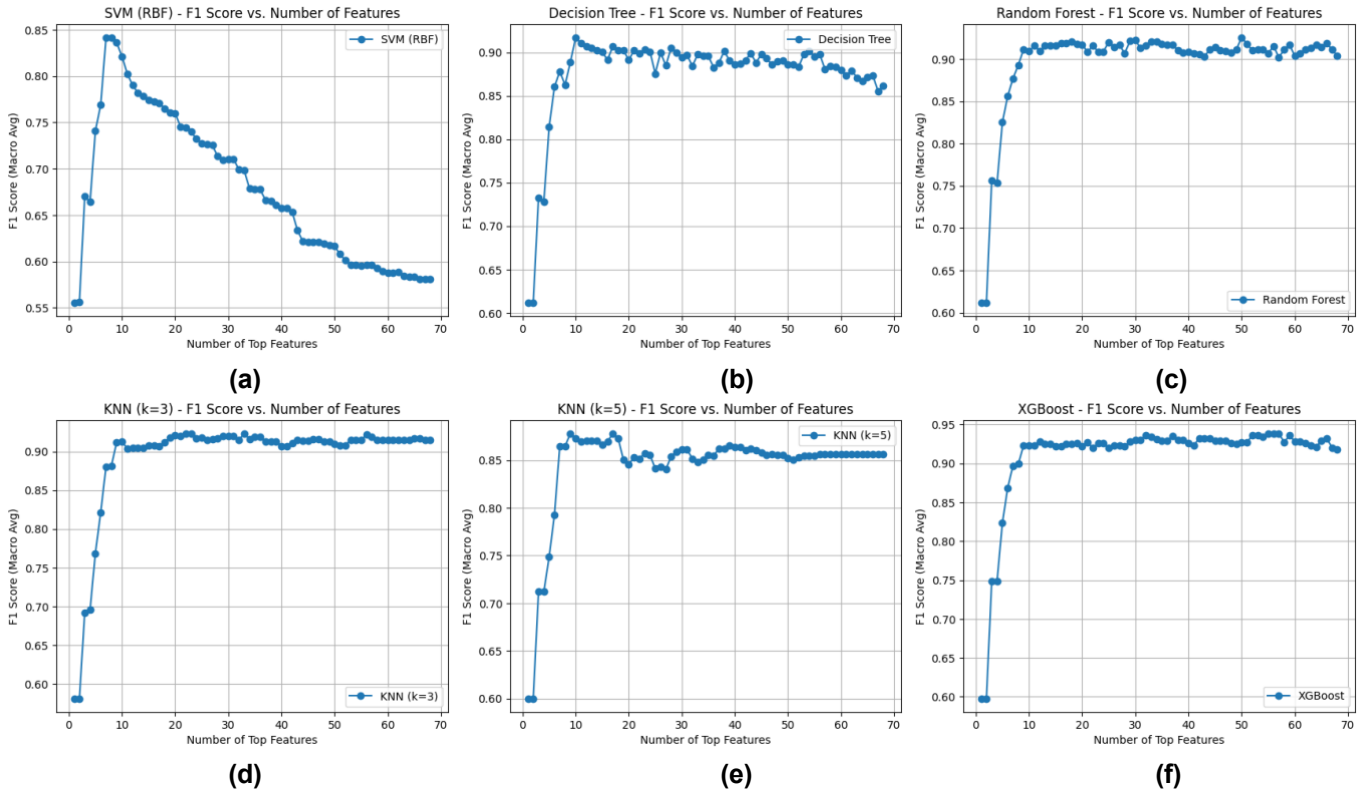


Fig. 6. Evaluation of Selected Feature in Training Set. (a) SVM RBF, (b) Decision Tree, (c) Random Forest, (d) KNN=3, (e) KNN=5, (f) XGBoost

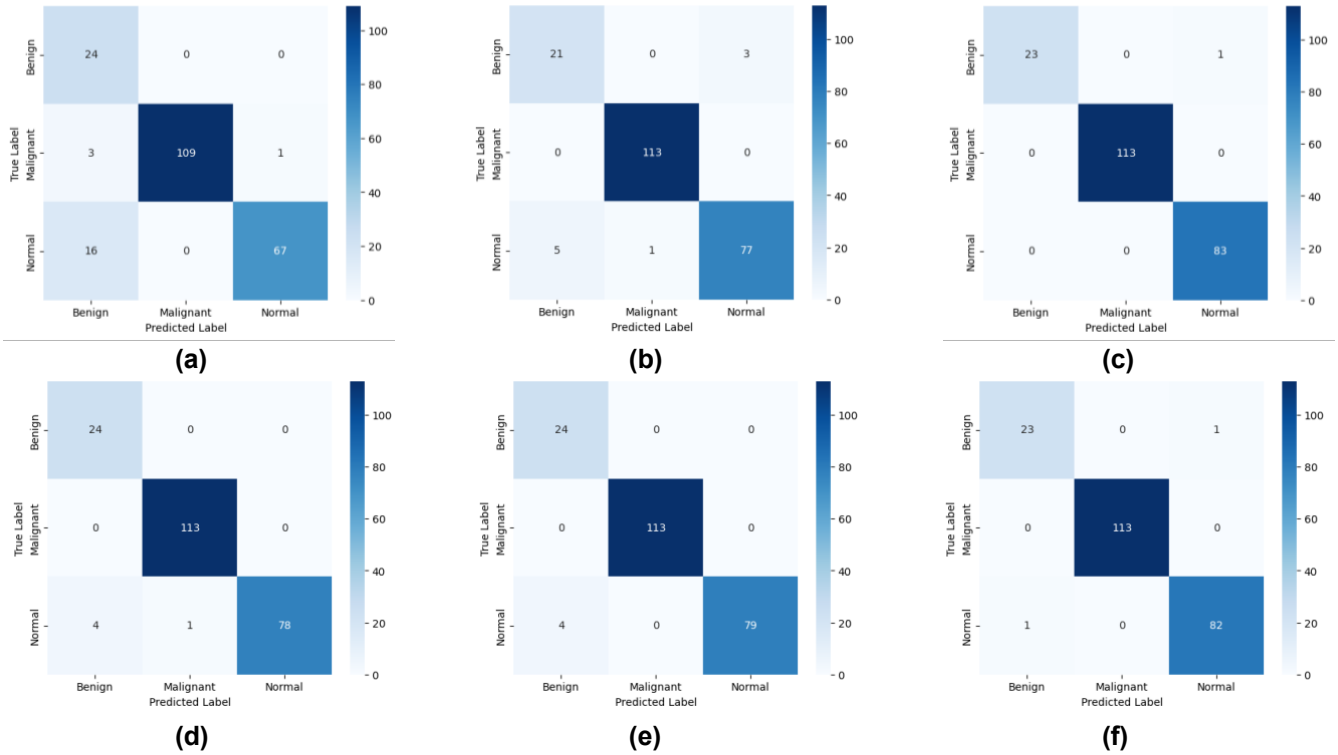


Fig. 7. Classification Confusion Matrix. (a) SVM RBF, (b) Decision Tree, (c) Random Forest, (d) KNN=3, (e) KNN=5, (f) XGBoost

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

Table 9. Classes Result Comparison

Method	F1-Score			Precision			Recall		
	Benign	Malignant	Normal	Benign	Malignant	Normal	Benign	Malignant	Normal
SVM RBF	0.72	0.98	0.89	0.56	1.00	0.99	1.00	0.96	0.81
DT	0.84	1.00	0.94	0.81	0.99	0.96	0.88	1.00	0.93
RF	0.98	1.00	0.99	1.00	1.00	0.99	0.96	1.00	1.00
KNN = 3	0.92	1.00	0.97	0.86	0.99	1.00	1.00	1.00	0.94
KNN = 5	0.92	1.00	0.98	0.86	1.00	1.00	1.00	1.00	0.95
XGBOOST	0.96	1.00	0.99	0.96	1.00	0.99	0.96	1.00	0.99

C. Classification

The classification in this research uses several machine learning methods, including SVM RBF, RF, DT, KNN=5, KNN=3, and XGBoost. Evaluation was carried out on a test set with 113 malignant images, 83 normal images, and 24 benign images. The evaluation results are shown in Table 8 with the bolded number being the highest value.

Table 8. Relief-F Ranked Features

Method	Accuracy	Precision	Recall	F1-Score
SVM RBF	0.9091	0.9462	0.9091	0.9173
DT	0.9591	0.9604	0.9591	0.9595
RF	**0.9955**	**0.9955**	**0.9955**	**0.9954**
KNN = 3	0.9773	0.9799	0.9773	0.9776
KNN = 5	0.9818	0.9844	0.9818	0.9823
XGBOOST	0.9909	0.9909	0.9909	0.9909

For each machine learning method used, ensemble methods such as RF and XGBOOST show excellent value. Both methods have evaluation results above 99%, RF produces the best evaluation value with the highest value of the other methods, this indicates that the RF method is almost perfect in distinguishing each image class. Clinically, enabling early detection with low false positives through high precision and detecting every cancer occurrence through high recall could improve prognosis by having a better treatment plan. Correspondingly, misdiagnoses of the cases may interfere with the treatment process. Thus, minimizing misclassification is important before implementing the method in real medical applications. In addition to the evaluation table, the analysis of the results was also carried out on the class of image shown in Table 9 and the confusion matrix shown in Fig. 7.

The confusion matrix shows the classification results with the original label of the test set. The diagonal column, from the top left to the bottom right, shows the number of images classified correctly according to the original label. Each method can perfectly classify for the malignant class, except for the SVM RBF method with an FN of 4 images. Meanwhile, for the KNN=3 and KNN=5 methods, the method can perfectly classify the images for benign and malignant classes, with the FN on the normal class images classified into benign classes. For the DT method, Fig. 7(b) shows a classification error that is almost similar, showing 3 FP and 6 FN for normal imagery.

Fig. 7(c) and Fig. 7(f) show excellent results. The RF method has only 1 FN for the benign class, whereas, the XGBoost method only has 1 FN and 1 FP for the benign class. Some of the methods shown in Fig. 7 often show FP for benign classes, this is because there is much less training data for benign class images compared to other classes, thus, there is less information related to benign imagery as well. Table 9 also supports this observation, showing that the benign class records the lowest F1-score and precision among all classes due to the high number of false positives.

IV. DISCUSSION

Other than other filter algorithm approaches, Relief-F is used in this research as it has shown promising results in [37]. The study specifically mentioned its performance in accuracy, prediction score, stability, and reliability due to its filter approach with good generalization power rather than other selection algorithm such as mRMR that yields slower execution time. Generalization power is important as it represents how well a selected feature subset performs across different classifiers, ensuring optimal accuracy [38].

The result of this study showed that lowest F1-Score and precision for image classification is from benign image class, this indicates the weakness of this research due to using ROS method in handling data imbalances is still not optimal. One of the things that can be done to address this is to use datasets with an even distribution and more for each class so that the model can classify the images with better information. In addition, the use of other sampling methods with better result can also be done, as balanced performance is really important for clinical diagnosis.

Table 10 shows the evaluation of the use of feature sets with RF classifier. The table shows the fewest feature sets, namely Haralick with 13 features, still has more features with a lower evaluation metric value than the proposed method. In addition, the HLK+GLCM set that had the same evaluation value used much more features than the proposed method. This shows that the approach of the proposed method that combines all features and feature selection results in the optimal approach than the method without feature selection.

A comparison of the performance of the image feature classification with the other same IQ-OTH/NCCD dataset is shown in Table 11, and the bolded numbers show the highest values. A comparison can be seen from the evaluation metrics and the number of features used,

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

Table 10. Feature Set Comparison

No	Feature Set	Number of Features	Accuracy	Precision	Recall	F1-Score
1	HLK	**13**	99.09	99.10	99.09	99.09
2	GLCM	24	98.18	98.27	98.18	98.16
3	GLRLM	16	93.18	93.17	93.18	93.12
4	GLSZM	15	93.64	93.58	93.64	93.50
5	HLK + GLCM	37	**99.55**	**99.55**	**99.55**	**99.54**
6	HLK + GLRLM	29	98.64	98.67	98.64	98.65
7	HLK + GLSZM	28	97.73	97.77	97.73	97.74
8	GLCM + GLRLM	40	97.27	97.29	97.27	97.18
9	GLCM + GLSZM	39	97.27	97.31	97.27	97.23
10	GLRLM + GLSZM	31	95.00	94.98	95.00	94.95
11	HLK + GLCM + GLRLM	53	98.18	98.18	98.18	98.18
12	HLK + GLCM + GLSZM	52	99.09	99.11	99.09	99.08
13	HLK + GLRLM + GLSZM	44	98.18	98.18	98.18	98.18
14	GLCM + GLRLM + GLSZM	55	97.73	97.77	97.73	97.73
15	HLK + GLCM + GLRLM + GLSZM	68	98.64	98.63	98.64	98.63

referring to Table 11, the proposed method can produce excellent results, with the value of all metrics reaching 99% with only 10 features.

The method proposed has different approach with research [7] and [8] who uses deep learning as their method in feature extraction and classification. Although it used more complex approach, the proposed method can show better performance, where the method proposed in [7] although it produces perfect precision, the accuracy value achieved is 93.21%, the recall is 90.71% and the F1-Score is 92.72%. The lower recall value of this precision is due to the large number of classification image results that are FN. Meanwhile, for the method in [8] the VGG-19 deep learning model that has been fine-tuned is used to extract features combined with Principal Component Analysis (PCA) and ANFIS to reduce the dimension of features still produces a lower evaluation value with the method proposed in this research with the values for accuracy, precision, recall, and F1-Score respectively being 92.73%, 91.22%, 91.96%, and 91.59%. This lower result may be caused by inefficient or naïve deep learning architecture that still needs more tuning to be able to capture the needed features. In this case, the proposed method outperforms by only using the heavily weighted features through feature selection.

The last deep learning method compared is the method by [9], which uses an autoencoder for feature extraction and SVM as the classification method. Although it showed a better evaluation score than the previous two deep learning methods, the still demonstrated a lower score, with accuracy, precision, recall, and F1-Score measuring 97%, 96%, 97%, and 96%, respectively. This approach of using deep learning also has shortcomings, primarily due to its complexity, making it resource-intensive.

This study also compares the methods proposed with a simpler approach by [14], which uses the GLCM and Gabor Filter features for its classification process with SVM Linear, RBF, and Polynomial. The amount of features used in this method cannot be specified specifically because the number of Gabor Filter features is not explicitly mentioned. However, the performance of this method can still be evaluated, where the combination of GLCM and Gabor Filter results in an accuracy value of 89.89, which is much less than the proposed method, where the accuracy obtained by proposed method reaches 99%. Similar conclusions were also reached for other evaluation metrics in the study, such as 98.55% precision, 97.14% recall, and 97.84% F1-score.

The feature extraction approach was also carried out in [15], who used SVM with 594 features consisting of 20 GLCM features and 574 LBP features. These features are more than the proposed method, showing better application method efficiency by using only 10 features. In addition, using GLCM and LBP features still yielded lower evaluation results, with accuracy, precision, recall, and F1-score recorded at 93%, 94%, 93%, and 92.5%, respectively. Both study [14] and [15] uses GLCM like the proposed method. Similarly, both of these approaches also achieve above 90% in most of the performances, showing the effectiveness of the selected texture features in the model.

For the feature extraction approach, the method in [16] generated the second-best evaluation value. The study

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

used the Euclidean Distance Classifier and the BCR feature with as many as 64 features. The number of features used in the study is higher than that of the proposed method, indicating the study model's lower efficiency in feature usage. With evaluation values of 98.32% accuracy, 98.1% precision, 96.18% recall, and 97.05% F1-score, the evaluation score of the method is still lower than that of the proposed method. Overall, the proposed approach can achieve the highest performance compared to previous studies due to its improvement by only using the relevant features. The proposed method shows a promising opportunity for future clinical applications with more intensive research. This can be done by delving into the specific region of interest and extracting the texture features.

These reduction number of features from the study [15] and [16] are purely statistical and did not increase nor decrease the clinical interpretability. Instead, it shows how these statistical features' numerical value portrays the characteristics of each lung tumor malignancy. These values give further understanding in terms of statistical interpretation of CT Scans.

Feature extraction is done through Kaggle's default notebook environment. Next, from feature selection to testing, Google Colab Notebook is used under Google Compute Engine with CPU as its hardware accelerator. Table 12 shows the computational burden through execution time and peak memory usage using Python's time library and tracemalloc. Every process only uses a small amount of time and peak memory usage, showing its promising excellence in resource-limited settings.

Table 12. Computational Cost

Step	Execution Time (seconds)	Peak Memory Usage (MB)
Feature Extraction	1142.29	52.90
Feature Selection	211.26	11.07
Classifier Training (average)	0.53	0.27

Despite its excellent result in Table 11, before integrating this study into real-world clinical workflows, this study has its limitations as it relies solely on The IQ-OTH/NCCD dataset, which may overfit, create bias, and reduce its generalization capability in classifying real-world cases. Thus, in future research, testing with independent datasets or real clinical data is needed to validate the proposed method, along with collaboration with healthcare providers to ensure the robustness and its possible application.

V. CONCLUSION

This study uses the combination of 1 GLCM feature, Cluster Prominence, 2 features of GLRLM, Gray Level Variance and Gray Level Non Uniformity Normalized, 3 features of GLSZM, Gray Level Variance, Gray Level Non Uniformity Normalized, and Low Gray Level Zone Emphasis, lastly, 4 features of Haralick, Informational Measure Of Correlation 1, Inverse Difference Moment, Difference Variance, and Angular Second Moment. These features were obtained by the selection process using Relief-F to select the most weighted features in the classification process. The training and evaluation were carried out with the same dataset, namely the IQ-OTH/NCCD lung cancer dataset, showing improved performance and efficiency in using features. The research showed that the proposed method achieved an accuracy of 99.55%, precision of 99.55%, recall of 99.55%, and an F1-score of 99.54% using only 10 features. These results show that the hybrid feature method between GLCM, GLRLM, GLSZM, and Haralick, along with the Relief-F feature selection process and the random forest classification method, has relatively high effectiveness and efficiency. Future research using a more balance and diverse dataset are necessary to prove the method robustness, using a more advance sampling technique can also be implemented to handle imbalance data.

REFERENCES

[1] A. Patel, "Benign vs Malignant Tumors," *JAMA Oncol*, vol. 6, no. 9, pp. 1488–1488, Sep. 2020, doi: 10.1001/JAMAONCOL.2020.2592.

[2] J. Ferlay *et al.*, "Trachea Bronchus And Lung Fact Sheet," Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. Accessed: Feb. 16, 2025. [Online]. Available: https://gco.iarc.who.int/media/globocan/factsheets/cancers/15-trachea-bronchus-and-lung-fact-sheet.pdf

[3] J. Ferlay *et al.*, "Indonesia fact sheet," Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. Accessed: Feb. 16, 2025. [Online]. Available: https://gco.iarc.who.int/media/globocan/factsheets/populations/360-indonesia-fact-sheet.pdf

[4] C. Zhang *et al.*, "Enhancing lung cancer diagnosis with data fusion and mobile edge computing using DenseNet and CNN," *Journal of Cloud Computing*, vol. 13, no. 1, pp. 1–10, Dec. 2024, doi: 10.1186/S13677-024-00597-W/TABLES/2.

[5] O. D. Asmara *et al.*, "Lung Cancer in Indonesia," *Journal of Thoracic Oncology*, vol. 18, no. 9, pp. 1134–1145, Sep. 2023, doi: 10.1016/j.jtho.2023.06.010.

[6] M. Vettoretti, M. Gouveia, T. Mendes, E. M. Rodrigues, H. P. Oliveira, and T. Pereira, "Comparing 2D and 3D Feature Extraction Methods for Lung Adenocarcinoma Prediction Using CT Scans: A Cross-Cohort Study," *Applied Sciences 2025, Vol. 15, Page 1148*, vol. 15, no. 3, p. 1148, Jan. 2025, doi: 10.3390/APP15031148.

[7] T. I. A. Mohamed, O. N. Oyelade, and A. E. Ezugwu, "Automatic detection and classification of lung cancer CT scans based on deep learning and ebola optimization search algorithm," *PLoS One*, vol. 18,

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

no. 8, p. e0285796, Aug. 2023, doi: 10.1371/JOURNAL.PONE.0285796.

[8] K. Nithish Kumar, V. C. Sai Santhosh, A. Mane, K. Rahimunnisa, G. Nirmal, and M. Vaghulade, "Enhancing Lung Cancer Detection and Localization with a Hybrid VGG-19 and Adaptive Neuro-Fuzzy Inference System (ANFIS) Approach on Imaging Data," *2024 Systems and Information Engineering Design Symposium, SIEDS 2024*, pp. 280–285, 2024, doi: 10.1109/SIEDS61124.2024.10534680.

[9] A. Lakshmanarao, N. Gopal, N. Vullam, M. Sridhar, M. K. Kanth, and U. M. Rayudu, "Lung cancer detection using hybrid integration of autoencoder feature extraction and ML techniques," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 37, no. 1, pp. 416–424, Jan. 2025, doi: 10.11591/ijeecs.v37.i1.pp416-424.

[10] H. T. Gayap and M. A. Akhloufi, "Deep Machine Learning for Medical Diagnosis, Application to Lung Cancer Detection: A Review," *BioMedInformatics 2024, Vol. 4, Pages 236-284*, vol. 4, no. 1, pp. 236–284, Jan. 2024, doi: 10.3390/BIOMEDINFORMATICS4010015.

[11] H. Wei *et al.*, "A texture feature extraction method considering spatial continuity and gray diversity," *International Journal of Applied Earth Observation and Geoinformation*, vol. 130, p. 103896, Jun. 2024, doi: 10.1016/J.JAG.2024.103896.

[12] S. G. A. Usha and S. Vasuki, "Significance of texture features in the segmentation of remotely sensed images," *Optik (Stuttg)*, vol. 249, p. 168241, Jan. 2022, doi: 10.1016/J.IJLEO.2021.168241.

[13] Z. Sadeghian, E. Akbari, H. Nematzadeh, and H. Motameni, "A review of feature selection methods based on meta-heuristic algorithms," *Journal of Experimental & Theoretical Artificial Intelligence*, Jan. 2025, doi: 10.1080/0952813X.2023.2183267.

[14] H. F. Kareem, M. S. AL-Husieny, F. Y. Mohsen, E. A. Khalil, and Z. S. Hassan, "Evaluation of SVM performance in the detection of lung cancer in marked CT scan dataset," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 3, pp. 1731–1738, Mar. 2021, doi: 10.11591/ijeecs.v21.i3.pp1731-1738.

[15] K. C. Ahnaf and E. S. Wahvuni, "Comparative Analysis of Image Processing Methods using GLCM and LBP Feature Extraction for Lung Cancer Detection," *6th International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2023 - Proceeding*, pp. 513–518, 2023, doi: 10.1109/ISRITI60336.2023.10467244.

[16] S. Saechueng and U. Suttapakti, "Binary Count Ratio for Lung Cancer Classification in Computerized Tomography Scan Images," *6th International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2024*, pp. 70–74, 2024, doi: 10.1109/ICAIIC60209.2024.10463210.

[17] S. O. Shim, M. H. Alkinani, L. Hussain, and W. Aziz, "Feature Ranking Importance from Multimodal Radiomic Texture Features using Machine Learning Paradigm: A Biomarker to Predict the Lung Cancer," *Big Data Research*, vol. 29, p. 100331, Aug. 2022, doi: 10.1016/J.BDR.2022.100331.

[18] hamdalla alyasriy and M. AL-Huseiny, "The IQ-OTH/NCCD lung cancer dataset," vol. 4, 2023, doi: 10.17632/BHMDR45BH2.4.

[19] C. L. Chowdhary and D. P. Acharjya, "Segmentation and Feature Extraction in Medical Imaging: A Systematic Review," *Procedia Comput Sci*, vol. 167, pp. 26–36, Jan. 2020, doi: 10.1016/J.PROCS.2020.03.179.

[20] B. Dolly and D. Raj, "Texture Based Image Retrieval Using GLCM and LBP," *Advances in Intelligent Systems and Computing*, vol. 1312 AISC, pp. 35–45, 2021, doi: 10.1007/978-981-33-6176-8_5.

[21] S. Dash and M. R. Senapati, "Gray level run length matrix based on various illumination normalization techniques for texture classification," *Evol Intell*, vol. 14, no. 2, pp. 217–226, Jun. 2021, doi: 10.1007/S12065-018-0164-2/FIGURES/13.

[22] M. E. Mayerhoefer *et al.*, "Introduction to Radiomics," *J Nucl Med*, vol. 61, no. 4, pp. 488–495, Apr. 2020, doi: 10.2967/JNUMED.118.222893.

[23] M. N. Winnarto, M. Mailasari, and A. Purnamawati, "Ekstraksi Fitur Dan Implementasi Machine Learning Untuk Klasifikasi Jenis Tumor Otak," *Jurnal Infortech*, vol. 6, no. 1, pp. 65–70, Jun. 2024, doi: 10.31294/INFORTECH.V6I1.21987.

[24] A. Rahman *et al.*, "Interpretable wood chip moisture content prediction through texture analysis," *Expert Syst Appl*, vol. 275, p. 126989, May 2025, doi: 10.1016/J.ESWA.2025.126989.

[25] L. Guomin, Hujing, and H. Guichuan, "Research on the Key Feature Selection of Supercharger Based on Relief-F Arithmetic," *IOP Conf Ser Earth Environ Sci*, vol. 826, no. 1, p. 012033, Jul. 2021, doi: 10.1088/1755-1315/826/1/012033.

[26] K. Li *et al.*, "Lung adenocarcinoma identification based on hybrid feature selections and attentional convolutional neural networks," *Mathematical Biosciences and Engineering 2024 2:2991*, vol. 21, no. 2, pp. 2991–3015, 2024, doi: 10.3934/MBE.2024133.

[27] T. Pasangthien and B. Yimwadsana, "Rebalancing Clinical Data with Probabilistic Random Oversampling," *Journal of the Thai Medical Informatics Association*, vol. 8, no. 2, pp. 68–72, Oct. 2022, Accessed: May 01, 2025. [Online]. Available: https://he03.tci-thaijo.org/index.php/jtmi/article/view/480

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101

[28] M. Simsek and A. S. Das, "The Effect of Handling Imbalanced Datasets Methods on Prediction of Entrepreneurial Competency in University Students," *Turkish Journal of Forecasting*, vol. 06, no. 2, pp. 53–60, 2022, doi: 10.34110/forecasting.1185545.

[29] Y. A. Sir and A. H. H. Soepranoto, "Pendekatan Resampling Data Untuk Menangani Masalah Ketidakseimbangan Kelas," *Jurnal Komputer dan Informatika*, vol. 10, no. 1, pp. 31–38, Mar. 2022, doi: 10.35508/jicon.v10i1.6554.

[30] R. E. Nogales and M. E. Benalcázar, "Analysis and Evaluation of Feature Selection and Feature Extraction Methods," *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, pp. 1–13, Dec. 2023, doi: 10.1007/S44196-023-00319-1/TABLES/6.

[31] J. Wainer and P. Fonseca, "How to tune the RBF SVM hyperparameters? An empirical evaluation of 18 search algorithms," *Artif Intell Rev*, vol. 54, no. 6, pp. 4771–4797, Aug. 2021, doi: 10.1007/S10462-021-10011-5/METRICS.

[32] S. Misra and H. Li, "Noninvasive fracture characterization based on the classification of sonic wave travel times," *Machine Learning for Subsurface Characterization*, pp. 243–287, Jan. 2020, doi: 10.1016/B978-0-12-817736-5.00009-0.

[33] R. K. Halder, M. N. Uddin, M. A. Uddin, S. Aryal, and A. Khraisat, "Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications," *Journal of Big Data 2024 11:1*, vol. 11, no. 1, pp. 1–55, Aug. 2024, doi: 10.1186/S40537-024-00973-Y.

[34] M. Kumar, P. Samui, D. R. Kumar, and P. G. Asteris, "State-of-the-art XGBoost, RF and DNN based soft-computing models for PGPN piles," *Geomechanics and Geoengineering*, Nov. 2024, doi: 10.1080/17486025.2024.2337702.

[35] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data 2021 8:1*, vol. 8, no. 1, pp. 1–74, Mar. 2021, doi: 10.1186/S40537-021-00444-8.

[36] L. Alzubaidi *et al.*, "A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications," *J Big Data*, vol. 10, no. 1, p. 46, Apr. 2023, doi: 10.1186/s40537-023-00727-2.

[37] M. C. Barbieri, B. I. Grisci, and M. Dorn, "Analysis and comparison of feature selection methods towards performance and stability," *Expert Syst Appl*, vol. 249, p. 123667, Sep. 2024, doi: 10.1016/J.ESWA.2024.123667.

[38] S. Salesi and G. Cosma, "Generalisation Power Analysis for finding a stable set of features using evolutionary algorithms for feature selection," *Knowl Based Syst*, vol. 231, p. 107450, Nov. 2021, doi: 10.1016/J.KNOSYS.2021.107450.

AUTHOR BIOGRAPHY

Rizki Dwi Rahmawan is an informatics undergraduate student from Universitas Sebelas Maret rolled in 2021. His interest in biomedical field is shown by his final project on lung tumor classification, exploring texture features and lightweight machine learning methods as alternative to a heavy and resource-intensive deep learning. Creating a more accessible method for places with limited resources. He received Indonesian Deposit Insurance Corporation (IDIC) scholarship for 1 year in 2023.

Umi Salamah is a dedicated lecturer at the Department of Informatics within the Faculty of Information Technology and Data Science at Universitas Sebelas Maret (UNS) in Surakarta, Indonesia. Umi Salamah received her Master's and Doctoral degrees from Institut Teknologi Sepuluh Nopember Surabaya, Indonesia in 2002 and 2018 respectively. Her research interests include image processing and pattern recognition in medical images, especially computer-aided diagnosis of malaria.

Ery Permana Yudha is a lecturer at Universitas Sebelas Maret focusing on image processing, machine learning, and information retrieval. He earned his Bachelor's and Master's degrees from Institut Teknologi Sepuluh Nopember. His research focuses on feature matching and information retrieval. Committed, his dissertation contributes to increasing the effectiveness of feature matching through hypergraph representation. He has successfully published more than 10 articles in national/international reputed journals/conferences.

Corresponding author: Rizki Dwi Rahmawan, rizkidwirahmawan@student.uns.ac.id, Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Surakarta, Indonesia.

DOI: https://doi.org/10.35882/ijeeemi.v7i2.101