INTEGRATING DOMAIN KNOWLEDGE INTO DEEP NETWORKS FOR LUNG ULTRASOUND WITH APPLICATIONS TO LUNG CANCER

Guide:
Mrs. P. Deepika
By :
A. CHARAN (20R01A0501)
AR. CHINMAYI (20R01A0502)
B.NAGA VARSHA (20R01A0512)

INTRODUCTION

1.Our project, titled " **Integrating domain knowledge into Deep Networks for lung ultrasound with applications to Lung cancer**," aims to detect lung cancer at an early stage using ultrasound images.

2. Artificial Intelligence (AI) is increasingly transforming the medical field, offering promising avenues for diagnosis and patient care.

3. The project's primary objective is to leverage AI algorithms SVM and KNN to enhance the early detection of lung cancer through the analysis of medical imaging data.

4. By providing healthcare professionals with a powerful diagnostic tool, the project aims to significantly improve patient outcomes, underscoring AI's potential in modern healthcare.

ABSTRACT

- Our project, titled " **Integrating domain knowledge into Deep Networks for lung ultrasound with applications to Lung cancer**,“ focuses on utilizing machine learning algorithms to predict lung cancer based on patient data, with a structured workflow that includes data preparation, algorithm execution, and accuracy assessment.
- **Data Preparation:** We begin by uploading a lung cancer dataset, which serves as the foundation of our analysis.
- **Data Split:** The dataset is then split into training and testing sets to facilitate algorithm evaluation.
- **Algorithm Execution:** We execute Support Vector Machine (SVM) and k-Nearest Neighbors (KNN) algorithms on the data.
- **Lung Cancer Prediction:** Using these algorithms, we predict the likelihood of lung cancer in patients based on their data.
- **Accuracy Assessment:** We visualize the accuracy of these algorithms through graphical representations, aiding in the assessment of their performance.

PROBLEM STATEMENT

Despite significant advances in medical technology, early detection of lung cancer remains a critical challenge. Existing methods often lack the sensitivity and accuracy required for timely diagnosis. Current approaches often fall short in providing precise and reliable predictions, potentially leading to delayed diagnoses and compromised patient outcomes. In light of these shortcomings, our project aims to address this pressing issue by employing advanced machine learning algorithms to significantly enhance the accuracy of lung cancer prediction, ultimately contributing to more effective early detection and treatment.

EXISTING SYSTEM

The existing system for detecting early-stage lung cancer in CT scan images undergoes the following steps:

1.Image Preparation: First, the CT scan image is processed to improve its quality.
2.Lung Area Isolation: Then, the region of interest (ROI), which is the part containing the lungs, is separated from the rest of the image.
3.Image Compression: They use a method called Discrete Waveform Transform to compress the image, which helps in reducing its size while keeping important information.
4.Feature Extraction: Next, they extract features (like patterns or textures) from the image using a technique called GLCM (Gray-Level Co-occurrence Matrix).
5.Classification: Finally, these extracted features are used as input to a classifier called SVM (Support Vector Machine) to decide if the lung image shows signs of cancer.

DISADVANTAGES

The existing system has some drawbacks:

1.Complexity: The process involves multiple steps, making it somewhat complex.

2.Information Loss: Image compression may result in some loss of information.

3.Accuracy: The existing approach often falls short in providing precise ,accurate and reliable predictions, potentially leading to delayed diagnoses and compromised patient outcomes

These disadvantages suggest the need for improvements in terms of accuracy and simplicity in detecting lung cancer from CT scan images.

PROPOSED SYSTEM

- The proposed system for detecting early-stage lung cancer in CT scan images introduces several improvements compared to the existing system:

1. **Dataset Upload:** The project initiates with the uploading of a lung cancer dataset, which serves as the foundational data source for the analysis.
2. **Pre-processing and ROI Separation:** After dataset upload, the CT scan images undergo pre-processing, ensuring they are optimized for subsequent analysis. The region of interest (ROI), specifically the lung area, is isolated during this stage.
3. **Segmentation with DWT:** The Discrete Wavelet Transform (DWT) is then applied to perform image segmentation. This step helps in distinguishing different parts of the lung more accurately.
4. **Feature Extraction with GLCM:** Feature extraction follows segmentation, utilizing the Gray Level Co-occurrence Matrix (GLCM).
5. **Classification Using SVM and KNN:** Finally, the extracted features are employed as input for a Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) classifiers and the lung nodules are analyzed whether they are cancerous or not.

REQUIREMENTS

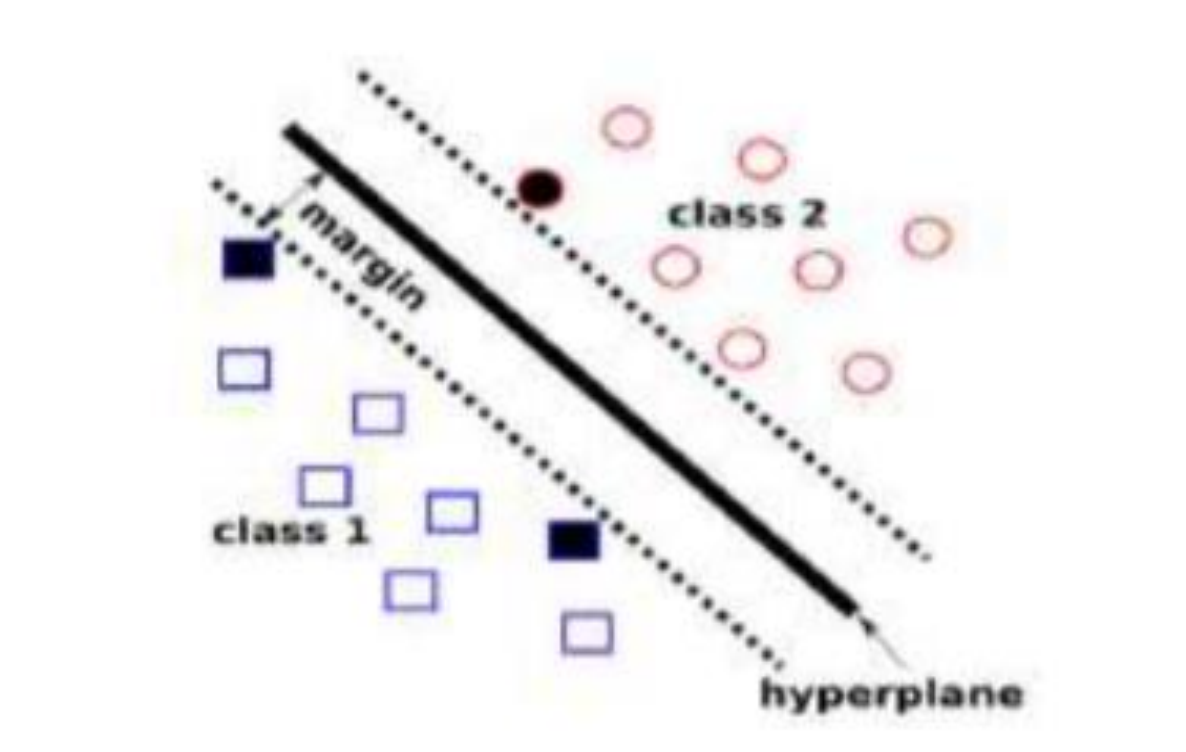
HARDWARE REQUIREMENTS:

- System : Pentium IV 2.4 GHz.
- Hard Disk : 40 GB.
- Floppy Drive : 1.44 Mb.
- Ram : 512 Mb

SOFTWARE REQUIREMENTS:

- Operating System**:** Windows
- Coding Language: Python 3.7

ARCHITECTURE



MODULES USED:

- **Tensorflow**
- **Numpy**
- **Scikit-learn**
- **Pandas**
- **Matplotlib**

ALGORITHMS USED

K-Nearest Neighbors (KNN):

1.Classifies by asking neighbors for help.
2.Looks at close neighbors to decide a category.
3.Imagine borrowing ideas from nearby friends.

Support Vector Machine (SVM):

1. Draws a line to separate different groups and finds the best line to keep things apart.
2. Think of drawing a clear line between two teams.

THANK YOU