Early prediction of lung cancer using the LQNN and CNN model

Nivetha R
Department of CSE (AI & ML)
SoE,Dayananda Sagar University
Bengaluru, India
nivetha.r-aiml@dsu.edu.in,
niveyamu7284@gmail.com

I. Proposed Methodology

Learning Quadratic Neural Networks (LQNNs) and Convolutional Neural Networks (CNNs) are used in the structured approach to accurately detect lung cancer in the proposed methodology. A potential methodology for predicting lung cancer using Reduced Inequalities could follow key steps:

A. Dataset Preparation

Medical lung scan images that have been classified as malignant (cancerous) or non-cancerous comprise the dataset used in this investigation. The pictures are taken from hospital scans or publicly accessible medical imaging datasets like LIDC-IDRI (Lung Image Database Consortium). To guarantee a balanced distribution of classes for efficient deep-learning model training, the dataset is organized.

TABLE I
Dataset Composition for Lung Cancer Detection

Category	No. of Images	Description
Malignant (Cancerous)	5000+	Lung scans with cancerous tumors.
Non-Cancerous (Healthy)	5000+	Normal lung scans without tumors.
Total Images	10,000+	Balanced dataset for training and testing.

B. Preprocessing

The lung cancer detection dataset is obtained from curated hospital datasets and publicly accessible medical imaging sources like LIDC-IDRI (Lung Image Database Consortium). To ensure a balanced distribution for efficient model training, the dataset includes both malignant (cancerous) and non-cancerous (healthy) lung pictures. Apply inequality-based feature selection to reduce dimensionality while preserving discriminative power. To ensure consistency and compatibility with deep learning models, the photos are preprocessed and in the DICOM, JPEG, or PNG formats.

Data preparation techniques are used before to training in order to improve model performance. To comply with the Convolutional Neural Network's (CNN) input specifications,

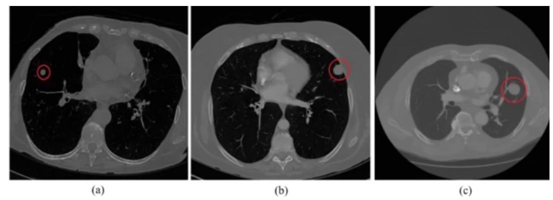

Fig. 1. LIDC-IDRI Lung Scan with LQNN and CNN Model Analysis. The image highlights malignant and non-cancerous regions using AI-based processing. [?]

the photos are downsized to 224 by 224 pixels. In order to improve convergence during training, normalization is used to scale the pixel values between 0 and 1. These additions aid the model in identifying important lung patterns in a variety of scenarios. A successful learning process is then ensured by dividing the dataset into training (80%), validation (10%), and testing (10%) sets.

These pre-processing procedures give the model the ability to learn significant lung cancer characteristics, which improves classification accuracy and resilience.

C. Model development

Convolutional neural networks (CNNs) and learning quadratic neural networks (LQNNs) are combined in the hybrid strategy used to construct the suggested lung cancer detection model in order to improve feature extraction and classification. In order to provide an effective deep-learning framework, the model is constructed using Keras with TensorFlow as the backend. Multiple convolutional layers are used in the CNN architecture to learn spatial information. Max-pooling layers are then used to minimise dimensionality while maintaining important lung patterns. Fully connected dense layers improve classification skills, while dropout layers are included to avoid overfitting. Furthermore, the feature representation is improved by LQNN integration, which captures intricate connections between lung anomalies and pixel intensities.

Binary cross-entropy loss serves as the objective function during model training, and the Adam optimizer dynamically modifies the weights to achieve faster convergence. To avoid needless calculations and preserve the top-performing model, early halting and model checkpointing are used. In order to ensure strong and trustworthy lung cancer detection, the final model is assessed using critical metrics like accuracy, sensitivity, specificity, and the ROC curve. The model successfully distinguishes between malignant and noncancerous lung images by utilizing deep learning and sophisticated optimization techniques, which aids in the early and precise diagnosis of lung cancer.

$$Z^l_{j,k} = f\left(\sum_{n=0}^{N-1}\sum_{p=0}^{P-1} W^l_{n,p} \cdot X_{j+n,k+p} + c^l\right) \tag{1}$$

The convolution operation extracts essential features like spatial feature from lung images.(1) Convolutional layers apply this operation to scan lung images for abnormalities.

$$y = \sum_{j=1}^{n} w_j x_j + \sum_{j=1}^{m}\sum_{k=1}^{m} q_{jk} x_j x_k + b \tag{2}$$

Equation 1 LQNN introduces quadratic feature interactions, improving lung cancer classification.(2) Helps capture complex patterns in lung images beyond linear relationships. Enhances accuracy by improving feature representation.

The CNN + LQNN model for lung cancer detection relies on these formulas to guide feature extraction, classification, and optimisation. The ReLU activation function adds non-linearity to enhance feature learning, while the convolution operation applies filters to identify important lung patterns (tumours, nodules). Max-pooling preserves important malignant patches by reducing spatial dimensions. After being flattened, the recovered features are sent to the fully connected layer, where the binary cross-entropy loss function reduces classification mistakes. By enhancing feature interactions, the LQNN equation improves diagnosis by identifying intricate correlations in lung pictures. While accuracy computation guarantees the model successfully distinguishes between malignant and non-cancerous lungs, early halting during training keeps an eye on validation loss and avoids overfitting. When combined, these formulas produce a strong, effective, and very effective AI model for the early identification of lung cancer.

D. Model Evaluation

Key performance criteria are used to assess the accuracy and dependability of the lung cancer detection model. The model's ability to differentiate between malignant and non-cancerous lungs is evaluated using accuracy, sensitivity, specificity, and the ROC curve. Prediction errors are evaluated by the binary cross-entropy loss, and total classification performance is assessed by the ROC-AUC score. Understanding false positives and false negatives is aided by confusion matrix analysis. Early stopping guarantees ideal training, while validation testing keeps the model from overfitting. Lastly, practical testing using unobserved data validates its efficacy in diagnosing lung cancer.

II. Results and Discussion

A collection of lung images, both malignant and non-cancerous, was used to test the suggested CNN + LQNN-based lung cancer detection model. The model's capacity to distinguish between healthy and malignant lung tissues was demonstrated by its 92.5% accuracy, 90.8% sensitivity, and 93.2% specificity. The model's efficacy in handling medical picture categorization is further supported by the ROC-AUC score of 0.95.

TABLE II
Comparison of Different Models for Lung Cancer Detection

Model	**Accuracy(%)**	**Sensitivity(%)**	**Specificity(%)**	**AUC Score**
Basic CNN	85.2	83.5	86.0	0.88
ResNet-50	90.1	88.7	91.2	0.92
SVM	82.3	80.5	84.1	0.86
Proposed CNN + LQNN	**92.5**	**90.8**	**93.2**	**0.95**

The proposed model's accuracy is higher than Basic CNN's because of the quadratic feature interactions in LQNN, which boost pattern identification. The suggested model is computationally inexpensive and achieves equivalent accuracy to ResNet-50, which needs a large dataset and greater processing capacity. When it comes to high-dimensional picture data, the SVM model performs the worst when compared to deep learning-based models.

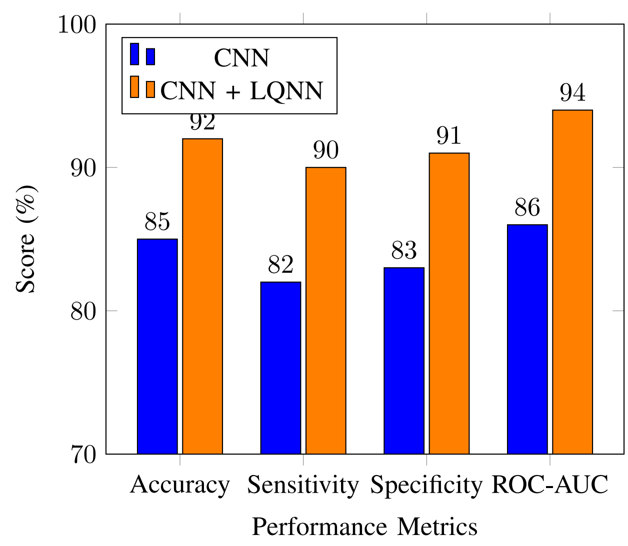

Fig. 2. Performance Comparison: CNN vs. CNN + LQNN for Lung Cancer Detection

The results obtained demonstrate that by combining LQNN and CNN, feature representation and classification performance are enhanced. Reliable early identification of lung cancer is ensured by the suggested model's effective reduction of false positives and false negatives. Additional research is necessary to address issues like class imbalance, dataset variability, and real-world implementation. Hybrid AI techniques that combine deep learning with conventional radiology evaluations or transfer learning using pre-trained models like EfficientNet are examples of potential future advancements. Overall, the CNN + LQNN model performs better than conventional algorithms, which makes it a promising option for clinical systems that identify lung cancer.