Journal of Computer Science

Original Research Paper

Lung Cancer Detection Using Regularized Extreme Learning Machine and PCA Features

[1]K. Kavitha, [2]Dr. V. Saravana Kumar, [3]V. Bhoopathy, [4]Dhanalakshmi S., [5]Syed Arfath Ahmed and [6]N. Valarmathi

[1]Department of CSE-AI&ML, GMR Institute of Technology, Rajam, India
[2]Department of AI &DS, Rajalakshmi Engineering College, Chennai, India
[3]Department of CSE, Sree Rama Engineering College, Tirupati, Andhra Pradesh, India
[4]Department of III/ECE, Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, India
[5]Department of CSE, Maulana Azad National Urdu University, Polytechnic, Gachibowli, Hyderabad, Telangana, India
[6]Department of Information Technology, M.Kumarasamy College of Engineering, Karur, Tamilnadu, India

Article history
Received: 11-12-2023
Revised: 27-04-2024
Accepted: 22-05-2024

Corresponding Author:
V. Bhoopathy
Department of CSE, Sree Rama Engineering College, Tirupati, Andhra Pradesh
Email: v.bhoopathy@gmail.com

Abstract: When finding abnormalities in target images is time-sensitive, as it is with many cancer tumors, image processing techniques have recently found widespread use across various medical industries to improve images for early detection and treatment stages. In the setting of cancer, where time is of the essence in detecting anomalies within medical imaging, our research takes on further urgency. In the medical field, image processing techniques have taken center stage, with the goal of improving image quality for the purpose of early identification and treatment planning. Our suggested approach incorporates multiple stages of image processing, such as feature extraction, morphological techniques, segmentation, and histogram equalization, with a focus on CT scan pictures. Finding better ways to interpret images for early detection in medical imaging is the driving force behind this research. Feature extraction, morphological algorithms, segmentation, and histogram equalization are some of the image-processing methods used in the study. In order to make the estimation process faster and more accurate, we also use Principal Component Analysis (PCA) and a Regularized Extreme Learning Machine (RELM). The suggested model performs admirably, with an accuracy of around 99.7%. When put to the test against popular models such as CNN, SVM, SVM-RBF, and RELM, the proposed method clearly comes out on top. This study's findings provide a more effective and efficient way for abnormality detection in medical imaging, which is a major advancement in the area. Early diagnosis and treatment planning in medical situations can be directly influenced by the integration of PCA and RELM, which shows promise for enhancing the speed and precision of estimation procedures.

Keywords: Lung Cancer Detection, Feature Extraction, Segmentation, Reinforcement Learning Machine (RLM), Convolutional Neural Network (CNN)

Introduction

When abnormal cells in the lungs proliferate and expand, this is known as lung cancer. Using image processing methods for the diagnosis of lung cancer lymph fluid, which surrounds lung tissue and blood both have the potential to transport cancer cells away from the lungs. The lymphatic system consists of lymphatic veins that carry lymph to lymph nodes in the chest cavity and lungs. As lymph moves out of the lungs, it naturally moves toward the center of the chest, which is where lung cancer tends to spread. When a malignant cell spreads from its original location to a lymph node or another organ via the bloodstream, this is called metastasis. When cancer starts in the lungs, it is called primary lung cancer. It is possible to classify many subtypes of lung cancer into two main groups: Squamous cell carcinoma, adenocarcinoma and carcinoma are subtypes of non-small cell lung cancer, one of the two major forms of lung cancer. Small cell lung cancer is the other. Among males and females alike, lung cancer has the second highest incidence rate. Out of all the newly diagnosed cancers, lung cancer makes up about 13.55%. Lung cancer is anticipated to cause 228,160 new cases in 2019 (116,444 in men and 111,720 in women), with an estimated 142,680 fatalities (76,620 in men and 66,120 in women) according to the American cancer society. The prognosis for a cancer patient following a diagnosis is very context-dependent. Early detection has the potential to reduce the increasing death toll from lung cancer. Lung nodules or tumors that have metastasized can now be evaluated using Low Dose-Computed Tomography (LDCT) or a chest CT scan, which allows clinicians to determine the size, location, and shape of the tumor (Narayanan and Jeeva, 2015; Teramoto *et al*., 2016). CT scan images are chosen because of their high resolution, great contrast in soft tissues, adaptability, speed, lack of invasiveness, ability to slice images into many views, broad use in clinical practice, and compatibility with image processing methods. All things considered; CT scans are the way to go for our early detection research that relies on image processing techniques. The term "cancer" covers a wide range of cancers. As a medical term, it is a cancer that develops when biological alterations lead to unchecked cell proliferation and division. Most cells in the body have certain purposes and die off at specific intervals. Apoptosis, on the other hand, is a natural process that involves cell death (Anushkannan *et al*., 2023; Sarkar *et al*., 2023). A cell obeys instructions to commit suicide so that it can be replaced with a new one that is more efficient. The systems that normally instruct cells to stop dividing and die are dysfunctional in cancerous tissues. As a result, they proliferate inside the body, consuming the oxygen and nutrients meant for healthy cells (Subramanian *et al*., 2023). Tumors, immune system dysfunction, and other abnormalities can all be caused by cancerous cells (Hsieh, 2001; Bunn Jr and Kelly, 1995). A malignant tumor of the lungs, lung cancer is characterized by the unchecked proliferation of cells. The leading cause of death from cancer is lung cancer. According to the stage at which the cancer cells in the lungs are discovered, lung cancer is the deadliest and most common cancer worldwide (Gopinath *et al*., 2023a-b). Hence, early detection of the cancer is crucial in avoiding the more severe and spread later stages of the cancer. A broad lung cancer background allows a smooth transition to the research's main focus processing tools for early diagnosis. Due to this strategic alignment, all of the data provided will help understand and appreciate the planned study objectives. In this proposed method we have preprocessed using Gabor filter. For the segmentation process, this study used the Otsu thresholding algorithm.

In this research to improve medical picture analysis, the Gabor filter for pre-processing and the Otsu algorithm for segmentation were chosen for their advantages and compatibility. The gabor filter's ability to capture texture and small details makes it suitable for medical imaging, which needs identifying microscopic anomalies. Convolutionally processing the input image using gabor filter kernels at various scales and orientations reveals texture patterns connected to anomalies. Otsu is a popular segmentation algorithm for images with different brightnesses. Otsu segmentation divides the better image after preprocessing. By enhancing the early detection and atreatment planning stages through modern image processing techniques, this research shows great promise for the treatment of lung cancer. Our study is of utmost importance in the setting of cancer, when the prompt identification of abnormalities in medical imaging is vital. Our research seeks to improve the interpretation of medical pictures for early detection by integrating many phases of image processing, including feature extraction, morphological approaches, segmentation, and histogram equalization, which are specifically designed for CT scan images. When it comes to lung cancer, these methods are crucial for finding anomalies early on, when the chances of a successful therapy are higher. In addition, our suggested model expedites and improves the accuracy of estimation by incorporating Principal Component Analysis (PCA) and a Regularised Extreme Learning Machine (RELM). With an astounding accuracy rate of almost 99.7%, this combination of sophisticated algorithms proves to be exceptionally effective. Our suggested approach outperforms more traditional models like CNN, SVM, and SVM-RBF. A major step forward in the area, this study's results provide a more efficient and accurate method for detecting abnormalities in medical imaging. In the end, there's a lot of hope that combining PCA and RELM may help with early lung cancer detection and treatment planning by making it faster and more accurate. Better patient outcomes and a decrease in the burden of this terrible disease are the ultimate goals of our study, which strives to improve diagnosis and intervention times.

Materials

Over the course of three months in the autumn of 2019, the Iraq-Oncology Teaching Hospital/National Center for Cancer (IQ-OTH/NCCD) gathered data on lung cancer cases at the aforementioned specialized facilities in Iraq. Images from healthy individuals and computed tomography scans from patients with lung cancer at different stages are included. At these institutes, IQ-OTH/NCCD slides were annotated by radiologists and cancer specialists. There is a grand total of 1190 photos in the dataset, which depict CT scan slices from 110 separate cases. To characterize these cases, three broad categories have been defined: Normal, benign and malignant. Among them, forty have been found to be malignant, fifteen to be benign and the other fifty-five to be normal. Using medical image processing technology, we study CT scan anomalies. The suggested method uses contemporary image processing for feature extraction, morphology, segmentation, and histogram equalization. These methods improve image interpretation to detect medical imaging problems early. PCA and a RELM in the image processing pipeline solve anomaly detection's temporal sensitivity in our investigation. These strategies speed up and improve estimation, making it more precise and efficient.

Methods

Over the course of three months in the autumn of 2019, the Iraq-Oncology Teaching Hospital/National Center for Cancer (IQ-OTH/NCCD) gathered data on lung cancer cases at the aforementioned specialized facilities in Iraq. Images from healthy individuals and computed tomography scans from patients with lung cancer at different stages are included. At these institutes, IQ-OTH/NCCD slides were annotated by radiologists and cancer specialists. There is a grand total of 1,190 photos in the dataset, which depict CT scan slices from 110 separate cases (AL-Yasriy, 2017). To characterize these cases, three broad categories have been defined: Normal, benign, and malignant. Among them, forty have been found to be malignant, fifteen to be benign and the other fifty-five to be normal. Using medical image processing technology, we study CT scan anomalies. The suggested method uses contemporary image processing for feature extraction, morphology, segmentation, and histogram equalization. These methods improve image interpretation to detect medical imaging problems early. PCA and a RELM in the image processing pipeline solve anomaly detection's temporal sensitivity in our investigation. These strategies speed up and improve estimation, making it more precise and efficient.

Preprocessing

The system begins with the preprocessing phase. The input image in this case is a CT scan. Throughout this process, we will be converting the image to a grayscale format in order to filter out any unwanted noise. As part of the preprocessing phase, various filters, including some filters that can be used to eliminate undesired noise are Gabor, median, wiener, and erosion filters (Yamini *et al*., 2023). For the purpose of this research, the Gabor filter is used to eliminate background noise. It makes the picture look even better by increasing the contrast. Even though we know that medical photographs are often grayscale, we still convert the raw CT scan images utilized as input photos to that color in order to enhance the quality of the images.

Histogram Equalization

The input image is first preprocessed and then the histogram is equalized. Image processing technique that takes advantage of an image's histogram to modify contrast. The image's histogram is smoothed down and its global contrast is boosted, both of which contribute to a more even distribution of brightness. By smoothing down the peaks and troughs in the histogram of the chest CT scan's input image, we can see the individual regions of the lungs more clearly as shown in Fig. 1. It takes a grayscale image and brings out the essential details to create an improved rendition of the image.

Segmentation

The image is then transformed into a binary image, sometimes known as a black-and-white image. Since it is a digital image, each pixel in a binary image can be represented by two possible values-0 1. The presence of white data is represented by a 1 and the absence of black data by a 0. To make it easier to identify and study individual nodules in a digital image, segmentation divides the image into numerous sections. This strategy employs Otsu's thresholding method for segmentation. A CT scan's input image is segmented by removing pixels that are either below or above a certain level, or threshold value. A grayscale image is the threshold in this investigation. As shown in Fig. 2 the threshold binary image only has two levels, 0 and 1, making it easier for the radiologist to analyze the information needed to detect the nodules, in contrast to a grey level picture, which usually has 256 steps.

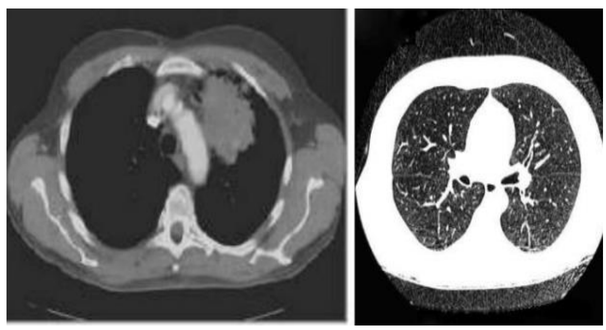

Fig. 1: Input and preprocessing by gabor

Fig. 2: Segmentation of image

Feature Extraction

In this study, we focus on the extraction and application of two feature sets, namely Gabor filter and GLCM (Mohanaiah and Sathyanarayana, 2013) features. Image processing automatically chooses the appropriate threshold for foreground and background segmentation using Otsu's thresholding. This method maximizes class variation, making it useful for distinguishing objects from the background in bimodal pixel-intensity histograms. The Gabor filter captures detailed texture patterns needed to identify lung cancer in CT scans, while Otsu's thresholding automatically divides regions of interest from the background. These approaches extract key aspects for detection and decision-making. The frequency spectrum as a whole, together with the amplitude and phase of each pair, are all captured by the Gabor filter. The image has been distorted using each Gabor filter group at all pixels for the purpose of Gabor feature extraction (Wang *et al*., 2005). This data set is commonly referred to as a Grey Level Co-occurrence Matrix (GLCM). Texture analysis, multi-resolution, orientation sensitivity, and feature extraction enhancement make the Gabor filter useful for CT scan lung cancer identification. The pre-processing pipeline's detailed image analysis may improve detection and classification algorithms. Many fields that make use of image analysis, including medicine, employ the two-stage process. In this feature extraction method, the initial stage (Rajadell *et al*., 2009) is to calculate the GLCM matrix and the second step is to determine the texture features using the GLCM matrix as a basis. According to GLCM, the frequency of each grey level at a pixel in a given geometric location in relation to all other pixels is described as a function of the grey level. Here are the traits that were discovered.

Contrast

To measure contrast in a picture, we look at how the surrounding pixels compare to the reference pixel in terms of brightness or grayscale:

$$\sum_{k}\sum_{i}(k-i)^{2}D(k,i) \tag{1}$$

k and i indicate pixel intensity. D (k, i) represents the normalized joint probability distribution of image intensities k and i. Pixel intensities that fluctuate greatly increase contrast, suggesting features or edges. A greater contrast indicates more texture or detail.

Homogeneity

Homogeneity is a measure of how evenly distributed the GLCM matrix elements are around the diagonal:

$$\sum_{k}\sum_{i}\frac{D(k,i)}{1+|k-i|} \tag{2}$$

Here, k and i stand for the intensity values of pixels. The joint probability distribution, D (k, i), has been normalized. A more uniform texture is suggested by an increase in homogeneity as pixel intensities converge. The image's regularity or smoothness is shown by it.

Energy

Energy is represented by the sum of the squared elements of the GLCM matrix. We got it by applying the ASM, or Second Moment of Angle. Here we may see an approximate representation of the regional consistency of the greyscale. The ASM value will be high for pixels that are very similar to one another and low for pixels that are very different:

$$\sum_{k}\sum_{i}\frac{D(k,i)}{1+|k-i|} \tag{3}$$

The joint probability distribution, D (k, i), has been normalized. When the intensities of pixels are concentrated in certain patterns, it creates a more organized and organized image, which in turn increases the energy. The homogeneity of the pixel values is measured. In order to reduce the quantity of raw image data, feature extraction (GLCM) evaluates specific values or characteristics that can be utilized for picture classification.

Classification of the Model

Efficient Feature Selection

When processing neuroimaging data, one could encounter the so-called "curse of dimensionality" due to the potentially excessively high number of features per subject in relation to the total number of persons. Our method for accomplishing this task is Principal Component Analysis (PCA), which is widely used to handle data that is high dimensional, such as images[14] to implement a feature selection process that yields desirable results. Medical picture classification might be difficult due to high-dimensional data. For effective classification, feature selection reduces dimensionality while maintaining critical information. High-dimensional datasets may have redundant or unhelpful features. Dimensionality reduction is essential for model generalization and computing efficiency. The principle components that capture the highest data variance are identified by PCA. PCA retains components with substantial variance to guarantee the smaller feature set has correct classification information. As a result, only the dimensions that adequately characterize the data are kept. New features are created by linearly combining the original features with each occurrence of the supplied dataset that occurs in an e-dimensional space and each instance is mapped to an N-dimensional subspace where $N<E$. By maximizing variance after accounting for the

variance in all preceding components, a collection of N new dimensions called Principal Components (PCs) is generated. Therefore, the first component explains the most variation, while the successive components explain a decreasing amount of variance.

Typically, computers are depicted as:

$$PC_j = b_1Y_1 + b_2Y_2 + ... + b_eY_e \tag{4}$$

In which PC_j denotes the j^{th} PC, Y_i denotes the initial feature of i and b_i denotes the numerical coefficient for Y_i.

RELM

Among the several machine learning techniques used today, one that stands out is the Back Propagation (BP) learning algorithm, which is employed by Single hidden-Layer Feed-forward neural Networks (SLFNs). Although these algorithms optimize the hidden layer biases and input weights to minimize the cost function, they incur a computational cost that is tolerable for an acceptable range of accuracy. Medical image classification pipelines use RELM. Highly efficient and effective in handling high-dimensional data, RELM is an ensemble learning technique. The reduced feature set using PCA is used to create predictions. RELM uses the reduced feature set for classification and PCA's informative features. In summary, PCA selects features, reduces dimensionality, and improves interpretability, while RELM is an efficient and regularised classifier that predicts based on selected features. These methods form a solid medical picture categorization method. As a learning algorithm, the ELM avoids repeatedly adjusting the artificial hidden nodes to decrease computing time. When dealing with SLFNs, ELM is recommended.

In order to represent the SLFN with an activation function $E(y)$ and M hidden nodes, we compose:

$$Z_M(Y) = \sum_{k=1}^{M} \alpha_k g_k(y) = g(y)\alpha_k \tag{5}$$

where, $\alpha = [\alpha_1, \dots . \alpha_m]$ is a hidden-to-output node weight matrix. The output of the secret node is $g_k(y)$. The hidden layer's parameters, including the input weight V_k and the hidden layer biases C_k, don't have to be tweaked like they are in SVM and other BP-based algorithms, therefore they can be produced at random before the training examples are gathered. The learning problem s_i *and* Z_i is solved by ELM, which, given M training samples, $\{(y_i, s_i)\}^{M}_{i=1}$ such that the value is minimized:

$$\left\|G\left(V_1,...,V_{\tilde{M}},c_1,...,c_{\tilde{M}}\right)\hat{\alpha} - S\right\| = \min_{\alpha}\left\|G\hat{\alpha} - S\right\| \tag{6}$$

$$G\left(V_1,...,V_M,c_1,...,c_M\right) = \begin{bmatrix} h(V_1.y_1+c_1)\ ...\ h(V_M.y_1+c_M) \\ h(V_1.y_M+c_1)\ ...\ h(V_M.y_M+c_M) \end{bmatrix} \tag{7}$$

$$\alpha = \begin{bmatrix} \alpha_1^S \\ . \\ . \\ . \\ \alpha_M^S \end{bmatrix} \tag{8}$$

$$S = \begin{bmatrix} S_1^S \\ . \\ . \\ . \\ S_M^S \end{bmatrix} \tag{9}$$

The output matrix from the hidden layer is denoted by G. α is the output weights' mathematical formula:

$$\alpha = G^{+}S \tag{10}$$

where, G^{+} has the speed advantages of an inverse of the matrix G that is generalized by Moore and Penrose, when working with big datasets and neuroimaging, ELM performs exceptionally well in multiclass and binary classification problems. By contrast, the accuracy suffers as a result of a rise in output error as computation time is reduced. This combination of ELM and sparse representation is intended to improve precision. The classification task is accomplished in two stages by this hybrid method (Peng *et al*., 2013; Qureshi *et al*., 2016). To begin, conventional techniques are used to train the ELM network. On the other hand, when testing is underway, reliability-based classification is used. In reliability-based classification, the ELM classifier is used if the test data is correctly classified; in sparse representation-based classification, it is used otherwise (Huang *et al*., 2011). As an additional step, a regularization term is introduced to strengthen the solution and enhance generalization performance. At last, the RELM's output weight can be written as:

$$\alpha = \left(\frac{I}{D} + G^sG\right)^{-1} G^sS \tag{11}$$

Results and Discussion

We tested the proposed method's performance with the RELM classifier. A large number of parameters influence the model's generalizability, fit quality, and training efficacy. The amount of input dimensions should not be more than the number of features or hidden nodes. Stay away from both of those things because they're bad for you. The trials showed that the 3-fold cross-validation method produced the best results, thus that's what the

CKELM training process would employ. There is continuous research into how the hidden layer affects accuracy and one training parameter is the amount of features. To determine the overall number of hidden nodes, the following methods are employed: 0,200, 400, 600, and 800. We next use 3-fold cross-validation to train three separate models using CNN, SVM, and RELM, all of which are convolutional neural networks. These training outcomes are shown in Fig. 3.

Figure 3 shows that as the number of experimental duplicates increases, the precision decreases in a linear fashion. The buried layer can accommodate up to 300 more nodes. Levels 400 and 600 typically exhibit a very consistent level of precision. The capacity to generalize decreased and overfitting became an issue after the 500th node. According to our findings, there are about 500 nodes that can be buried most efficiently. A number of hidden nodes influence a neural network's complex data pattern learning. Too few or too many nodes can cause underfitting and overfitting. The model's 500 nodes demonstrate its necessity to record complicated medical imaging data. Performance and empirical testing may select 500 hidden nodes. Model training and validation on a portion of the dataset with varied hidden node counts may reveal that 500 nodes optimize underfitting-overfitting. The experiment also showed that using just one CKELM classifier did not yield consistent results. The addition of a convolutional kernel ELM classifier greatly improves the CKELM classification model's accuracy.

After settling on 500 hidden nodes as the optimal number, this study uses four-fold cross-validation to compare the testing dataset's accuracy with and without model cross-validation, as well as the accuracy of each fold individually. Accuracy drops after 500 buried nodes. Overfitting may cause the decline. Hidden nodes may make the model too good at training data, causing noise and hindering generalization to unseen data. You may find a synopsis of the results in Table 1. According to experiments, using 3-fold cross-validation improves the accuracy of the resultant model on the test set.

In Table 2, you'll see that CNN, SVM, and RELM were picked as the most promising models. From the results of the experiments presented above, it is clear that the model based on the RELM algorithm achieves the best outcomes.

Using a dataset, Fig. 4 compares and contrasts the performance of different machine learning models. The accuracy percentage is indicated by the height of the line that connects each model, which is represented by a point. In general, the graph makes it easy to see how each model compared in terms of accuracy and it clearly shows that RELM was the most accurate one.

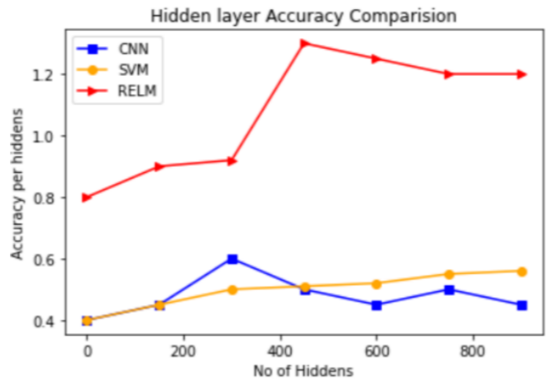

Fig. 3: Comparison of model accuracy under various hidden layers

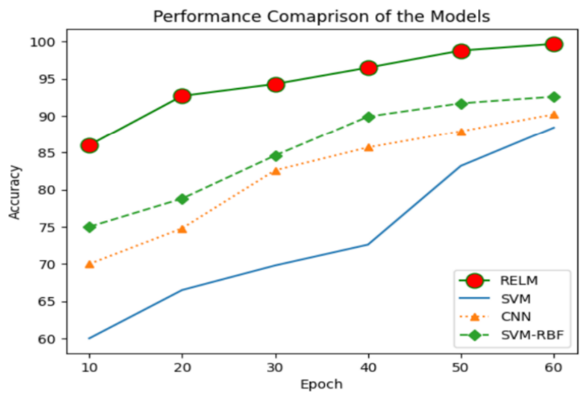

Fig. 4: Model accuracy comparison

Table 1: Comparison of model's cross-validation per fold and accuracy (%)

Models	1st fold	2nd fold	3rd fold	With CV	Without CV
SVM	80.45	86.74	88.43	80.63	75.00
CNN	83.56	87.76	90.17	88.25	80.34
SVM-RBF	85.28	89.92	92.62	86.28	81.57
RELM	90.43	97.32	99.68	98.48	97.78

Table 2: Performance comparison of the models (%)

Model	Accuracy	F1-score	Recall	Precision	F2-score
SVM	88.4	89.5	86.3	84.6	86
CNN	90.2	91.4	88.4	85.8	87.9
SVM-RBF	92.6	93.7	89.5	87.3	82.8
RELM	99.7	99.7	96.3	93.8	95.8

Conclusion

In terms of incidence and severity, lung cancer ranks first among cancers. The further advanced the cancer process was when it was identified, the higher the risk. This study suggests a new database of CT scans from lung cancer patients, called the IQ-OTH/NCCD dataset. This data analysis was used to develop a system that can detect lung cancer automatically. This study uses a Gabor filter for preprocessing in order to eliminate noise from the image. The segmentation process made use of Otsu's thresholding method. For feature extraction, the GLCM approach was used. Lastly, we will use RELM to train the model. Gabor characteristics represent lung tissue's intricate textures, enriching later phases. Otsu's approach improves lung picture contrast by calculating an ideal threshold, making possible anomalies easier to spot. GLCM-derived statistical descriptors including contrast, homogeneity, and energy characterize spatial dependencies in the image, helping identify aberrant patterns. The classification model, RELM, uses extracted features to identify pictures as malignant or not. GLCM, RELM, gabor filter, and Otsu's thresholding work synergistically. Each stage provides distinct information and these strategies are combined to diagnose lung cancer more effectively. The results showed that a RELM-based classifier worked very well for both binary and multiclass classification tasks. We also learned that feature selection using Principal Component Analysis (PCA) can improve accuracy slightly. One last step is to train the model using a CNN, SVM, SVM-RBF, and RELM combination. According to the results, RELM outperforms CNN and SVM in terms of accuracy. An accuracy of nearly 99.7% is the result of substantial improvements over earlier approaches.

Acknowledgment

We express our sincere gratitude to the institutions and individuals who contributed to the completion of this research. We extend our heartfelt thanks to our respective departments for their support and guidance throughout this study.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

K. Kavitha and V. Bhoopathy: Proposed work.

V. Saravana Kumar: Simulation.

Dhanalakshmi S.: Literature review.

Syed Arfath Ahmed: Dataset.

N. Valarmathi: Existing work.

Ethics

The research here reported has been respectful of any ethical consideration concerning neural network research.

References

AL-Yasriy, H. F. (2017). *The IQ-OTH/NCCD lung cancer dataset*. https://www.kaggle.com/datasets/hamdallak/the-iqothnccd-lung-cancer-dataset

Anushkannan, N. K., Balde, G. H., Suganthi, D., Pandian, P. M., Kaur, B., & Sagar, K. V. D. (2023). A Novel Method for Categorizing Brain Tumors using the Hybrid ALO-ELM Model. *2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)*, 1467–1472. https://doi.org/10.1109/icoei56765.2023.10125907

Yamini, B, S., K., Walid, Md. A. A., Prasad, J., Aparna, N., & Chauhan, A. (2023). Innovative Method for Detecting Liver Cancer using Auto Encoder and Single Feed Forward Neural Network. *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, 156–161. https://doi.org/10.1109/icaaic56838.2023.10140207

Bunn Jr, P. A., & Kelly, K. (1995). New treatment agents for advanced small cell and non-small cell lung cancer. *Seminars in Oncology*, *22*(3 Suppl 6), 53–63. https://doi.org/10.1016/0169-5002(96)84339-x

Gopinath, A., Gowthaman, P., Gopal, L., Abul Ala Walid, Md., Manju Priya, M., & Keshav Kumar, K. (2023a). Enhanced Lung Cancer Classification and Prediction based on Hybrid Neural Network Approach. *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, 933–938. https://doi.org/10.1109/icces57224.2023.10192798

Gopinath, A., Gowthaman, P., Alam, M. S., Prasad, J., Senthurya, S., & Vasudha. (2023b). R-LSTM-CNN Framework Based Lung Cancer Detection and Classification from Chest CT Images. *2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*, 575–580. https://doi.org/10.1109/icssas57918.2023.10331833

Hsieh, W. W. (2001). Nonlinear principal component analysis by neural networks. *Tellus A*, *53*(5), 599–615. https://doi.org/10.1034/j.1600-0870.2001.00251.x

Huang, G.-B., Wang, D. H., & Lan, Y. (2011). Extreme learning machines: A survey. *International Journal of Machine Learning and Cybernetics*, *2*(2), 107–122. https://doi.org/10.1007/s13042-011-0019-y

Subramanian, M, Ala Walid, Md. A., Sarada Prasanna Mallick, Dr., Rastogi, R., Chauhan, A., & Vidya, A. (2023). Melanoma Skin Cancer Detection using a CNN-Regularized Extreme Learning Machine (RELM) based Model. *2023 2nd International Conference on Electronics and Renewable Systems (ICEARS)*, 1239–1245. https://doi.org/10.1109/icears56392.2023.10085489

DOI: 10.3844/jcssp.2024.1243.1250

Mohanaiah, P., Sathyanarayana, P., & L.G. (2013). Image texture feature extraction using GLCM approach. *International Journal of Scientific and Research Publications*, *3*(5), 1–5.

Narayanan, L. A., & Jeeva, J. B. (2015). A Computer Aided Diagnosis for detection and classification of lung nodules. *2015 IEEE 9th International Conference on Intelligent Systems and Control (ISCO)*, 1–5. https://doi.org/10.1109/isco.2015.7282242

Peng, X., Lin, P., Zhang, T., & Wang, J. (2013). Extreme Learning Machine-Based Classification of ADHD Using Brain Structural MRI Data. *PLoS ONE*, *8*(11), e79476. https://doi.org/10.1371/journal.pone.0079476

Qureshi, M. N. I., Min, B., Jo, H. J., & Lee, B. (2016). Multiclass Classification for the Differential Diagnosis on the ADHD Subtypes Using Recursive Feature Elimination and Hierarchical Extreme Learning Machine: Structural MRI Study. *PLOS ONE*, *11*(8), e0160697. https://doi.org/10.1371/journal.pone.0160697

Rajadell, O., García-Sevilla, P., & Pla, F. (2009). Textural Features for Hyperspectral Pixel Classification. In *Pattern Recognition and Image Analysis* (208–216). Springer. https://doi.org/10.1007/978-3-642-02172-5_28

Sarkar, A., Badrinath, N., Kapse, P., Bhalerao, D. M., Singh, K., & Sagar, K. V. D. (2023). An Effective Method for Skin Cancer Detection using Convolutional Kernel Extreme Learning Machine. *2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)*, 1088–1093. https://doi.org/10.1109/icoei56765.2023.10126032

Teramoto, A., Fujita, H., Yamamuro, O., & Tamaki, T. (2016). Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional neural network technique. *Medical Physics*, *43*(6Part1), 2821–2827. https://doi.org/10.1118/1.4948498

Wang, X., Ding, X., & Liu, C. (2005). Gabor filters-based feature extraction for character recognition. *Pattern Recognition*, *38*(3), 369–379. https://doi.org/10.1016/j.patcog.2004.08.004