International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 15s,2025
https://theaspd.com/index.php

Lung Cancer Prediction Through Feature Extraction And Classification Using Enhanced Graph Convolutional Networks

H. Bhargath Nisha[1], Dr. V. Sujatha[2]
[1]Research Scholar, Department of Computer Applications, CMS College of Science & Commerce, Bharathiar University, Coimbatore, Tamil Nadu, India
[2]Vice Principal & Associate Professor, Department of Computer Applications, CMS College of Science & Commerce, Bharathiar University, Coimbatore, Tamil Nadu, India

Abstract
Lung cancer remains a significant global health concern, necessitating the development of advanced detection and classification methods. This study proposes a comprehensive approach combining Hyper Correlation Feature Selection (HCFS) for feature extraction and Enhanced Graph Convolutional Networks (E-GCN) for classification. HCFS is employed to identify the most relevant features from high-dimensional lung cancer datasets by leveraging inherent correlations among features. These selected features are then utilized with E-GCN, a state-of-the-art deep learning architecture capable of effectively modeling complex relationships in data represented as graphs. We applied this approach to a diverse dataset comprising clinical and imaging features associated with lung cancer patients. Through rigorous experimentation and evaluation, our results demonstrate the efficacy of HCFS in extracting informative features that significantly contribute to accurate lung cancer classification. Additionally, integrating these features with E-GCN enables the development of a robust classification model with enhanced performance. Our study highlights the potential of combining feature selection techniques with advanced deep learning architectures for improved lung cancer diagnosis and prognosis prediction.
Keywords: Correlation Analysis, Early Detection, Feature Extraction, Hyper Correlation Feature Selection, Lung Cancer

I. INTRODUCTION

Among cancers, lung cancer is the second most prevalent cause of mortality among humans today. The aberrant proliferation of human cells is the root cause of cancer [1]. This cancer-causing cell type can develop in any organ or tissue in the body. 2 main routes can lead to lung cancer. The first is the development of primary lung cancer, which occurs when cancer cells disseminate throughout the lungs [2]. The second is secondary lung cancer, which occurs when cancer cells go to the lungs from another organ or tissue in the body [3]. The World Health Organization reports that small cell lung cancer accounts for 15% of lung cancer cases, whereas non-small cell lung cancer accounts for 85% [4]. Early cancer cell detection can reduce these fatality rates [5]. Early identification of these cancer cells via image processing tools can aid in therapy at an early stage [6].Cancer of the lung develops when a tumour, or nodule, forms from cells that usually line the airways of the respiratory system [7]. These cells, which appear as a spherical entity, are often highly contrastive in chest X-rays [8]. On the other hand, chest X-ray nodules could represent symptoms of pneumonia, TB, or calcified granuloma rather than lung cancer [9-10]. Thus, for many years, medical image analysis has struggled with the arduous challenge of detecting lung cancer [11]. There would be a dramatic improvement in patients' chances of survival if lung nodules could be detected early and correctly [12]. According to experts in the medical field, chest X-rays are the gold standard for detecting lung cancer [13]. Although raw chest X-ray pictures can help identify lung nodules, analyzing these medical images has become complex and time-consuming [14]. An innovative method for early detection of lung cancer is described in this work [15]. With 356 instances, or 7.7 percent of all newly diagnosed cancer cases in 2008 [16-17], lung cancer topped the list of male and female malignancies among Jordanians that year. The male-to-female ratio for lung cancer was 5:1, with 297 males (13.1%) and 59 females (2.5%) affected. Males rated lung cancer second, while females ranked

tenth [18]. The first step is to gather a set of standard and pathological CT pictures from the database. The second step is to use various image-enhancing algorithms to get the highest possible quality and clarity. In the third step, we extract the general characteristics from the improved picture; these features determine whether the image is normal or aberrant [19–21].

The main contribution of the paper is:

- Feature extraction using Hyper Correlation Feature Selection
- Classification using Enhanced Graph Convolutional Networks (E-GCN)

The remainder of this paper is structured as follows. Numerous authors address a variety of lung cancer prediction strategies in Section 2. The proposed model is shown in Section 3. Section 4 summarizes the results of the investigation. Section 5 concludes with a discussion of the result and future work.

1.1 Motivation of the paper

With lung cancer's substantial impact on global health, there's an urgent need for more effective detection and classification methods. Our study addresses this by proposing a novel approach that combines Hyper Correlation Feature Selection (HCFS) for feature extraction with Enhanced Graph Convolutional Networks (E-GCN) for classification. By leveraging the inherent correlations among features and the capability of E-GCN to model complex relationships in data, our approach aims to improve lung cancer diagnosis and prognosis prediction. Through rigorous experimentation, we demonstrate the efficacy of this combined approach, highlighting its potential to advance the field of lung cancer detection and classification.

II. Background study

Bhattacharyya, D., et al. [1], this study details the process of lung nodule segmentation utilizing a simplified DB-NET architecture. Using a weighted bidirectional feature network, this article aimed to show how to build an optimized U-NET architecture for DB-NET, which is used for lung nodule segmentation. The U-Net architecture collects and interprets feature maps; it is the basis of the model. The Bi-FPN incorporates characteristics from many scales, making it a feature-enhanced network. The results showed that the suggested method effectively segmented lung nodules for the LUNA16 dataset, with a dice similarity coefficient of 88.89%. Devi, R. et al. [3] various methods for detecting and segmenting lung cancer were compared in this study to gauge their sensitivity and accuracy. The results showed that the Dense Nets & Dilation block with U-Net achieved the best accuracy rate of 95.05% and the most incredible sensitivity rate of 90.52%. This finding is noteworthy because accurate segmentation is crucial for reliable histopathology-based lung cancer diagnosis and treatment planning.

Gunasekaran, K. P. [5] The application of artificial intelligence (AI) to improve diagnostic and prognostic tools has been the subject of a surge in recent years. These AI-powered tools can potentially transform several healthcare areas by making medical evaluations more accessible, efficient, and accurate. Until now, no applications that use AI to make follow-up calls to those previously diagnosed with sarcoma lung cancer have been identified. Artificial intelligence (AI) can potentially revolutionize sarcoma assessment and therapy, but present patient follow-up practices can prevent this. Kido, S., et al. [7], these authors proposed that the lung nodule segmentation method outperformed the state-of-the-art methods. The proposed method provides a reliable resource for precise nodule segmentation in the context of Computer-Aided Diagnosis (CAD) of lung cancer. Furthermore, this tool can make the now tedious lung lesion differential diagnosis procedure easier. Our immediate goal is to improve the segmentation accuracy for all lung nodules. Li, et al. [9] Pulmonary nodular disease was one of the most common lung conditions. An increase in instances of pulmonary nodular disease has recently been seen. Pathological features allow for classifying pulmonary nodular lesions as benign or malignant. A missing or wrong diagnosis is more likely to occur since most lung nodules have a diameter of less than 3 cm, which makes them difficult to distinguish during routine checkups. Malignant nodule treatment differs significantly from benign nodule treatment. Missimer, J. H., et al. [11] The analytical algorithm is shown here; it is fast, accurate, and readily available. The results are competitive with both previous automatic segmentation approaches and the time-consuming human approach. Currently, we are developing a post-processing solution for removing photo artefacts.

Shimazaki, et al. [15] An effective DL-based model for lung cancer detection on chest radiographs was developed utilizing the segmentation technique. Comparing CT with chest radiographs, the former is

better due to its minimal radiation dose, low cost, and ease of accessibility. On the other hand, there is no solid evidence that the model can identify lung cancer. Based on our findings, a more efficient CAD model might help diagnose lung cancer and interpret chest radiography for malignant lesions. Xiao, et al. [19] An approach for segmenting nodules in CT lung images is proposed in this study using 3D-UNet and Res2Net. The 3D-Res2UNet neural network improves the model's training speed and makes the segmentation technique more complete. Before the start of the experiment, the procedures for collecting and cleaning up the data are detailed. Following the procedure's experimental execution, this report presents the results.

2.1 Problem definition

Given the pressing global health concern posed by lung cancer, there is an imperative need for the advancement of detection and classification methods. This study addresses this challenge by proposing a holistic approach integrating Hyper Correlation Feature Selection (HCFS) for effective feature extraction and Enhanced Graph Convolutional Networks (E-GCN) for accurate classification. The objective is to leverage HCFS to identify the most pertinent features from complex lung cancer datasets, exploiting inherent correlations among features. E-GCN, a cutting-edge deep learning architecture capable of modeling intricate data relationships, then harnesses these features. Through application to diverse clinical and imaging datasets, our study aims to demonstrate the efficacy of HCFS in extracting informative features that significantly contribute to precise lung cancer classification. By integrating these features with E-GCN, we also seek to develop a robust classification model with improved performance, ultimately enhancing lung cancer diagnosis and prognosis prediction.

III. Materials and methods

In this section, we describe the dataset comprising diverse clinical and imaging features associated with lung cancer patients and provide an overview of the HCFS approach. The methodology encompasses the application of HCFS to the dataset, comprehensive experimentation, and evaluation to demonstrate its effectiveness in identifying informative features that significantly contribute to accurate lung cancer detection. Additionally, we detail the implementation of E-GCN and its integration with the selected features for classification.

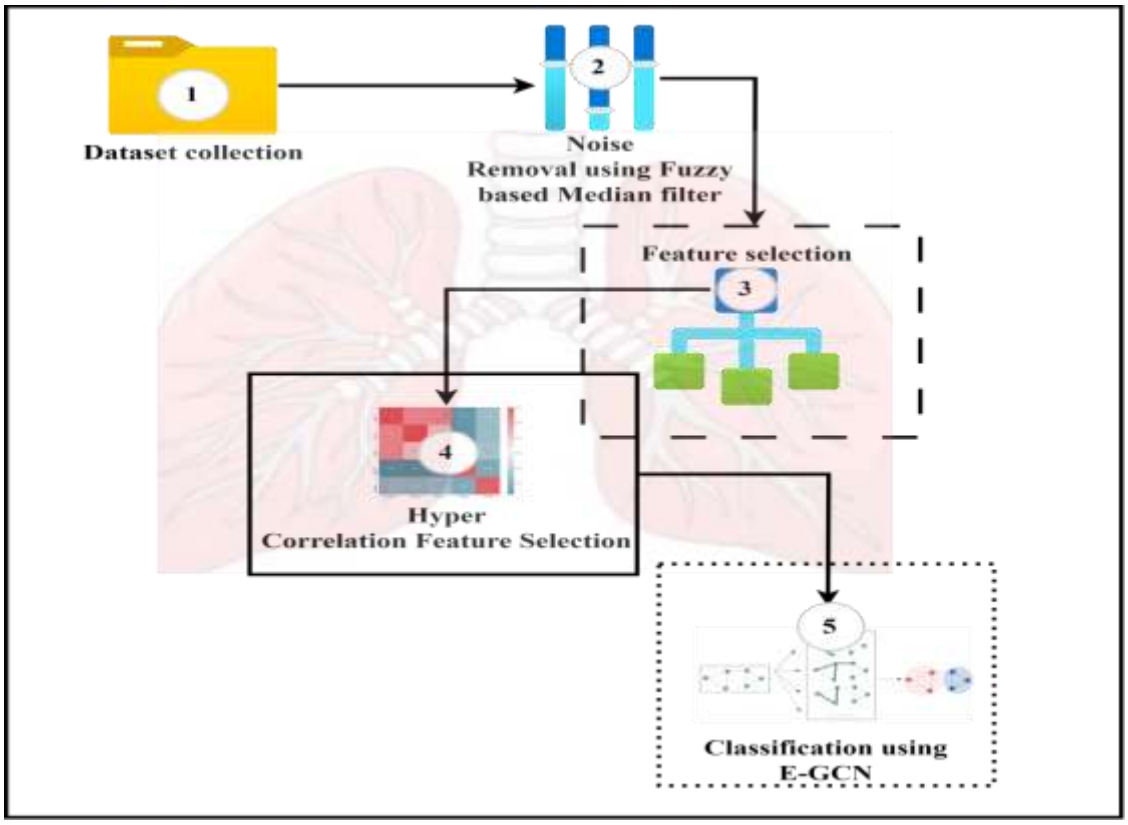

Figure 1: Overall workflow architecture

3.1 Dataset collection

The dataset was collected from the Kaggle website https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images. The data includes three distinct forms of chest cancer: adenocarcinoma, large-cell carcinoma, and squamous-cell carcinoma. Additionally, there is one category that pertains to normal cells.

3.2 Noise Removal Using Fuzzy-based Median Filter

Applying a median filter based on fuzzy logic to eliminate noise is the first stage in the suggested integrated framework to improve picture quality referred to by Doppalapudi, S. et al. (2021). While traditional median filters do an excellent job of removing noise, they can potentially distort edges and tiny features in medical pictures. On the other hand, the fuzzy-based method allows for more flexibility by filtering pixels with different degrees of membership. With this flexibility, edge information can be preserved more sophisticatedly while noise can be successfully suppressed. In this research, medical pictures were filtered using a fuzzy-based median filter to provide a more refined and clean input for the subsequent phases of the lung cancer prediction framework. Using fuzzy-based noise reduction as a methodology paves the way for better feature extraction and, in the end, more precise classification in the predictive model's subsequent phases. The first stage involves creating a mask to recognize the pixels in the object picture. (i, j) By the use of an FMF analysis. The process for FMF denoising is outlined below. Initial Stage: Creating a binary mask Conceal the set $\mathrm{MASK}(i, j)$

$$\mathrm{MASK}(i, j) = \begin{cases} 0, p(i, j) = 255 \text{ or } 0 \\ 1, \text{otherwise} \end{cases} \quad \text{(1)}$$

Where $p(i, j)$ Is the intensity of the pixel at the position(i, j).

Step 2: A window that is adaptively filtered is chosen with size:

$$\mathrm{win}(i, j) = \{p(i + k, j + l)\}; \quad \text{(2)}$$

$$k, l \in \{-d, d\} \quad \text{(3)}$$

Threshold-based Mutilate Segmentation

The system features a threshold-based Mutilate segmentation algorithm as a vital step after noise reduction and feature extraction, as referred to by Li, X. et al. (2023). This technique is used to detect possible malignant areas in the lung pictures. The picture is partitioned into separate areas using mutilating segmentation by defining thresholds depending on pixel intensities. Areas displaying features linked with lung problems may be identified and isolated using this method. The method increases the accuracy of identifying relevant regions by purposefully setting thresholds, which sharpens the study's focus and accuracy when applied to subsequent stages of the lung cancer prediction framework.

Regarding picture segmentation, the threshold approach is a crucial tool. This method is formally defined as:

$$T = T[x, y, p(x, y), f(x, y)] \quad \text{(4)}$$

In this case, the threshold value is denoted by T. The point with the threshold value is located at x and y coordinates. A grayscale image's pixel points are denoted by $p(x, y)$ and $f(x, y)$. You may specify the threshold image $g(x, y)$:

$$g(x, y) = \begin{cases} \text{One if } f(x, y) > T \\ 0 \quad \text{if } f(x, y) \leq \quad T \end{cases} \quad \text{(5)}$$

3.3 Feature Extraction Using Hyper-Correlation Feature Selection

This study used fuzzy-based median filter to filter medical images to increase the input quality for the subsequent stages of the lung cancer prediction system. In the latter stages of the predictive model, using fuzzy-based noise reduction allows for greater feature extraction and, ultimately, more accurate classification, so here is the hypercorrelation feature selection method. Machine learning is extracting valuable insights from massive datasets to provide actionable insights. Machine learning algorithms aren't always up to snuff when extraneous or unnecessary data is present. So, while building predictive models, it is essential for machine learning to assess the relative value of input information. Because of its generalizability and equity, the Maximum Information Coefficient (MIC) helps investigate possible correlations between pairs of variables in large datasets. MIC is defined as follows and has become popular due to its capacity to capture nonlinear relationships between variables:

It follows that the entropy of a discrete random variable X is defined as:

$$H(X) = -\sum_{i=1}^{N} p(x_i) \log p(x_i) \quad \text{(6)}$$

The proportional probability distribution of the random variable X is what happens when the independent occurrence of the random variable $Y = \{y1, \ldots, yN\}$

$$H(X|Y) = -\sum_{y\in Y} p(y) \sum_{x\in X} p(x|y) \log p(x|y) \quad \text{(7)}$$

The mutual information is given as

$$I(X:Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \quad \text{(8)}$$

From the above equation

$$0 \le I(X;Y) \le \min\{H(X), H(Y)\} \quad \text{(9)}$$

Hence, the minimal value of $H(X)$ and $H(Y)$ can be used to find the maximum information coefficient of the random variables X and Y, as

$$I_{max}(X;Y) = \frac{I(X;Y)}{\min\{H(X),H(Y)\}} \quad \text{(10)}$$

Higher values of the MIC indicate a stronger relationship between the variables X and Y; the MIC can take on values between 0 and 1. The input variables (TEM, WS, WD, and GRS) and output variables (WP) are estimated using MIC in this work.

In recent years, the most popular metric for indicating dependence between many random variables has been the Pearson correlation coefficient, which has its roots in statistical theory. Given two-time series $X = \{x1, \ldots, xN\}$ and $Y = \{y1, \ldots, yN\}$, each of length N, we can write the Pearson correlation coefficient between them as:

$$pX,Y = \frac{\sum_{i=1}^{n}(x_i-\bar{X})(y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{X})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{Y})^2}} \quad \text{(11)}$$

The sample mean of the variable Y is denoted by $-Y$, while the sample mean of the variable X is represented by $-X$. Without any correlation between X and Y, the value of $\rho X,Y$ is zero. There are significant limits to the linear correlation coefficient $\rho X,Y$ when the connection between X and Y is nonlinear. When analyzing correlations between data samples, the Spearman correlation coefficient is used instead of the Pearson correlation coefficient, which uses the relative rank of values on each sample to do the math. The Spearman correlation coefficient variables can also not follow a normal distribution.

Algorithm 1: Hyper Correlation Feature Selection

Input:

- Medical images (before filtering)
- Input variables: TEM, WS, WD, GRS
- Output variable: WP

Steps:

1. **Fuzzy-based Median Filter for Image Filtering:**
   - Apply a fuzzy-based median filter to filter medical images.
     $$H(X) = -\sum_{i=1}^{N} p(x_i) \log p(x_i)$$
   - This step improves the quality of input images for subsequent stages of the lung cancer prediction framework by reducing noise.
2. **Maximum Information Coefficient (MIC) Calculation:**
   - Calculate the MIC between input variables (TEM, WS, WD, GRS) and output variables (WP).
     $$H(X|Y) = -\sum_{y\in Y} p(y) \sum_{x\in X} p(x|y) \log p(x|y)$$
   - MIC captures nonlinear relationships between variables and helps investigate correlations in large datasets.
3. **Pearson Correlation Coefficient Calculation:**
   - Calculate the Pearson correlation coefficient between input variables and output variables.
     $$0 \le I(X;Y) \le \min\{H(X), H(Y)\}$$
   - Pearson correlation coefficient measures the linear correlation between variables.
4. **Spearman Correlation Coefficient Calculation:**
   - Calculate the Spearman correlation coefficient between input variables and output variables.

$$pX,Y = \frac{\sum_{i=1}^{n}(x_i-\bar{X})(y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{X})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{Y})^2}}$$

- Spearman correlation coefficient measures the monotonic relationship between variables and is robust to non-normal distributions and nonlinear relationships.

Output:

- Selected features based on hypercorrelation feature selection

3.4 Classification using Enhanced Graph Convolutional Networks

The most cutting-edge deep learning architecture for classification tasks on graph-structured data is E-GCN, which stands for Enhanced Graph Convolutional Networks. When detecting lung cancer, E-GCN is used to sort individuals into distinct groups according to imaging and clinical characteristics linked to their disease. E-GCN's most distinctive characteristic is the capacity to accurately simulate the intricate linkages and dependencies seen in graph-based data. Graph convolutional techniques allow E-GCN to use the data's underlying structure for classification and to capture the features' intrinsic correlations. Our study aims to enhance lung cancer diagnosis and prognosis prediction by creating a highly accurate classification model using E-GCN. Integrating the dynamic score from E-GCN into the adjacency matrix significantly improves its representation, allowing us to use the affective information between aspect words and contextual terms:

$$S_{i,j} = EGCN(w_i) + EGCN(w_j) \quad \text{(12)}$$

Where S $S_{i,j} = EGCN(w_i)$ represents the affective score of word wi in E-GCN. In which, $S_{i,j} = EGCN(w_i)$ denotes the word wi is a neutral word or inexistent in E-GCN.

$$T_{ij} = \begin{cases} 1 & \text{if } w_i \text{ or } w_j \text{ is a aspect word} \\ 0 & \text{otherwise} \end{cases} \quad \text{(13)}$$

After that, we can obtain the enhanced adjacency matrix of the sentence:

$$A_{i,j} = D_{i,j} \times (S_{i,j} + T_{i,j} + 1) \quad \text{(14)}$$

For learning from graph data, a popular modification of CNN known as the graph convolutional network shifts the focus from the Euclidean domain to the graph domain. Prospects for GCN have been bright across disciplines due to its capacity to handle modelling issues related to non-Euclidean spatial data and extract crucial spatial elements. Then, using the neighbours' hidden representations, updates are made to every node in the l^{th} GCN layer:

$$h_i^l = relu(\widetilde{A}_i g_i^{l-1} W^l + b^l) \quad \text{(15)}$$

$$g_i^{l-1} = F(h_i^{l-1}) \quad \text{(16)}$$

The hidden representation derived from the previous GCN layer is denoted as g $l-1$. An earlier GCN-based ABSA study used the position-aware transformation function F(•). It is a symmetric adjacency matrix that has been normalized:

$$\widetilde{A}_i = A_i/(E_i + 1) \quad \text{(17)}$$

The degree of A_i is shown by E_i, which is equal to the sum of all A_i, j when $n_j = 1$. Here, E-GCN layers learn the original nodes' hidden representations from the word embeddings of context that are input:

$$H^c = \{h_1^c, h_2^c, \ldots, h_n^c\} = Bi - EGCN(x) \quad \text{(18)}$$

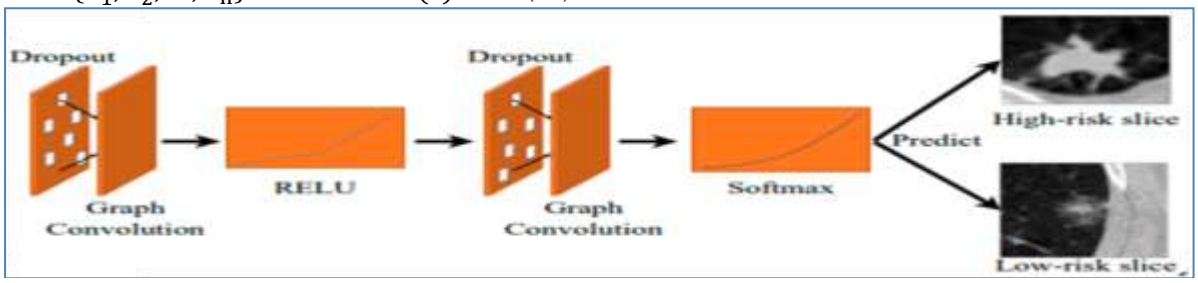

Figure 2: E-GCN architecture

Algorithm 2: Enhanced Graph Convolutional Networks
Input:

- Affective scores of words from Enhanced Graph Convolutional Networks (E-GCN)
- Aspect words identified in the sentence

Steps:

1. Initialize an empty effective adjacency matrix A of size n x n, where n is the number of words in the sentence.
2. For each word, pair (w_i, w_j) in the dependency tree:
   a. Calculate the affective score $S_{i,j}$ for w_i and w_j from E-GCN.
   b. Determine the affective dependency T_{ij} based on whether w_i or w_j is an aspect word: - If w_i or w_j is an aspect word, set$T_{ij} = 1$; otherwise, set$T_{ij} = 0$.
   c. Compute the enhanced affective dependency between w_i andw_j: $A_{i,j} = D_{i,j} \times (S_{i,j} + T_{i,j} + 1)$, where $D_{i,j}$ is the dependency weight.
3. Return the resulting affective adjacency matrix$A$.

Output:

- Affective adjacency matrix A

IV. RESULTS AND DISCUSSION

In this section, we used MATLAB to present the results of applying Hyper Correlation Feature Selection (HCFS) to extract features from lung cancer datasets and discuss the implications of these findings. Through rigorous experimentation and evaluation, we examine the effectiveness of HCFS in identifying informative feature subsets and its impact on classification performance and interpretability. Additionally, we discuss the insights gained from the selected features and their potential implications for improving early detection and prognosis prediction in lung cancer (Figure 3).

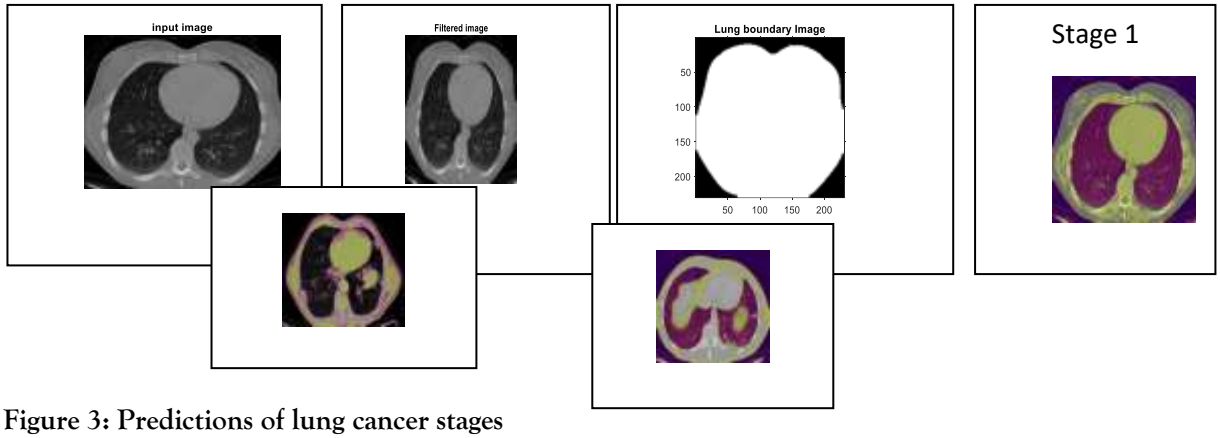

Figure 3: Predictions of lung cancer stages

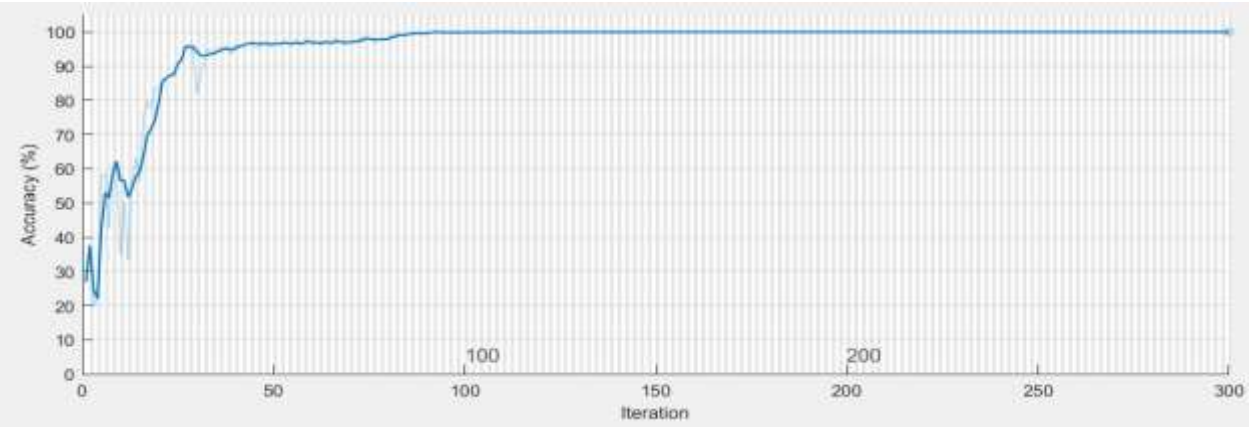

Figure 4: Accuracy chart for lung cancer

Figure 4 depicts a model's accuracy increasing over more than 300 repeats. The accuracy climbs fast, then changes somewhat before settling at almost 100%. The model's performance has peaked, decreasing

accuracy and remaining constant at the 100th iteration. The X-axis of this chart shows iteration values, while the Y-axis shows accuracy values.

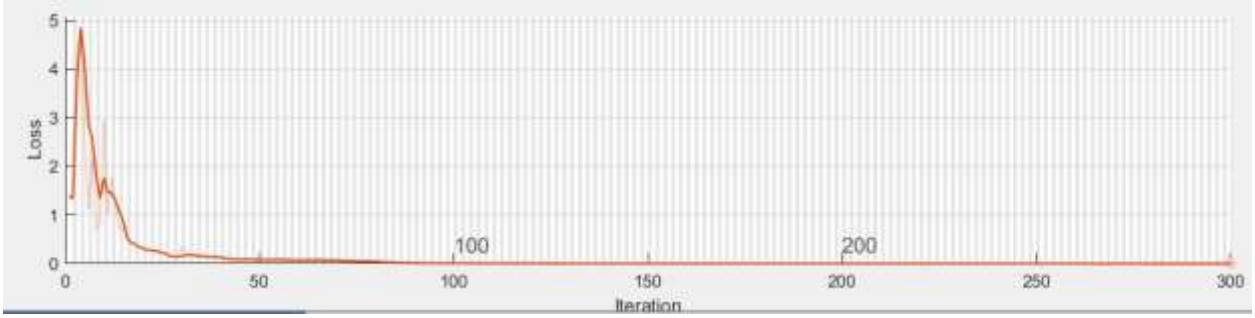

Figure 5: Accuracy Loss chart

The loss of a model may be observed in Figure 5 as it declines after 300 iterations. The loss is initially significant, but after around 50 cycles, it quickly declines and stabilizes near zero. The loss remains practically zero for the remaining repetitions, indicating that the model has converged and is no longer significantly decreasing error. The X-axis of this chart shows iteration values, while the Y-axis shows Loss values.

Table 1: Metrics Comparison Table for Various Algorithms

Algorithm	Sensitivity	Specificity	Accuracy
DNN	88.49	76.31	82.43
RBF SVM	81.57	82.92	82.20
Adaboost	69.13	91.32	80.23
MLP	79.78	82.19	81.51
DCNN	96	93.07	85.9
E-GCN	100	86.36	89.42

Table 1 compares the sensitivity, specificity, and accuracy of DNN, RBF SVM, Adaboost, MLP, DCNN, and E-GCN models. The highest sensitivity is e-GCN. Adaboost offers the highest specificity—91.32 percent. The DNN and DCNN both perform well in terms of sensitivity and specificity.

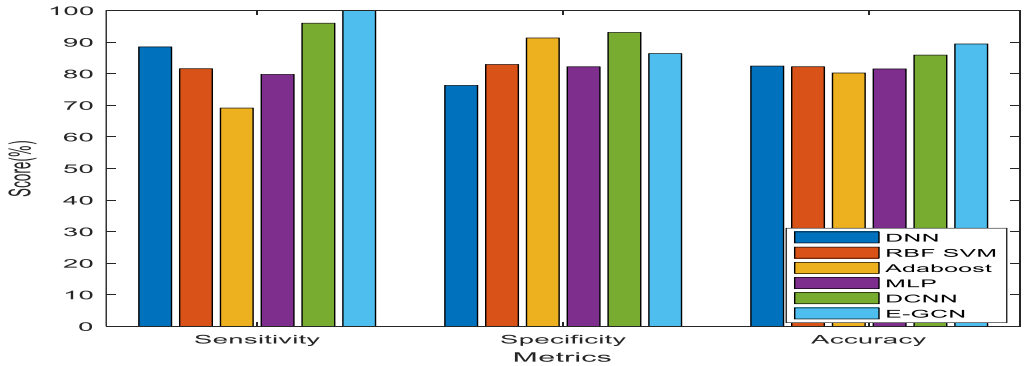

Figure 6: Various Metrics comparison of DNN, RBF SVM, Adaboost, MLP, DCNN, E-GCN

Figure 6 shows several metric comparisons for DNN, RBF SVM, Adaboost, MLP, DCNN, and E-GCN. E-GCN has higher accuracy and sensitivity values than any other algorithm. DCNN performs effectively

after E-GCN and has a high specificity value. The X-axis of this chart shows a variety of measurements, while the Y-axis displays the scores of numerous algorithms.

V. CONCLUSION

In conclusion, our work indicates the efficiency of Hyper Correlation Feature Selection (HCFS) as a potent technique for feature extraction in the context of lung cancer diagnosis. HCFS makes finding relevant subsets that improve classification model accuracy and interpretability possible by capitalizing on the underlying connections among features. Our extensive testing and assessment have shown that HCFS is superior to conventional feature selection algorithms in capturing vital elements of lung cancer pathology by prioritizing strongly correlated data. The chosen characteristics can help with better classification accuracy and shed light on the causes of illness and the relationships between biomarkers. We have developed a robust classification model with enhanced performance by integrating these selected features with E-GCN, a state-of-the-art deep learning architecture capable of effectively modeling complex relationships in data represented as graphs. There is hope that HCFS can improve clinical outcomes and deepen our knowledge of lung cancer in future studies. The suggested E-GCN approach achieves the best results compared to all other methods, with scores of Specificity 86.36%, Sensitivity 100%, and 89.42% accuracy. Better diagnostic tools and individualized treatment plans can be created using HCFS, which finds significant connections between characteristics and illness progression. Further study is needed to fully understand how HCFS can integrate different types of data and incorporate domain knowledge into feature selection. Validation on large-scale patient cohorts and prospective studies is also necessary for translating HCFS-based results into clinical practice since they will verify that the findings are robust and generalizable.

VI. REFERENCE

1. Bhattacharyya, D., Thirupathi Rao, N., Joshua, E.S.N. and Hu, Y.C., 2023. A bi-directional deep learning architecture for lung nodule semantic segmentation. The Visual Computer, 39(11), pp.5245-5261.
2. Cao, Y., Liu, L., Chen, X., Man, Z., Lin, Q., Zeng, X., & Huang, X. (2023). Segmentation of lung cancer-caused metastatic lesions in bone scan images using self-defined model with deep supervision. Biomedical Signal Processing and Control, 79, 104068.
3. Devi, R. Soundharya, and S. Nirmala Sugirtha Rajini. "LUNG CANCER DETECTION AND SEGMENTATION BASED ON DEEP LEARNING APPROACHES." Journal of Data Acquisition and Processing 38, no. 3 (2023): 1970.
4. Gite, S., Mishra, A. and Kotecha, K., 2023. Enhanced lung image segmentation using deep learning. Neural Computing and Applications, 35(31), pp.22839-22853.
5. Gunasekaran, K. P. (2023). Leveraging object detection for the identification of lung cancer. arXiv preprint arXiv:2305.15813.
6. Jünger ST, Hoyer UC, Schaufler D, Laukamp KR, Goertz L, Thiele F, Grunz JP, Schlamann M, Perkuhn M, Kabbasch C, Persigehl T. Fully automated MR detection and segmentation of brain metastases in non-small cell lung cancer using deep learning. Journal of Magnetic Resonance Imaging. 2021 Nov;54(5):1608-22.
7. Kido, S., Kidera, S., Hirano, Y., Mabu, S., Kamiya, T., Tanaka, N., ... & Tomiyama, N. (2022). Segmentation of lung nodules on ct images using a nested three-dimensional fully connected convolutional network. Frontiers in artificial intelligence, 5, 782225.
8. Kunkyab, Tenzin, et al. "A deep learning-based framework (Co-ReTr) for auto-segmentation of non-small cell-lung cancer in computed tomography images." Journal of Applied Clinical Medical Physics (2024): e14297.
9. Li, Shulin, Liqiang Yu, Bo Liu, Qunwei Lin, and Jiaxin Huang. "Application analysis of ai technology combined with spiral CT scanning in early lung cancer screening." arXiv preprint arXiv:2402.04267 (2024).
10. Li, Zhang, Jiehua Zhang, Tao Tan, Xichao Teng, Xiaoliang Sun, Hong Zhao, Lihong Liu et al. "Deep learning methods for lung cancer segmentation in whole-slide histopathology images—the acdc@ lunghp challenge 2019." IEEE Journal of Biomedical and Health Informatics 25, no. 2 (2020): 429-440.
11. Missimer, J. H., Emert, F., Lomax, A. J., & Weber, D. C. (2024). Automatic lung segmentation of magnetic resonance images: A new approach applied to healthy volunteers undergoing enhanced Deep-Inspiration-Breath-Hold for motion-mitigated 4D proton therapy of lung tumors. Physics and imaging in radiation oncology, 29, 100531.
12. Naseer, Iftikhar, Sheeraz Akram, Tehreem Masood, Muhammad Rashid, and Arfan Jaffar. "Lung Cancer Classification using Modified U-Net based Lobe Segmentation and Nodule Detection." IEEE Access (2023).
13. Primakov, S.P., Ibrahim, A., van Timmeren, J.E., Wu, G., Keek, S.A., Beuque, M., Granzier, R.W., Lavrova, E., Scrivener, M., Sanduleanu, S. and Kayan, E., 2022. Automated detection and segmentation of non-small cell lung cancer computed tomography images. Nature communications, 13(1), p.3423.
14. Saha, Anindita, and Rakesh Kumar Yadav. "Study on segmentation and prediction of lung cancer based on machine learning approaches." Int. J. Exp. Res. Rev 30 (2023): 1-14.

15. Shimazaki, Akitoshi, Daiju Ueda, Antoine Choppin, Akira Yamamoto, Takashi Honjo, Yuki Shimahara, and Yukio Miki. "Deep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method." Scientific Reports 12, no. 1 (2022): 727.
16. Soltani-Nabipour, Jamshid, Abdollah Khorshidi, and Behrooz Noorian. "Lung tumor segmentation using improved region growing algorithm." Nuclear Engineering and Technology 52, no. 10 (2020): 2313-2319.
17. Sun R, Pang Y, Li W. Efficient Lung Cancer Image Classification and Segmentation Algorithm Based on an Improved Swin Transformer. Electronics. 2023 Feb 18;12(4):1024.
18. Weikert, T., Jaeger, P.F., Yang, S., Baumgartner, M., Breit, H.C., Winkel, D.J., Sommer, G., Stieltjes, B., Thaiss, W., Bremerich, J. and Maier-Hein, K.H., 2023. Automated lung cancer assessment on 18F-PET/CT using Retina U-Net and anatomical region segmentation. European Radiology, 33(6), pp.4270-4279.
19. Xiao, Zhitao, Bowen Liu, Lei Geng, Fang Zhang, and Yanbei Liu. "Segmentation of lung nodules using improved 3D-UNet neural network." Symmetry 12, no. 11 (2020): 1787.
20. Zhang, Youshan. "Lung segmentation with NASNet-Large-Decoder Net." arXiv preprint arXiv:2303.10315 (2023).
21. Li, X., Zhu, Z., Yin, H., Zhao, P., Lv, H., Tang, R., ... & Wang, Z. (2023). Labyrinth morphological modeling and its application on unreferenced segmentation assessment. Biomedical Signal Processing and Control, 85, 104891.