International Journal of Innovative Research in Science and Engineering
Vol. No.2, Issue 10, October 2016
www.ijirse.com
ISSN (O) 2454-9665
ISSN (P) 2454-0663

DESIGN OF COPD BIGDATA HEALTHCARE SYSTEM THROUGH MEDICAL IMAGE ANALYSIS USING IMAGE PROCESSING TECHNIQUES

B.Srinivas Raja[1], Ch.Chandrasekhar Reddy[2]

[1]Research Scholar, Department of ECE, Acharya Nagarjuna University,Guntur, Andhra Pradesh.

[2]Asst.Professor, Department of ECE, KKR & KSR

Institute of Technology and Sciences, Guntur, Andhra Pradesh

ABSTRACT

Chronic obstructive pulmonary disease (COPD) is a heterogeneous disease with various clinical presentations. The basic abnormality in all patients with COPD is airflow limitation. The objective of the work is to help the hospitals and medical experts at all levels of healthcare to diagnose and manage COPD in a scientific manner with analytics. The proposed work allows to design a healthcare system with the application of Bigdataanalytics, medical image analysis using image processing techniques and decrease the risk in COPD patients. Image processing techniques are used to detect and classify COPD as per the severity levels by effective feature extraction, feature selection through an algorithm and the results are classified by the implementation of an effective classifier.

The preprocessing techniques will remove the noises and the feature extraction are done to extract the useful features in the image and the feature selection technique will optimize features relevant for the image and the classifiers are employed to classify the image. Datasets are used to compare and diagnose the level of severity of the disease with relevance to lung parameters. Data-driven innovation approach on medical image analysis provides solutions to severe problems faced by COPD patients.

I. INTRODUCTION

Chronic Obstructive Pulmonary Disease (COPD), a common preventable and treatable disease, is characterized by persistent airflow limitation that is usually progressive and associated with an enhanced chronic inflammatory response in the airways and the lung to noxious particles or gases. Exacerbations and comorbidities contribute to the overall severity in individual patients.

The goals of COPD assessment are to determine the severity of the disease, including the severity of airflow limitation, the impact on the patient’s health status, and the risk of future events (such as exacerbations, hospital admissions, or death), in order to guide therapy [1].

ISSN (O) 2454-9665
ISSN (P) 2454-0663

CLASSIFICATION OFAIRFLOW LIMITATION IN COPD BASED ON FEV1		
GOLD 1	MILD	FEV1 >=80% PREDICTED
GOLD 2	MODERATE	50% <= FEV1 < 80% PREDICTED
GOLD 3	SEVERE	30% <= FEV1 < 50% PREDICTED
GOLD 4	VERY SEVERE	FEV1 <30% PREDICTED

Comorbidities occur frequently in COPD patients, including cardiovascular disease, skeletal muscle Dysfunction, metabolic syndrome, osteoporosis, depression, and lung cancer. Given that they can occur in patients with mild, moderate and severe airflow limitation and influence mortality and hospitalizations independently, comorbidities should be actively looked for, and treated appropriately if present.

Hence there is a need for patient diagnosis regarding lung parameters.the combination of medical imaging, image processing techniques and Big Data can help the medical experts to provide quick solutions to various complex problems.

1.1 Role of Health care

The complexity of health care makes it a perfect domain to explore the potential for prescriptive analytics and imaging. Health care has been a pioneer in capturing rich imaging information and built databases to develop a variety of statistical medical norms. The next step is to use this image analytics to provide real-time insight to healthcare providers during diagnosis and treatment. Medical decision support systems link knowledge bases to multiple clinical databases and they are in turn linked to a patient's data.

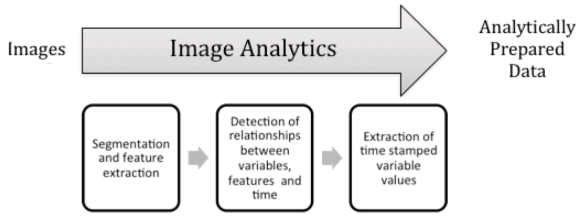
These complex systems have varying schemata, comparative image banks and discipline. Image analytics reduces varying subjective interpretation and experts error, thereby accelerating the process of treatment and recovery. The increasing role of algorithmic diagnosis and treatment creates the perfect opportunity to integrate images with prescriptive analytics.

1.2 Medical Image processing Techniques

Image analytics is the automatic algorithmic extraction and logical analysis of information found in image data using digital image processing techniques. The objective of image analytics is to bring an unstructured rendition of reality in the form of images and videos into a machine analyzable representation of a set of variables.

A variable is represented by a series of values related to an entity.image analytics continues to be a set of transformations on image-input that add value and create a rich set of time series as analytically prepared data output. The first transformation step segments images into structured elements and prepares them for feature extraction – i.e., the identification of low-level features in the image.

The second transformation step is the detection of relationships between these features, variables and time. The third transformation step is the extraction of variables with time-stamped values.



1.3 Segmentation and Feature Extraction

Images and videos are segmented using algorithms and digital processing techniques known as image segmentation. Segments are spatially relevant regions of image or video scenes that have a common set of features. These can be color distributions, intensity levels, texture, moving and other criteria. Image segmentation algorithms are used to extent of COPD. Feature extraction is next in the process. To assist in the detection of higher-level characteristics, low-level features are extracted and stored with each instance.

1.4 Relationships among Variables, Features and Time

To detect relationships between variables, features and time, an artificial intelligence sub-discipline known as Machine Learning is combined with Applied Statistics to create the relationship "intelligence" that is the core of the image analytics process. The relationships among variables, features and time in image analytics are represented as a predictive model. Before a predictive model can be created, a set of instances are extracted from all the given images and/or all the given videos being analyzed. A machine learning or statistical modeling algorithm trains a predictive model based on the set of annotated training instances. Modeling algorithms are based on known techniques including neural networks, scalable vector machines, function learning, Bayesian networks, regression and many more. Test instances are used to calculate the accuracy of a predictive model created by a modeling algorithm. The training process is often repeated with different sets of training and test instances and/or algorithm parameters until the accuracy of the predictive model is at an acceptable level. After the predictive model has been trained, it is used to classify predicted instances in a process.

INPUT SCANNED IMAGE

FEATURE EXTRACTION

FEATURE SELECTION

CLASSIFIER

Normal **COPD Lung**

II. PROPOSED SYSTEM

The proposed system has to establish the dataset of COPD patient network image database, drug network, disease network, Gene network.Big data is not just about size. It can find insights from complex, noisy, heterogeneous,longitudinal, and voluminous data. It aims to answer questions that were previously unanswered.

ELECTRONIC HEALTH RECORD (Patient data-EHR)–
Systematic collection of individual patients or population
ICD - International Classification of Diseases

PATIENT SIMILARITY ANALYTICS

PREDICTIVE MODELING

CLINICAL PATHWAY ANALYSIS

DISEASE PROGRESSION ANALYSIS

In patient similarity analytics, a big data approach is needed where all known attributes about patients are taken into consideration, in order to account for all potential confounding factors. This poses two challenges. First, since the number of attributes can be very large (e.g., in the order of tens of thousands), how to define distance, or similarity metric, in this high dimensional space is a challenging mathematical problem. Second, the notion of patient similarity is context dependent. Healthcare analytics research increasingly involves the construction of predictive models for disease targets across varying patient cohorts using observational data such as EHR.

Extracting insights from temporal event sequences, such as mining frequent patterns, is an important challenge in healthcare. However, despite the availability of temporal data and the common desire to extract knowledge, mining patterns from temporal event sequences is still a fundamental challenge in data mining.By featuring a novel frequent sequence mining algorithm to handle multiple levels-of-detail, temporal context, concurrency, and outcome analysis. Frequency also features a visual interface designed to support insights, and support exploration of patterns of the level-of- detail relevant to users.

Disease progression procedure of a disease with computational technologies, is one important technique that can help realize disease early detection. Key challenges in developing DPM methodologies include: (1) Multiple Covariates. The progression of disease usually involves the evolution of many different types of covariates. In general it is not know which one or which group of variables are important. (2) Progression Heterogeneity. The patient disease conditions can progress differently for different individuals, and the patient records are not necessarily aligned. (3) Incomplete Records. The patient records are not complete, meaning that in most of the cases we are not able to get the patient records from the beginning stage of the disease to its end stage. (4) Irregular Visits. The patient only has medical records when he/she pays visit to medical facilities. Most of the times patients visits are at irregular time stamps due to various reasons. (5) Discrete Observation. Although the disease progression is a continuous time procedure, the patient records are only observed on certain discrete time stamps or intervals. (6) Limited Supervision.

III. CONCLUSION

Healthcare has undergone a tremendous growth in the use of EHR systems to capture patient disease and treatment histories. advanced data driven analytics and visualization methodologies and systems need to be developed in order to convert the source material into meaningful insights. Bigdata in COPD reduces the number of unnecessary hospitalizations,improving the health of patients while decreasing the costs of care.The aim of the paper is to provide a powerful tool that can be leveraged by researchers everywhere to speed up the development of data-driven analytics that can lead to better deliver of care at lower cost. Efficient models for comprehensive risk assessment models to better predictive risk in a cross domain Environment, disease modeling,automatic disease progression model.

REFERENCES

[1] Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2016.link:http://www.goldcopd.org/guidelines-global-strategy-for-diagnosis-management.html.

[2] BIGDATA Analytics for healthcare Weblink: http://dmkd.cs.wayne.edu/TUTORIAL/Healthcare.

[3] Giardina,C.R. and E.R.Dougherty, Morphological Methods in Image and SignalProcessing", Englewood Cliffs, New Jersey: Prentice–Hall. 321, 1988.

[4] Uppaluri R, Hoffman EA, Sonka M, Hartley PG,Hunninghake GW, McLennan G,"Computer recognition of regional lung disease patterns" American Journal of Respiratory Critical Care Medicine 1999;160:pg:648–54.

[5] Manish Kakara, Dag Rune Olsen "Automatic segmentation and recognition of lungs and lesion from CT scans of thorax "IEEE transactions on Computerized Medical Imaging and Graphics 33 (2009),pg: 72–82.

[6] R. M. Summers, "Road maps for advancement of radiologic computer-aided detection in the 21st century," Radiology, vol.229, no. 1, pp. 11–13, 2003.

[7] A. Jain and D. Zongker, "Feature Selection: Evaluation Application, and Small Sample Performance", IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, Vol. 19, No. 2, pp. 153-158.