Performance analysis of texture characterization techniques for lung nodule classification

Ishan Devdatt Kawathekar1, Anu Shaju Areeckal1*

1Department of Electronics & Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Karnataka, India

*Corresponding author: anu.areeckal@manipal.edu

Abstract. Lung cancer ranks very high on a global index for cancer-related casualties. With early detection of lung cancer, the rate of survival increases to 80-90%. The standard method for diagnosing lung cancer from Computed Tomography (CT) scans is by manual annotation and detection of the cancerous regions, which is a tedious task for radiologists. This paper proposes a machine learning approach for multi-class classification of the lung nodules into solid, semi-solid, and Ground Glass Object texture classes. We employ feature extraction techniques, such as gray-level co-occurrence matrix, Gabor filters, and local binary pattern, and validate the performance on the LNDb dataset. The best performing classifier displays an accuracy of 94% and an F1-score of 0.92. The proposed approach was compared with related work using the same dataset. The results are promising, and the proposed method can be used to diagnose lung cancer accurately.

Keywords: Lung cancer, Nodule, Texture, Feature extraction, Classification, CT images

1. Introduction
Lung cancer is one of the deadliest forms of cancer and is responsible for the majority of cancer-related deaths. In 2020, around 2.2 million patients were diagnosed with lung cancer, and 1.8 million deaths were attributed to lung cancer.

Early diagnosis of lung cancer is paramount in the early diagnosis and treatment of patients. Accurate detection techniques have led to an improved patient survival rate. Low-dose helical Computed Tomography (CT) is widely accepted as the gold standard for screening patients. Nevertheless, this screening technique is riddled with false positives. Automating or semi-automating the detection of lung cancer can reduce the time taken for manual detection and also bring down the number of false positives considerably.

The conventional technique of lung cancer detection is an initial manual annotation of lung CT scans by a radiologist, which is later confirmed by two more radiologists individually. The radiologists try to locate the cancerous lung nodules, then classify the nodules based on the texture of the nodules in the CT images. Lung nodules are typically very small in size, approximately 2 to 3 mm in diameter. Hence, the location and classification of lung nodules becomes a daunting task for radiologists. Therefore, there exists a clinical necessity for an automated algorithm for lung nodule classification.

This paper aims to compare various texture analysis techniques for the multi-class classification of lung nodules. The discrimination ability of the extracted features is evaluated using a correlation test. The extracted features are used to train three classifiers to classify the lung nodules into solid, semi-solid, and Ground Glass Object. The performance analysis helps to identify the best texture features for an accurate diagnosis of lung cancer.
This paper is organized into the following sections. The second section describes the previous work done on this research problem. The third section elucidates the proposed methodology. The fourth section discusses the results obtained by the trained classifiers, and the final section concludes the work.

2. Related Work
Automated lung cancer detection has seen growing attention in recent years. However, the early works done in automated classification dates back to the early 2010s. Al Tarawneh et al. proposed Gaussian filters for the detection of lung nodules [2]. The nodules were also segmented using the same filter. Binarization was used as the sole feature extraction technique [3]. The proposed methodology displays an accuracy of ~70%. In addition, their work was done on X-Rays instead of the clinically used CT scans. Hence, their technique cannot be used effectively in a clinical setting.

Sharma et al. have used various image processing methods such as morphological erosion, median filtering, and dilation for extracting the region-of-interest (ROI) [4]. They have used SIFT blob detection for segmentation of the lung nodules. Furthering this, they have classified the lung nodules manually [5].

Gajdhane et al. have used eccentricity as a measure for cancerous nodule detection [6]. Eccentricity is used to determine how circular an object is. They worked under the assumption that a cancerous nodule is more circular in shape than non-cancerous nodules. The main objective of the work was to detect the stage of cancer progression from only the CT scans. Once they had extracted the nodule, they used feature extraction techniques such as wavelets. They used support vector machines for classification into various stages.

More recently, deep learning approaches have dominated the field of feature extraction and classification. Convolutional neural networks (CNN) have been widely employed for these tasks. Wafaa et al. used 3D CNN for feature extraction and classification of chest cavity CT scans [7]. A U-net was employed for the 3D segmentation of nodules [8]. The nodules were cropped and fed into the Convolution Neural Network (CNN) for the final output. Their method showed an accuracy of 86%, which is not very high, considering the complexity of the algorithm developed.

The best-performing methods for lung cancer classification on the LNDb dataset include works done by Atwal and Phoulady (2020), Chen et al. (2020), Galdran and Bouchachia (2020), Kaluva et al. (2020), Rasadin (2020), Sun et al. (2020) and Look (2020) [9]. All these related works have used variants of 3D convolutional neural networks. These constitute state-of-art methods, and our proposed methodology is compared against these works.

Our proposed method aims at comparing various texture analysis techniques based on which features are extracted. Various classification algorithms, namely Logistic Regression, K-Nearest Neighbours (KNN), and Support Vector Machines (SVM), are used for multi-class classification of lung nodules [10-12].

3. Methodology
The proposed methodology is subdivided into the following stages: manual nodule localization, automated feature extraction, feature selection, and classification. An overview of this methodology is given in figure 1.
3.1 Dataset
In this paper, the dataset used is the LNDb challenge dataset, which is publicly available for the registered teams of the challenge [13,9]. This dataset contains 294 CT scans of the lung cavity, out of which 58 CT scans were used as test data. The CT scans were read by radiologists to identify pulmonary nodules. The radiologists annotated the nodule lesions based on their size. The nodules were segmented and subjectively characterized based on the internal structure, texture, sphericity, calcification, etc. The location of the nodule centroids, segmentation masks, and the texture class to which each nodule belongs was available as ground truth labelled data. Figure 2 shows some examples of CT scans from the LNDb dataset.

Figure 1. Overview of the proposed approach.

Figure 2. Examples of input CT scans of the LNDb dataset.
3.2 Nodule Localization
In this work, a semi-automated approach is developed. The nodules' centroid locations and segmentation masks available with the ground truth data were used to find the nodule regions-of-interest. Features were extracted automatically for texture characterization of the lung nodules.

A cube of 40mm x 40mm x 40mm was cropped around the nodule centroids in the CT scan and its corresponding segmentation mask. Figure 3 shows the input CT scan and the corresponding segmentation mask for nodule localization. Figure 4 shows the localized nodule from different views of the CT scan.

![Figure 3. CT scan of lungs and nodule segmentation mask.](attachment:image1.png)

![Figure 4. View of the cropped nodule region-of-interest from different axes.](attachment:image2.png)
3.3 Feature Extraction
Various feature extraction techniques such as Gray Level Co-Occurrence Matrix (GLCM), Gabor filters, and Local Binary Patterns (LBP) were employed to determine multiple texture features from the nodule regions-of-interest [14-16].

3.3.1 Gray Level Co-occurrence Matrix. GLCM is a statistical approach for texture analysis. It considers the frequency of pairs of pixels occurring in a given spatial direction in an image.

GLCM feature extraction was set up to extract features from 0°, 45°, 90°, and 135° with four gray levels. The features obtained along the four directions were averaged while keeping in mind the symmetric nature of the data. Statistical features, namely contrast, energy, dissimilarity, homogeneity, correlation, and ASM, are extracted from the matrix.

3.3.2 Gabor filters with GLCM. Gabor filters are linear filters used for texture analysis. Gabor filters analyse a localized region for any specific frequency content in a fixed direction. Over the years, Gabor filters have been widely used for texture analysis as the filter representation is similar to the optical system of humans.

Gabor filters with a frequency of seven pixels were initially applied on the cropped volume, with 65 gray levels. GLCM feature extraction method was performed on the Gabor-filtered images, and statistical features, namely contrast, energy, dissimilarity, homogeneity, correlation, and ASM, were extracted.

3.3.3 Local Binary Patterns. LBP is a texture feature extractor that assigns a binary value to the image pixels by thresholding the pixels in the adjacent neighborhood. The operator is very robust to changes caused by differences in lighting or augmentation.

LBP operators with 15 circularly symmetric neighbour set points and a radius of seven pixels were used. The output of the operator are histogram graphs containing spatial information about the image. Various features such as mean, variance, and standard deviation were extracted from these histograms.

3.4 Feature selection
A correlation test was used to determine the degree of the relation of each feature to the classes and thus select the best features for training classifiers.

3.5 Classification
The target variable, texture, was provided as the ground truth label for texture characterization of the lung nodules. The texture had values ranging from 0 to 5, where 0 indicates a non-nodule. If the texture ranges between 0 and 1.3, the nodule is labelled semi-solid (class 1). If the texture lies in the range of 1.3 to 3.8, the nodule is labelled as solid (class 2). For values of texture greater than 3.8, the nodule is labelled as ground glass object (class 3).

The dataset was split into 80% as training set and 20% as the test set. Three classifiers, namely SVM, KNN, and logistic regression classifier, were used to perform predictive analysis on the training data.
4. Results

4.1 Evaluation metrics

Accuracy and F1-score are the metrics used to determine the efficacy of the proposed methodology. Accuracy is defined as the fraction of correctly predicted data to the total data.

\[ \text{Precision} = \frac{\text{True Positives}}{(\text{True Positives}) + (\text{False Positives})} \] (1)

\[ \text{Recall} = \frac{\text{True Positives}}{(\text{True Positives}) + (\text{False Negatives})} \] (2)

\[ F1 \text{ score} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision}) + (\text{Recall})} \] (3)

4.2 Correlation analysis

Correlation test was used to select the best set of texture features for training the classifiers. For the GLCM feature extraction method, GLCM matrix is determined from the original raw images of the nodule regions-of-interest, and statistical features are extracted. The correlation values of the extracted features with the nodule texture classes are shown in Table 1.

| Table 1. Correlation of GLCM features with texture classes. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|--------------------|
| Contrast        | Energy          | Dissimilarity   | Homogeneity     | Correlation     | ASM             | Nodule             |
| 1.00            | -0.92           | 1.00            | -1.00           | 0.47            | -0.93           | 0.54               |
| -0.92           | 1.00            | -0.92           | 0.92            | -0.37           | 0.99            | -0.33              |
| 1.00            | -0.92           | 1.00            | -1.00           | 0.47            | -0.93           | 0.54               |
| -1.00           | 0.92            | -1.00           | 1.00            | -0.47           | 0.93            | -0.54              |
| 0.47            | -0.37           | 0.47            | -0.47           | 1.00            | -0.39           | 0.62               |
| -0.93           | 0.99            | -0.93           | 0.93            | -0.39           | 1.00            | -0.36              |

Table 1 shows that the GLCM features, namely contrast, dissimilarity, homogeneity, and correlation features, correlate well with the nodule texture classes.

For the Gabor filtered GLCM method, the images are filtered using Gabor filter, and then the GLCM matrix is calculated from the filtered images. The correlation values of the extracted features with the nodule texture classes are shown in Table 2.

| Table 2. Correlation of Gabor filtered GLCM features with texture classes. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|--------------------|
| Contrast        | Energy          | Dissimilarity   | Homogeneity     | Correlation     | ASM             | Nodule             |
| 1.00            | -0.92           | 1.00            | -1.00           | 0.47            | -0.93           | 0.54               |
| -0.92           | 1.00            | -0.92           | 0.92            | -0.37           | 0.99            | -0.33              |
| 1.00            | -0.92           | 1.00            | -1.00           | 0.47            | -0.93           | 0.54               |
| -1.00           | 0.92            | -1.00           | 1.00            | -0.47           | 0.93            | -0.54              |
| 0.47            | -0.37           | 0.47            | -0.47           | 1.00            | -0.39           | 0.62               |
| -0.93           | 0.99            | -0.93           | 0.93            | -0.39           | 1.00            | -0.36              |

On comparison of tables 1 and 2, it is observed that the GLCM features extracted from Gabor filtered images show a higher correlation as compared to the GLCM features extracted from original images. Therefore, we chose Gabor filtered GLCM features to train the classifiers. The features with the highest correlation values were selected as the final feature set, namely contrast, energy, homogeneity, and correlation.
Table 2. Correlation of Gabor filtered GLCM features with texture classes.

|        | Contrast | Dissimilarity | Energy | Homogeneity | Correlation | ASM  | Nodule |
|--------|----------|---------------|--------|-------------|-------------|------|--------|
| Contrast | 1.00     | -0.93         | 1.00   | -1.00       | 0.53        | -0.94| 0.58   |
| Dissimilarity | -0.93    | 1.00          | -0.93  | 0.93        | -0.45       | 0.99 | -0.41  |
| Energy   | 1.00     | -0.93         | 1.00   | -1.00       | 0.53        | -0.94| 0.58   |
| Homogeneity | -1.00    | 0.93          | -1.00  | 1.00        | -0.53       | 0.94 | -0.58  |
| Correlation | 0.53     | -0.45         | 0.53   | -0.53       | 1.00        | -0.47| 1.00   |
| ASM      | -0.94    | 0.99          | -0.94  | 0.94        | -0.47       | 1.00 | -0.44  |

For the LBP histogram method, the three features extracted, namely histogram mean, variance, and standard deviation were used to train classifiers. Table 3 gives the correlation values of the LBP histogram features with the nodule texture classes.

Table 3. Correlation of LBP histogram features with texture classes.

|        | Mean       | Standard deviation | Variance | Nodule |
|--------|------------|--------------------|----------|--------|
| Mean   | 1.00       | -0.83              | -0.93    | -0.35  |
| Standard deviation | -0.83    | 1.00               | 0.93     | 0.64   |
| Variance | -0.93     | 0.93               | 1.00     | 0.40   |

4.3 Classification results

The proposed methodology involves three machine learning algorithms performed on the features extracted to classify lung nodules into three classes based on their texture. The three classes are semi-solid (class 1), solid (class 2), and ground glass object (class 3). Classifiers, namely logistic regression, KNN, and SVM, were trained on 1106 slices of CT scans and tested on 277 slices of CT scans. The results obtained for the test data are summarised in Table 4.

Table 4. Results of the classifiers for the test data.

| Feature extraction method | Classifier | Accuracy | F1-score |
|---------------------------|------------|----------|----------|
| Gabor filtered GLCM       | Logistic Regression | 0.91     | 0.89     |
|                           | KNN        | 0.91     | 0.89     |
|                           | SVM        | 0.86     | 0.84     |
|                           | Logistic Regression | 0.91     | 0.89     |
|                           | KNN        | **0.94** | **0.92** |
|                           | SVM        | 0.91     | 0.89     |

As shown in Table 4, the combination of Local Binary Patterns for feature extraction and K-Nearest Neighbours as the multi-class classification algorithm provided the best result for the evaluation metrics, i.e., accuracy of 94% and F1-score of 0.92. The confusion matrix for this is given below.

\[
C(i,j) = \begin{bmatrix}
85 & 0 & 6 \\
0 & 5 & 6 \\
1 & 0 & 174
\end{bmatrix}
\]

\( C(i,j) \) is a 3 x 3 confusion matrix since there are three classes for classification. \( C(i,j) \), when \( i = j \), represents the number of correctly identified samples where \( i \) represents the class to which it belongs.
C(1,3) shows that six samples that belong to class 1 have been incorrectly identified as class 3. C(2,3) shows that six samples that belong to class 2 have been incorrectly identified as class 3. C(3,1) shows that one sample belonging to class 3 has been incorrectly identified as class 1.

The better performance of the LBP feature extraction method can be attributed to the fact that the algorithm builds several local descriptors and combines all of them to form a global descriptor. This helps in preventing the loss of spatial information during feature extraction. Other feature extraction methods employed here work on a global scope due to which the spatial information is readily lost. As a result of this, efficient pattern recognition would fail due to sub-par feature extraction.

### 4.4 Comparison with previous works

The best performing method, i.e., Local Binary Patterns for feature extraction and KNN for multi-class classification, was compared against the results of the challenge teams who participated in the LNDb challenge [9]. The results are summarised in Table 5. The machine learning approach developed in this paper shows better results for the F1-score as compared to related works. It must be noted that the proposed work is a semi-automated approach and uses the segmented regions of the nodules provided in the dataset. In contrast, the challenge teams developed an automated method for the segmentation and classification of nodules.

| Related work                        | F1-score |
|-------------------------------------|----------|
| Atwal and Phoulady (2020)           | 0.61     |
| Chen et al., (2020)                 | 0.85     |
| Galdram and Bouchachia (2020)       | 0.86     |
| Rassadin (2020)                     | 0.72     |
| Sun et al., (2020)                  | 0.87     |
| Look et al., (2020)                 | 0.87     |
| **Proposed method**                 | **0.92** |

A limitation of our work is that the classifier cannot satisfactorily differentiate between class 2 (solid) and class 3 (ground glass object) nodules. This is apparent in the confusion matrix given in equation (4). The reason for this is fuzzy differentiation boundaries between solid and GGO nodules. One way to overcome this limitation and improve the overall model performance would be to use convolutional neural networks for feature extraction. The classification can be achieved using the CNN or machine learning classifiers as developed in the proposed work.

### 5. Conclusion

In this work, we present a comparison of various feature extraction and machine learning algorithms for texture analysis and classification of lung nodules. The best performing algorithm used a combination of Local Binary Patterns and K-Nearest Neighbours for feature extraction and classification. This proposed approach accrues an accuracy of 94% and an F1-score of 0.92. Our proposed methodology outperforms the state-of-the-art methods in the classification of lung nodules. Although our proposed methodology is semi-automated, the evaluation results point towards a promising approach that can be explored further using a larger dataset. Hence, it has the potential to be used in a clinical setting for fast and accurate lung nodule texture classification.
References

[1] Bastarrika G, García-Velloso MJ, Lozano MD, Montes U, Torre W, Spiteri N, Campo A, Seijo L, Alcaide AB, Pueyo J and Cano D 2005 Early lung cancer detection using spiral computed tomography and positron emission tomography American J. Respiratory And Critical Care Medicine 171 pp 1378-83

[2] Al-Tarawneh MS 2012 Lung cancer detection using image processing techniques Leonardo Electronic J. Practices and Technologies 11 pp 147-58

[3] Sauvola J and Pietikäinen M 2000 Adaptive document image binarization Pattern Recognition 33 pp 225-36

[4] Sharma D, Jindal G 2011 Identifying lung cancer using image processing techniques Proc. Int. Conf. on Computational Techniques and Artificial Intelligence (ICCTAI) vol 17 pp 872-80

[5] Lowe G 2004 Sift-the scale invariant feature transform Int. J. 2 p 2

[6] Gajdhane VA and Deshpande LM 2014 Detection of lung cancer stages on CT scan images by using various image processing techniques IOSR J. Computer Engineering 16 pp 28-35

[7] Alakwaa W, Nassef M and Badr A 2017 Lung cancer detection and classification with 3D convolutional neural network (3D-CNN) Lung Cancer 8 p 409

[8] Ronneberger O, Fischer P, and Brox T 2015 U-net: convolutional networks for biomedical image segmentation Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention (Springer, Cham) pp 234-41

[9] Pedrosa J, Aresta G, Ferreira C, Atwal G, Phoulady HA, Chen X, Chen R, Li J, Wang L, Galdran A and Bouchachia H 2021 LNDb challenge on automatic lung cancer patient management Medical Image Analysis 70 p 102027

[10] Suthaharan S 2016 Support vector machine Machine Learning Models and Algorithms for Big Data Classification (Integrated Series in Information Systems vol 36) (Springer, Boston, MA) pp 207-35

[11] Peterson LE 2009 K-nearest neighbor Scholarpedia 4 p 1883

[12] Kleinbaum DG, Dietz K, Gail M, Klein M and Klein M 2002 Logistic regression (New York: Springer-Verlag) p 536

[13] Pedrosa J, Aresta G, Ferreira C, Rodrigues M, Leitão P, Carvalho AS, Rebelo J, Negrao E, Ramos I, Cunha A and Campilho A 2019 LNDb: a lung nodule database on computed tomography arXiv preprint arXiv:1911.08434

[14] Pathak B and Barooah D 2013 Texture analysis based on the gray-level co-occurrence matrix considering possible orientations Int. J. Advanced Research in Electrical, Electronics and Instrumentation Engineering 2 pp 4206-12

[15] Biancon F and Fernández A 2007 Evaluation of the effects of Gabor filter parameters on texture classification Pattern recognition 40 pp 3325-35

[16] Guo Z, Zhang L and Zhang D 2010 A completed modeling of local binary pattern operator for texture classification IEEE Transactions on Image Processing 19 pp 1657-63