A Detection of Lung Cancer using Cat Swarm and BFO optimization Method

Vikas[1], Navjeet Kaur[2],
[1]*Doaba Institute of Engineering and Technology*
(E-mail: vr1574@gmail.com*)*

Abstract**—**Cancer is the leading life-threatening disease for people in today's world. Although cancer formation is different for each type of cancer, it has been determined by studies and research that stress also triggers cancer types. Lung cancer is very challenging to detect at its early stage with medical examinations. Early detection of lung cancer can help for the commendable decrease in the lung cancer mortality rate. In recent years, the image processing methodologies are widely used in medical image diagnosis, especially in detection of lung cancer tumors. One of the major reasons for non-accidental death is cancer. It has been proved that lung cancer is the topmost cause of cancer death in men and women worldwide. The death rate can be reduced if people go for early diagnosis so that suitable treatment can be administered by the clinicians within specified time. Cancer is, when a group of cells go irregular growth uncontrollably and lose balance to form malignant tumors which invades surrounding tissues. For any type of cancer, firstly image of internal parts of the body should be obtained. CT scan also known as X-ray computed tomography makes use of X-ray for capturing the images from various angles and merge these images to generate cross sectional tomographic image of particular areas of scanned tissues, i.e. it allows the person to see the status inside body without non-invasive techniques. Lung cancer is customarily a contagion which takes place because of the element linked with unimpeded cell or conveniently progress in zones present in lung area. This sort of evolution could possibly initiate metastasis that is also an occurrence linked with subsequent tissues and also access outer the lung region. Treatment along with diagnoses depend on histological sort of cancer malignancy degree, accompanied by the patient's efficiency. Some feasible treatment options comprises of surgical method, chemical therapy, along with radio therapy. In this the features are extracted using independent component analysis, feature optimization is done using firefly swarm optimization and classification is done using support vector machines. In the research work, the fuzzy technique has been used for the feature extractions of DICOM images and we have also optimized the best features for lung cancer detection using BFO and Cat Swarm Optimization. Further, the classification is performed using the SVM classifier and also computed the best survival rate and accuracy for the proposed research work.

Keywords**—***Cancer; Cat Swarm; BFO; Sensitivity*

I. INTRODUCTION

Cancer is still the major cause of death in the world, further more lung cancer is the most frequently seen type of cancer among others (WHO, 2015). As there is no cancer registry system in TRNC, there is not any official data about cancer statistics. Yet, lung cancer is the leading cancer in males in Turkey as it is in the world male rates (T.C. Saglik Bakanlıgı, 2016). Early diagnosis and proper treatment may pull down the death rates, hence the CAD systems are increasingly becoming the preferred aid in diagnostic procedures by the doctors (Doi, 2007). CAD becomes a significant research topic in the diagnostic radiology and medical imaging. In fact, CAD systems help the doctors in interpreting the images of computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, positron emission tomography (PET), conventional projection radiography as well as all other imaging methods. Practically, diagnosis process incorporates the assistance of computers from medical imagery, lab work, and electronic medical records and more. When it comes to radiology, CAD is the essential system of procedures in medicine that help doctors in the medical image interpretation. The use of the digital processing and hybrid optical technologies afford the reduction in a processing time as well as enabling more enhancements in specificity and sensitivity. The computer aided diagnosis holds the great potential for the radiology and its utilization is based on its capability to speed up a diagnostic process as well as lessen probable errors. The concept of the automated diagnosis exists from the year 1960, however the attempts in research and development were failed mostly (Doi, 2007). Now, there are many institutions all over the world that involved in the development and research of CAD aspects. Day by day, CAD systems are giving more confidence in the medical area therefore CAD systems become a superior method for the cancer detection in interpreting X-ray, CT, MRI and other medical images. Using the outputs of CAD systems as a reference helps the doctors not only to accomplish their tasks more accurately and precisely but also in a shorter time. The CAD systems ensure its reliability and efficiency to the integration of various scientific disciplines such as artificial intelligence, image processing, pattern recognition, etc. Although, CAD systems showed great improvement, it needs much to do in lung segmentation and in different shapes of nodule detection. CAD systems [2, 3] still have more false-positive results than experienced radiologist and have not achieved 100% accuracy, sensitivity and specificity which are very important measurements for the systems. This challenge is the motivation of this study in implementation of CAD system for lung cancer detection. The main purpose of the CAD system is to enhance a diagnostic accuracy as well as radiologist's image interpretation consistency with the help of computer output. This output is highly useful, since the radiologist's diagnosis is based on the subjective judgment. Generally, there are two general approaches that can be applied in the computerized schemes for computer aided diagnosis. First, to identify the lesions location like lung nodule in the chest image by looking isolated abnormal pattern with the computer. Then, the next thing is to measure the features of image of abnormal or/and normal pattern like lung texture concerned to the interstitial infiltrate in vessel sizes and chest image related to the angiograms stenotic lesions.

A. *What is Lung Cancer?*

The unrestrained expansion [5] of abnormal cells in lungs causes lung cancer. These abnormal cells disturb the smooth functioning and development of lung tissues. If this condition is left untreated, abnormal cells grow and form tumors and ultimately damage the lungs that are proving oxygen to the whole body via blood. Two main types of lung cancers are non-small cell and small cell. Because of the large size of lungs, nodules can grow for a time untills detecting them (Lung Cancer, 2016).

B. *Non-small cell lung cancer (NSCLC)*

The NSCLC is the commonly prevailing form of lung cancer, also according to American Cancer Society; NSCLC is responsible for 85 percent of the total lung cancers in America (American Cancer Society, 2016). The common tumors of lung cancers are following;

- Adenocarcinoma is the lung cancer in non-smokers, and equally found among men and women.
- Squamous cell carcinoma or sepidermoid carcinoma is the lung cancer that is positively correlated with the tobacco smokers. This tumor is formed mainly in at the center of large bronchi. Males are more vulnerable to this type of tumor.
- Large cell carcinomas are the tumor cells that have comparatively larger size with excess cytoplasm.

Unlike adenocarcinomas and epidermoid, these cells lack microscopic characteristics.

C. *Small cell lung cancer (SCLC)*

The remaining 15 percent contribution in lung cancers is of SCLC. Tobacco smoking is the leading cause of SCLC and gets birth quickly as compared to NSCLC. In the body, this type of cancer is relatively spread quickly, higher growth rate and shorter multiplying time. Chemotherapy is a more effective treatment for the SCLC [7].

D. *Screening for Lung Cancer*

The first step to diagnosing the lung cancer is the identification of symptoms. Symptoms are largely showing the damage to lungs and their functionality. Chest pain and cough are the most common symptoms of lung cancer. The cough gets worst on each passing day and also increases chest pain. Besides these, breath shortness, feeling weak, weight loss, blood in cough and fatigue are also commonly appeared symptoms among the lung cancer patients. Unfortunately, the scientific community has not developed any screening tool that could identify the lung cancer at early stage. Chest X-rays are commonly available tools for the screening, but they are not reliable enough yet. The development of a screening tool is the necessity of time as many researchers have concluded that early-stage tumors are easy to cure. The low-dose computed tomography (LDCT) is recommended screening on the annual basis to smokers and those who quit smoking with last 15 years. According to American Society of Clinical Oncologists, people who are in the age group of 55-74 and smoked more than 30 years are at more risk of lung cancer. In addition to LDCT, following are the some imaging technologies that are used for diagnosis and treatment [8].

E. *Chest x-ray*

X-ray machine discharges radiation that goes into the body and imaging picture of the organs on the film. To diagnose lung cancer, x-ray imaging is used as step that helps in the identification of lung tumors. As mentioned, x-rays are not the final authority because they are unable to differentiate between the cancer and other lung diseases.

Figure 1: Typical x-ray machine, (Frederick Memorial Hospital, 2016)

F. *CT scan*

CT scan [6] stands for computed tomography, and it is an extended version of X-ray in which computer is attached to the X-ray machine. Pictures that are taken from taken angles and distances are processed in the computer and presented in the 3-dimenionsal, cross-sectional (tomographic) and in slices form. In this way, bones, tissues, blood vessels, and organs are shown up clearly. The imaging of CT scan is useful for diagnosis, treatment and progress of medication. Recently, helical or multi-slice scanning is introduced that almost eliminated gaps in the collection of slides.

Figure 2: Typical CT machine, (Frederick Memorial Hospital, 2016)

G. MRI (magnetic resonance imaging)

It is imaging technique [9] in which radio waves and strong magnetic fields are used by the scanners to form the inside images of the body. The powerful magnet is aligned with the nuclei of atoms, and then magnetic field triggers atom to resonate. In this way, nuclei generate its own magnetic field and then the field is detected by the scanner for creating an image. The advancement in technology has helped to take detailed pictures from different angles. MRI is particularly helpful when there is a need for the identification of soft tissues. This is the reason; this technique has more reliability.

Figure 3: Typical MRI machine, (Frederick Memorial Hospital, 2016)

H. PET scan (positron emission tomography scan)

If X-ray or CT scan diagnoses or doctor predicts any chances of lung cancer, PET scan is suggested for detailed results. In this imaging technique, tracer or radioactive glucose is injected and then scanners are rotating to take pictures which tissues or organs used tracer (Mac Manus et al., 2003). When malignant tumor cells use glucose, they are showing up brighter and more active in images. The integration of PET-CT scan is very useful for detecting the cancer. The CT scan gives a detailed view of tissues and organs, and PET gives pictures of abnormal activities and active cells. Researchers also concluded that PETCT scan are producing more accurate results as compared to PET or CT scan alone.

Figure 4: Typical PET machine, (Frederick Memorial Hospital, 2016)

I. Digital Imaging and Communications in Medicine (DICOM)

Is a standard for handling, storing, printing, and transmitting information in medical imaging. It includes a file format definition and a network communications protocol. The communication protocol is an application protocol that uses TCP/IP to communicate between systems. DICOM files can be exchanged between two entities that are capable of receiving image and patient data in DICOM format. The National Electrical Manufacturers Association (NEMA) holds the copyright to this standard. It was developed by the DICOM Standards Committee, whose members are also partly members of NEMA.DICOM enables the integration of medical imaging devices – like scanners, servers, workstations, printers, network hardware, and picture archiving and communication systems (PACS) – from multiple manufacturers. The different devices come with DICOM Conformance Statements which clearly state which DICOM classes they support. DICOM has been widely adopted by hospitals and is making inroads in smaller applications like dentists' and doctors' offices [15].

J. Lung Cancer using Image processing

Lung cancer is a disease of abnormal cells multiplying and growing into a tumour. Cancer cells can be carried away from the lungs in blood, or lymph fluid that surrounds lung tissue. Lymph flows through lymphatic vessels, which drain into lymph nodes located in the lungs and in the center of the chest. Lung cancer often spreads toward the center of the chest because the natural flow of lymph out of the lungs is toward the center of the chest. Metastasis occurs when a cancer cell leaves the site where it began and moves into a lymph node or to another part of the body through the blood stream [1]. Cancer that starts in the lung is called primary lung cancer. There are several different types of lung cancer, and these are divided into two main groups: Small cell lung cancer and non-small cell lung cancer which has three subtypes: Carcinoma, Adenocarcinoma and Squamous cell carcinomas. The rank order of cancers for both males and females among Jordanians in 2008 indicated that there were 356 cases of lung cancer accounting for (7.7 %) of all newly diagnosed cancer cases in 2008. Lung cancer affected 297 (13.1 %) males and 59 (2.5%) females with a male to female ratio of 5:1 which Lung cancer ranked second among males and 10th among females [2]. Figure 1 shows a general description of lung cancer detection system that contains four basic stages. The first stage starts with taking a collection of CT images (normal and abnormal) from the available Database from IMBA Home (VIA-ELCAP Public Access) [3]. The second stage applies several techniques of image enhancement, to get best level of quality and clearness. The third stage applies image segmentation algorithms which play an effective rule in image processing stages, and the fourth stage obtains the general features from enhanced segmented image which gives indicators of normality or abnormality of images.

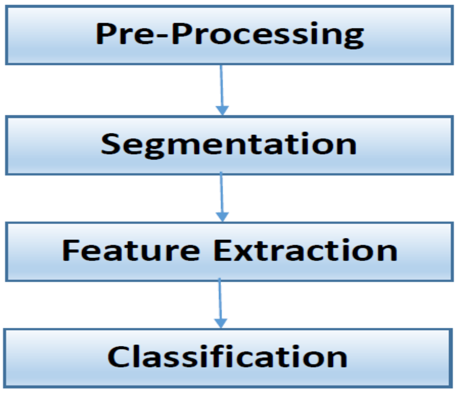

Figure 5: Lung cancer image processing stages

Lung cancer is the most dangerous and widespread cancer in the world according to stage of discovery of the cancer cells in the lungs, so the process early detection of the disease plays a very important and essential role to avoid serious advanced stages to reduce its percentage of distribution.

K. Lung cancer

Lung cancer is the type of cancer in which there is unchecked growth of unusual cells either in one or in both the lungs. These anomalous cells do not perform the functions of healthy human cells and do not mature into normal cells. This abnormality affects the proper regular functioning of the lung of supplying oxygen to the human body through blood. Though there are many advances in treatment procedures, the lung cancer which is at an advanced stage or late stage is not often easily curable. The field of lung cancer screening is a very challenging, well-paced and interdisciplinary. It is not free from controversies either. The perfect circumstance is to look the screening to permit it to develop as a general wellbeing technique.

L. Types of Lung Cancer

The lung cancers are usually classified into based on their appearance in micro-scope:

- Small Cell lung cancers
- Non-small cell lung cancers

The treatment of these two types of lung cancers is quite different as they have varied features for growth and spread. So it is extremely important to differentiate between Small cell lung cancers and Non-small cell lung cancers 1.

The Problem and Current challenges

Non-small cell lung cancer (NSCLC), the most widely recognized sort of lung tumor, is one of genuine ailments bringing on death for human beings. Computer-aided diagnosis and survival prediction of NSCLC is of great significance in diagnosis and treatment of people suffering from lung cancer [1]. The prognosis of lung malignancy is still poor, with year survival rate of roughly 10% in many nations. NSCLC accounts for the majority (84%) of lung cancer. Two major types of NSCL Care adenocarcinoma (including bronchi alveolar carcinoma) which is about 40% and squamous cell carcinoma about 25 - 30%.

Lung cancer imposes a major health care problem in many countries in the world. It is the most common solid cancer, with an estimated 1.6 million new cases annually worldwide (13% of total cancer diagnoses), and the leading cause of cancer death in western countries. In the Netherlands, 11,699 new cases of lung cancer were diagnosed in 2011, which reflects an increase of 30% since the beginning of the 21st century. A further increase is expected, mainly due to relative enlargement of the ageing population.

Smoking is the predominant cause of lung cancer. A lifetime smoker has a 20-30 fold increased risk of lung cancer compared to a lifetime non-smoker; of all smokers, 15% will ultimately develop lung cancer.6 Incidence numbers of lung cancer follow smoking trends in the population over time with a lag time of about 30 years. Tobacco consumption peaked by the middle of the 20th century, but declined at the end of the 20th century, which is now reflected in decreasing lung cancer incidence in males. However, as females started smoking several decades later than men, a 75% increase of lung cancer diagnoses among females was observed between 2000-2010. Quitting smoking results in risk reduction for lung cancer, but a former smoker will never reach baseline levels.

Although smoking is the major cause of lung cancer, in never-smokers (15% of lung cancers in males, 53% in females) lung cancer ranks the seventh cause of cancer death worldwide (before cervical cancer and prostate cancer). Gender and geographical variations may play a role as etiologic risk factor, as well as genetic predisposition, exposure to occupational carcinogens, and hormonal and environmental determinants.

M. Lung Cancer Classification: Histology and Genomics

Lung cancer compromises a heterogeneous spectrum of disease, pathologically classified according to the World Health Organization (WHO) standards. Recently, a modification of lung cancer classification was published, with emphasis on histological diagnosis and molecular testing for driver mutations, relevant in clinical decision-making and choice of treatment.

Histologically, lung cancer is divided broadly into non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). SCLC occurs less common (~15%), has a strong association with smoking (98%) and is characterized by an aggressive tumour behaviour.20 It responds initially well on chemotherapy, but has a high recurrence rate and overall has a similar dismal prognosis as NSCLC. NSCLC accounts for 85% of all lung cancers and 85% is related to smoking. Before molecular characterization of tumours became clinically relevant, NSCLC subtypes were lumped, because of similar natural history and therapy response to cytotoxic chemotherapy. Major subtypes are adenocarcinoma (AdC) and squamous cell carcinoma (SqCC); large cell carcinoma and miscellaneous other types (such as neuroendocrine tumours) are less frequently diagnosed. The diagnosis 'NSCLC not otherwise specified' should be avoided, as since the beginning of this century, the distinction between AdC and SqCC with specific mutational changes in subgroups provide markers for prediction of therapy response. AdCs are mostly peripherally located and arise from progenitor cells of the bronchioles (Clara cells), alveoli (Type II pneumocytes) or mucin-producing cells. It is the most common histologic type in never-smokers. A proportion of AdCs harbour specific mutations in the DNA of the cancer cells, which may play a key role in personalization of treatment. In this context, identification of genetic aberrations in EGFR and ALK genes are relevant for daily practice. For example, treatment with EGFR tyrosine kinase inhibitors (TKIs) as compared to classical chemotherapy has shown to increase fprogression-free survival in patients with lung cancer characterized by activating EGFR mutations. EGFR mutations are (in particular) associated with terminal respiratory unit type AdC, women, never-smokers and Asians and have shown to be mutually exclusive with various other mutations in primary lung cancer.

Some SqCC variants were not taken into account in the WHO classification. Traditionally, SqCCs were thought to arise centrally in the lung-airways, originating from metaplastic changes of the respiratory epithelium. However, peripheral SqCC have been described. Whether peripheral SqCC arises from metaplastic changes during peripheral AdC development or not, remains to be established. A comprehensive study was performed by the Cancer Genome Atlas Research Network, investigating aberrations on genomic level in squamous cell carcinomas: TP53 mutation was nearly universal and in 69% of tumours an altered PI3K/RTK/RAS signalling pathway was observed. In 7%, EGFR amplification was demonstrated. Promising therapeutic targets are FGFR1 amplification and DDR2 mutation. Interestingly, a molecular classification of lung cancer runs parallel with the conventional histologic classification.

Symptoms

The usual symptoms of squamous cell carcinoma 2 are coughing up blood, Wheezing (high pitched whistling sound occurring while a person breathing in or breathing out), constant cough. Squamous cell carcinoma is usually found in big airways. They therefore usually show symptoms at an early stage as compared to other forms of lung cancer. They usually obstruct the airways of the lungs causing infections like pneumonia and damage part of the lung. There is a syndrome called Pancoast syndrome. The syndrome starts at the beginning of the lungs and move on to other parts of the body adjacent to them, is mostly caused by Squamous cell carcinoma. People with squamous cell carcinoma are likewise more inclined to encounter a raised calcium level (Hypercalcemia) which can bring about weak muscles and other issues. Hypercalcemia is one of the side effects of paraneoplastic disorder, and is created by a tumor which secrets a hormone-like substance that brings the calcium level up in the blood.

N. Lung Cancer Diagnosis

The first step of detection of squamous cell carcinoma is through X-rays when there are any unusual abnormalities in the lungs. The other diagnostic methods are:

- CT Scan- The computed tomography of the chest aids pathologists visualizes the lungs and vessels inside them through non-invasive imaging techniques. The technique also involves injecting a dye known as contrast dye into the veins before scanning, so as to enable pathologists to clearly view the lungs.
- PET Scan-Positron Emission Tomography test is a radiology test which is commonly used alongside other diagnosis measures like CT scan.
- Bronchoscopy- A process in which large tube is inserted in the passage of mouth or through the nose in order to view the airways under medical supervision.

When a patient presents with symptoms, the diagnostic workup to establish lung cancer diagnosis involves different steps. Imaging techniques (chest X-ray (CXR), computed tomography (CT), positron emission scanning) are required to establish tumour location and assess tumour burden (staging, cTNM) and provides information on where to obtain tissue/cells for (molecular) pathological examination. Importantly, radiological imaging is non-specific for lung cancer. For confirmation of lung cancer diagnosis, additional invasive procedures as bronchoscopy with or without ultrasound/ CT guidance are performed to procure tumour material for histological/cytological/ molecular diagnosis. Limitations of these techniques are that not all patients are physically able to undergo these procedures and pulmonary AdCs are difficult to reach with bronchoscopy, as these are usually located in the periphery of the lung. In this case, other procedures such as transthoracic needle biopsy, mediastinoscopy, video-assisted thoracoscopic surgery (VATS), or even thoracotomy aid in procurement of tumour tissue.

Less or non-invasive tools to establish lung cancer diagnosis and/or to examine mutational status of the tumour may be the use of bronchoalveolar lavage and sputum. Sputum is an easily accessible biological specimen that can be obtained by simple and inexpensive means, and consists of a mixture of saliva and epithelial lining fluid. The thin layer on the surface epithelium of the respiratory tract has a volume of approximately 2 ml, and is composed of mucus secreted by goblet cells and proteins secreted by Clara cells and other cells.

The lining fluid is continuously moved by the ciliary system in proximal direction towards the larynx, and swallowed into the digestive tract. The airways are thus protected by the removal of inhaled and local debris. In general, this process occurs unnoticed. In the majority of smokers, hyperplasia of goblet cells in the respiratory epithelium leads to excess production of sputum. In order to prevent airway blockage, sputum is often expectorated. This is illustrated by the typical smoker's morning cough, to get rid of sputum that has accumulated during the night. Spontaneous sputum is usually easily obtained from smokers. For former and non-smokers it is more difficult to produce sputum, although this issue can be overcome by sputum induction. When the cells are collected in Saccomanno's fixative, the sputum can be stored at room temperature for years.

Sputum can be used for lung cancer diagnosis Sputum cytology reveals inflammatory and epithelial cells mostly from oropharynx. In lung cancer, less than 1% of sputum cells are exfoliated tumor cells. Sputum cytology has limited value for lung cancer diagnosis with a pooled sensitivity rate of 66% (95% confidence interval (95%CI): 42-97%), though with high specificity (pooled rate 99%, 95%CI: 68-100%) and is currently not part of routine diagnostics in western countries. It is most commonly successful in central endobronchial location of the tumor, SqCC, number of sputum samples collected per patient, large tumor size and/or advanced stage of tumor. Besides proteins, sputum may also contain fragments of tumor and inflammatory cell derived DNA. Interestingly, molecular tumor aberrations may occur in sputum before morphological changes are observed by cytological analysis, and can be detected using advanced molecular techniques with high sensitivity. DNA mutations, DNA hyper methylation and microsatellite aberrations have been detected in sputum. Prevalence for gene hyper methylation in sputum is higher than in serum.

O. Stages of Cancer

The squamous cell carcinoma of the lungs can be classified into 4 stages of cancer. In stage 1, the cancer is present within the lung and has not yet spread to other parts and lymph nodes. In stage 2, the tumor already affected the nearby lymph nodes or is in a particular part of the bronchi. In stage 3, the lungs are already affected by the carcinoma. And in stage 4, other regions of the body get affected.

II. LITERATURE REVIEW

Madhura J, et.al, (2017), have analyzed that lung cancer is considered as one of the fatal diseases that affects the lungs pulmonary nodules. The diagnostics, prognostic and follow-up along with image examination is one of the essential step of lung diseases processes. As compared to other types of cancer the survival rate of lung cancer is very low. It is very necessary to identity a lung cancer in very early stage and most of the researchers from medical image processing field are working on it. Several Computer aided systems have been intended to distinguish the lung cancer at its initial stage. Various types of images are used for detection of lung diagnosis. The most important challenging task is detection of lung nodule. Computed Tomography (CT) images are generally chosen due to less distortion, low noise, better clarity, less time consumption and low cost. The speckle, Gaussian, Impulse like different noises that is present in images obtained for lung mass detection. So, numbers of researchers are actively working on noise removal from images. In this paper, authors have given a review on lung cancer, types of noises in medical images and different methods of noise removal.

Manasee Kurkure, et.al, (2016), have found that cancer is one of the most commonly affected diseases and its early detection can protect their lives. As a part of routine checkups if cancer is also checkups than it can be detected in early stages . In this paper, authors' have worked on preliminary diagnosis and lung cancer detection system design and deployment. The X-Ray, CT and PET images are used and genetic algorithm is applied for getting optimized results. The optimization algorithm allows assistant doctors to identify the nodules present in the lungs at the early stages. The Naive Bayes classification and Genetic algorithm method best features are used to overcome the difficulty of manual interpretation. After that caner images different stages are classified accurately and will be less time consuming as well critical. They have obtained a intermediate results by experimented a proposed system that results in 80% accuracy.

Prajwal Rao, et.al, (2016), have analyzed that cancer diagnosis and cure has been become a very big challenge from last few years. All over the world number of lives can be saved by early detection of cancer. In this paper, authors have presented a Convolutional Neural Network (CNNs) approach that will classify tumors in lung cancer screening computed tomography scans as malignant or benign. The spatial invariance, multiple features can be extracted are some of the special properties of CNNs. There is increase in prediction accuracy when a deep CNNs is made by cascading such layers. The CNN used in this paper helps in analyzing CT scans with tumors in this they use both neural networks and medicine domain knowledge. The proposed network is analyzed in terms of classification accuracy as compared to other existing traditional neural networks. In most of the existing works authors have used CNNs for building image classification.

Sruthi Ignatious, et.al, (2015), have concluded that lung cancer is one of the most commonly types of cancers occur worldwide. The smoking is considered as one of the main causes of lung cancer. The tumor can accurately detected using good lung cancer detection system. The cancer is mainly detected using Computed Tomography (CT) images. The enhancement, segmentation and feature extraction are different stages of Computer Aided Lung cancer detection system. These stages are performed using different techniques. In this paper, authors have given a brief review on different techniques for performing those stages.

Fatma Taher, et.al, (2015), have proposed a new computer-aided diagnosis (CAD) system that helps in early detection of lung cancer. This system is based on analysis of sputum color images. A set of features is extracted from the nuclei of the sputum cells after applying a region detection process. The support vector machine (SVM) and artificial neural network (ANN) are two classification techniques used for system training and testing that results in increase of CAD system accuracy. The accuracy, specificity, precision and sensitivity are different criteria on the basis of which system performance is analyzed. The Receiver Operating Characteristic (ROC) curve is used for evaluation. The experimental results demonstrate the efficiency of SVM classifier over the ANN classifier with 97% of sensitivity and accuracy as well as a significant reduction in the number of false positive and false negative rates.

III. RESEARCH PROBLEM FORMULATION

A. Problem statement

Cancer is the most serious health problems in the world. In 2012, cancer is leading cause of death worldwide, accounting for 8.2 million deaths. The mortality rate of lung cancer is the highest among all other types of cancers, contributing about 1.3 million deaths/year globally [1]. There are many types of cancer, Lung cancer is one of the common types causing very high mortality rate. The best way of protection from lung cancer is its early detection and diagnoses. With the fast development of the technology of computed tomography (CT) technology, medical test images become one of the most efficient examination methods to detect clinically the lung disease. As the image processing field is growing day by day, researcher moves towards bio medical field to emerge new techniques and to diagnose various medical diseases using automated image processing algorithms. Among them Lung Cancer Segmentation also known as Cancer Detection and other lung cancer diseases (Silicosis, Interstitial Lung Disease) are very crucial. Initial stages of lung cancer having similar symptoms like silicosis, TB, Interstitial Lung Disease (ILD) due to which it result in delay in treatment process. Due to delay, treatment leads to high mortality rate. Therefore, early detection of lung cancer is the major factor in successful cancer treatment. Computational methods have good potential to predict cancer in early stages. But, these methods are not available today. So an efficient method is needed to which needs to be classified and detect the Lung Cancer disease whether it is in normal or abnormal state automatically with high precision and less error probability.

B. Objectives

Now a day lung cancer is the leading cancer among both men and women. Earlier detection of cancer is the only method to improve the survival rate. Presence of lung cancer can be diagnosed with the help of cancerous and non-cancerous images of lung. The research objectives of the work are:

- To study and analyzed the existing techniques of lung carcinoma.
- To input the DICOM images as a training set and perform pre-processing using segmentation of the image using edge detection
- To perform feature extraction and extract the feature vector using scale invariant feature transform
- To perform instance selection using BFO and Cat Swarm to optimize the features
- To perform classification to classify the normal lungs and abnormal lungs using Linear discriminant analysis and evaluate the performance in terms of recognition rate, sensitivity and specificity

IV. RESULTS

This section will give the brief description about the results obtained using proposed approach. The MATLAB software is used for simulation of proposed work. We have made Graphical user interface (GUI) and perform the task to obtained the improved results. As we know Cancer is the most serious health problems in the world. In 2012, cancer is leading cause of death worldwide, accounting for millions of death. There are different types of cancer out of which the mortality rate of lung cancer is the highest. So, the aim behind this research work to propose a technique that helps in detecting Lung cancer more accurately that leads in saving numbers of person life.

The complete work is divided into different sections. Firstly DICOM images are used for testing purpose that includes both cancerous as well as non-cancerous image. Firstly we will train the system for both types of images then we will test it to know the performance of proposed approach. In Training learning process human will interact with machine. In this part firstly pre-processing is performed using edge detection then features are extracted using scale invariant feature transform. After that relevant features are extracted using Cat Swarm optimization and BFO. This completes the whole process of training approach.

Then SVM classifier is used for testing purpose and proposed approach performance is evaluated in terms of recognition rate or accuracy, sensitivity and specificity. The accuracy or recognition rate will tell how efficient a system is in testing or how accurately it will give true results. In normal cases when values are above the cut off then it indicates a disease exist and vice a versa for values below cut off. But such perfect test does not exist in real life so we have to test it with some other way. Sensitivity and specificity are two parameters that help in testing the efficiency of approach.

In below section a proposed GUI MATLAB work for proposed approach is implemented. The below figure 6 shows the main panel made in MATLAB GUI in which number of pushbuttons are used. One pushbutton is used for training purpose, one for testing purpose and third is used for exit the task. In call back function for individual pushbutton coding is done to perform a different task after clicking on push button shown in GUI.

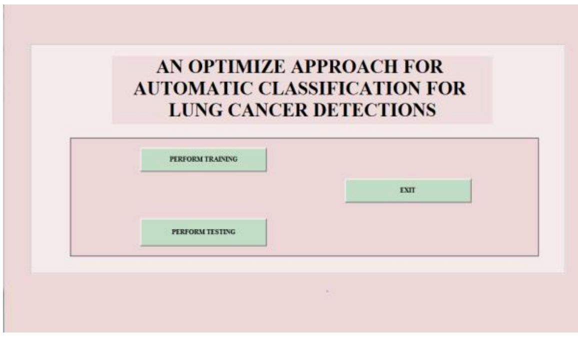

Figure 6: Main Panel

The below shown figure 7 is the training panel in which two categories are given one is cancerous and other is non-cancerous. In this firstly we have select one cancerous image from DCIM dataset images that comes when we click on Cancerous image pushbutton on Training panel pop up GUI. Once all the images will be uploaded and their features are extracted then we have to select the specific features by applying Hybrid optimization approach on it. We have done the hybridization of CSO and BFO. This operation is performed by clicking on Push button of feature optimization.

Figure 7:Training section Front Panel

The below figure 8 shows the GUI of the cancerous samples for the training scenario the same can be perform for noncancerous image. There are two axis on which the selected and EDGE detected image is shown. The edge are detcted to detect the boundary regions which will make the processing of the image easy. This is the pre-processing of samples that is performed before extracting features from it. The DICOM medical images are shown in GUI that is used for training purpose. There is one more axis that on which extracted features are uploaded for each samples side by side as the process of uploading is completed for the cancerous sample. The same is done for Non-cancerous images. The scale invariant feature transform is used for feature extraction which shows the independent characteristics values for each uploaded samples in terms of Intensity, Entropy and Contrast.

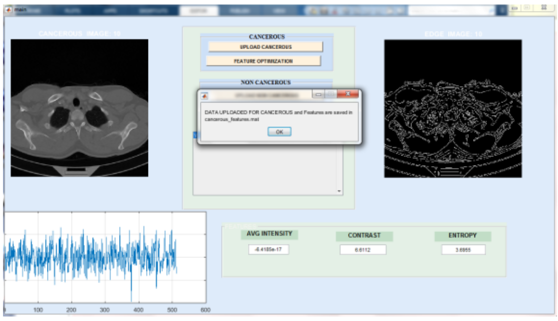

Figure 8: Feature Extraction

The below Figure 9 shows the optimized features that are obtained after passing extracted features from scale invariant transform. We did a hybridization of CSO and BFO. The CSO stands for Cat Swarm Optimization as we know cats spend most of their time in resting still it have high alertness and curiosity about their surroundings and objects move in their environment. Their behavior helps them in finding preys and hunt them. The Chu and Tsai has developed CSO by getting inspired from their hunting pattern. They have two modes in which they work the first is seeking mode and other is tracing mode. Below is the algorithm of CSO working to get optimized results.

Algorithm of CSO:

Start

Input parameters of the algorithm and the initial data

Initialize the cat population X_i *(i = 1, 2, ... , n), υ, and SPC*

While (the stop criterion is not satisfied or $I < I_{max}$*)*

Calculate the fitness function values for all cats and sort them

X_g*= cat with the best solution*

For = 1:

If SPC = 1

Start seeking mode

Else

Start tracing mode

End if

End for i

End while

Post-processing the results and visualization

End

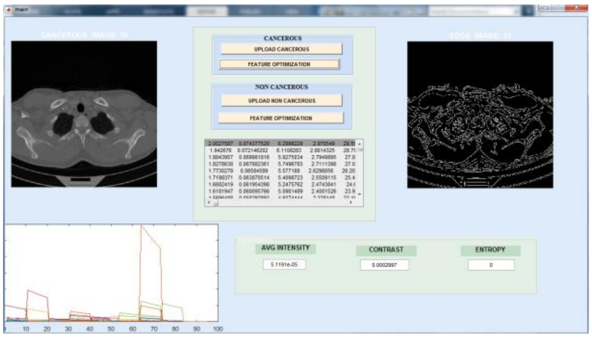

Figure 9: Optimized Features

In this work we have also used BFO that is stands for Bacterial Foraging Optimization. In 2002, Kevin Passino has proposed BFO that is a nature inspired optimization algorithms. Application of group foraging strategy of a swarm of E.coli bacteria in multi-optimal function optimization is the key idea of this new algorithm. Bacteria search for nutrients is a manner to maximize energy obtained per unit time. Individual bacterium also communicates with others by sending signals. A bacterium takes foraging decisions after considering two previous factors. The process, in which a bacterium moves by taking small steps while searching for nutrients, is called chemotaxis. The key idea of BFO is mimicking chemotactic movement of virtual bacteria in the problem search space.

Algorithm of BFO:

Initialize the parameters S, Nc, Ns, Nre, Ned, Ped and the C (i), (i =1, 2,…, S).

Elimination-Dispersal loop: l= l+1

Reproduction loop: k= k+1

Chemo taxis loop: j= j+1

(a) For i= 1, 2,….,S take a chemo tactic step for bacterium 'i' as follows:

(b) Compute cost J(i,j,k,l).

(c) Let J(i,j,k,l)=J(i,j,k,l)+Jcc(□ i (j,k,l),P(j,k,l))

(d) Let Jlast = J (i,j,k,l) to save this value since find better cost via a run

(e) Tumble: Generate a random vector

Compute J(i,j+1,k,l).

Swim;

If j<Nc go to step 3. In this case, continue chemo taxis, since the life of the bacteria is not over.

Reproduction

Step 8: Elimination-Dispersal

End

After feature extraction and saving those features next step is testing a proposed approach. For testing purpose firstly we will click on push button in Front panel GUI name perform test. In testing part we have to select one sample from images in Test folder. In this folder we have randomly taken few images of cancerous and few of noncancerous. When we select one image either can be cancerous or non-cancerous that will predicted and tested in terms of accuracy or recognition rate, Sensitivity and specificity. The below Figure 10 shows the image of selected image for testing purpose and feature samples are shown in figure 11.

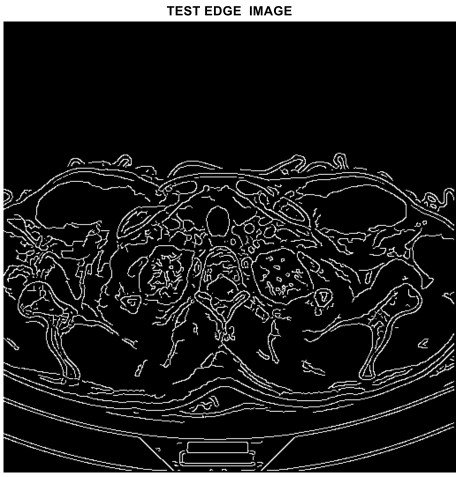

Figure 10: Selected sample

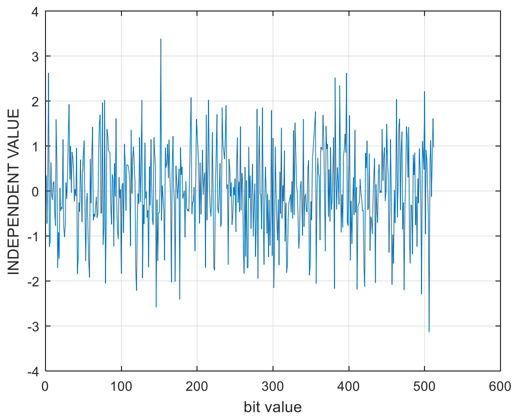

Figure 11: Selected test image feature samples

The figure 12 shows the detection of the category automatically by the system. Firstly the random image which is also known as the test sample is considered for the recognition of the right sample whether it is cancer or non cancer. The below message box shows that the uploaded test sample is the cenceorus. It is done using the classification approach which is known as the support vector machines and classifies the category based on the training optimized feature vector and test features.

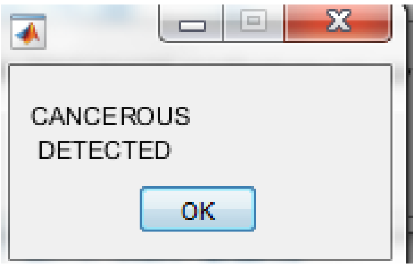

Figure 12: Detection

The below figure 13 and figure 14 shows the results of proposed approach in terms of sensitivity. It is the test ability of correctly determining the Lung cancer patients.

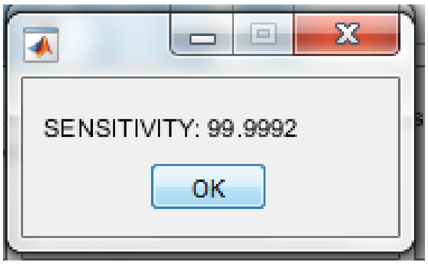

Figure 13: Screenshot of Senstivity

Mathematical formula used:

Sensitivity: 1-FAR

Where FAR is a false acceptance rate that deals with how much your system is able to accept the wrong samples based on feature extraction which must be low. As it can be depicted from Figure that there is decrease in senstivity as number of elements are increasing that depict a proposed approach is good.

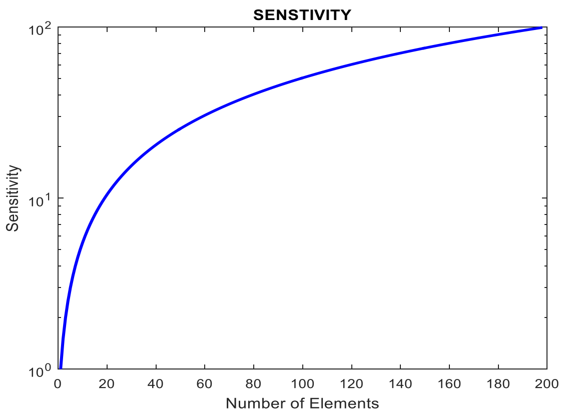

Figure 14: Sensitivity of proposed approach

The below figure 15 and figure 16 gives the tested results of proposed approach in terms of Specificity. It is the ability to test healthy cases correctly.

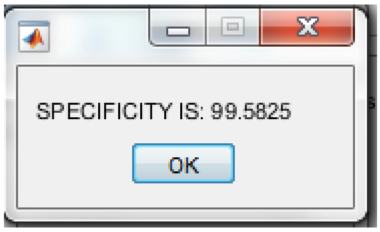

Figure 15: Screenshot of Specificity

Mathematical formula:

Specificity: 1-FRR

Where FRR is stands for false rejection rate that deals with how much your system is able to reject true samples which also must be low. The value of both FAR and FRR should be low for better system. As it can be depicted from Figure that there is increase in specificity as number of elements are increasing that depict a proposed approach is good.

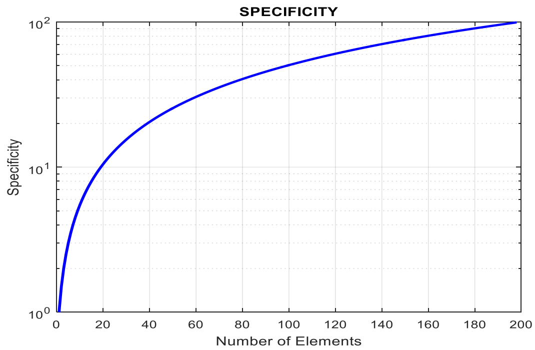

Figure 16: Specificity of proposed approach

The below figure 17 and figure 18 shows the recognition rate of proposed approach. It is the test ability of correct detection of patients.

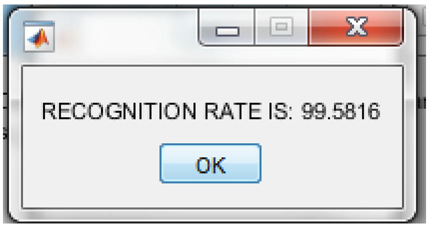

Figure17: Screenshot of Recognition Rate

Mathematical Formula:

Accuracy: 1-(FAR+FRR)

The accuracy or recognition rate should be high. We have to calculate FAR and FRR for testing the recognition rate or accuracy of a approach. As it can be depicted from Figure that there is increase in accuracy as number of elements are increasing that depict a high recognition rate in terms of testing the patient.

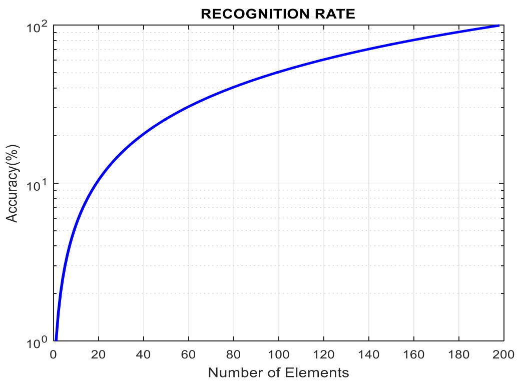

Figure 18: Recognition Rate of proposed approach

Table 1: Performance parameters:

Parameter	Cancerous	Non Cancerous
Sensitivity	99.9992	99. 3297
Specificity	99.5822	99.9987
Recognition Rate	99.5814	99. 3284

The above shown Table 1 shows that results obtained by proposed approach in terms of Specificity, Sensitivity and Recognition rate is very high. This result depict that proposed approach is efficient in terms of accuracy of recognition of cancerous or non cancerous image.

Table 2: Comparison results of existing and proposed approach

Parameter	Existing Approach	Proposed Approach
Sensitivity	98.3	99.9992
Specificity	98.6	99.5822
Recognition Rate	98.5	99.5814

This can be concluded from Table 2 that the results obtained using proposed approach is better than existing approach used by author in terms of Specificity, Sensitivity and Recognition rate.

CONCLUSION

As we know that unnecessary growth of tissues which are responsible for increasing cancerous is one of the major problem in medical field. Lung cancer is one of them. The proposed approach is divided into two phase. The very first is training because it is a learning process in which human is interacting with machine. The training approach deals with the feature extraction which is done by independent component analysis which is used to extract the feature vector which act as a characteristic value of the particular uploaded sample for the training purpose. Then the feature optimization is used to extract the relevant features form the feature vector using BFO and Cat swarm optimization. This completes the whole training approach. Then we have moved to the testing process in which the detection process is done using svm classification. The SVM acts as a classifier and the performance is evaluated using false acceptance rate, false rejection rate and accuracy or recognition rate. So from the above result and discussions the proposed approach achieves high recognitions with less error rate probabilities. In the future the work can be done for the comparative analysis in terms of feature extraction and classification processes and on the basis of that the performance parameters are evaluated. The another optimizations methods and feature to detect the cancer symptoms for lung cancer detections can also used.

REFERENCES

[1] Punithavathy Kannuswami, Sumathi Poobal, M. M. Ramya, “Artificial Neural Network Based Lung Cancer Detection for PET/CT Images”, Indian Journal of Science and Technology, vol. 42, 2017.

[2] Snehal Dabade, Shubhangi Chaudhari, Sneha Jadhav, Arjun Nichal, “A Review Paper on Computer Aided System for Lung Cancer Detection”, 2017 International Conference on Big Data, IoT and Data Science (BID), pp. 98-102, 2017.

[3] Agus Maman Abadi, Dhoriva Urwatul Wutsqa, Leonardus Ragil Pamungkas, “Detection of Lung Cancer using Radiograph Images Enhancement and Radial Basis Function Classifier”, 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI 2017), pp. 1-6, 2017.

[4] Qing Wu, Wenbing Zhao, “Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm”, 2017 International Symposium on Computer Science and Intelligent Controls, pp. 88-91, 2017.

[5] Ratih Wulandari, Riyanto Sigit, Setia Wardhana, “Automatic Lung Cancer Detection Using Color Histogram Calculation”, 2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC), pp. 120-126, 2017.

[6] Taolin Jin, Hui Cui, Shan Zeng , Xiuying Wang, “Learning deep spatial lung features by 3D convolutional neural network for early cancer detection”, IEEE, pp. 1-6, 2017.

[7] Madhura J, Dr .Ramesh Babu D R, “A Survey on Noise Reduction Techniques for Lung Cancer Detection”, International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2017), pp. 637-640, 2017.

[8] Manasee Kurkure, Anuradha Thakare, “Lung Cancer Detection using Genetic Approach”, Computing Communication Control and automation (ICCUBEA), 2016 International Conference, pp. 21-26, 2016.

[9] Bhagyarekha U. Dhaware, Anjali C. Pise, “Lung Cancer Detection Using Bayasein Classifier and FCM Segmentation”, 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), pp. 170-174, 2016.

[10] Avinash. S, Dr .K .Manjunath, Dr.S.Senthil kumar, “An Improved Image Processing Analysis for the Detection of Lung Cancer using Gabor Filters and Watershed Segmentation Technique”, Inventive Computation Technologies (ICICT), International Conference on, pp. 108-114, 2016.

[11] May Mohamed Abdelhamid, Prof. Dr. Abdelmegid M. Allam, “Detection of Lung Cancer Using Ultra Wide Band Antenna”, 2016 Loughborough Antennas & Propagation Conference (LAPC), pp. 121-126, 2016.

[12] Joseph B. Majeski, Kevin J. Lynch, George C. Giakos, “Lung Cancer Tissue Samples Discrimination through Statistical Analysis of Polarimetric Angular-Resolved Diffuse Measurements “,MIT Undergraduate Research Technology Conference (URTC), 2016 IEEE, pp. 21-25, 2016.

[13] Hugo Simoes, Mariana Alves Barros, Paulo Crespo, "OrthoCT for Tumor Lung Irradiation: a Simulation Study", Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD), pp. 21-26, 2016.

[14] Rachid Sammouda, "Segmentation and Analysis of CT Chest Images for Early Lung Cancer Detection", 2016 Global Summit on Computer & Information Technology, pp. 120-126, 2016.

[15] Prajwal Rao, Nishal Ancelette Pereira, Raghuram Srinivasan, "Convolutional Neural Networks for Lung Cancer Screening in Computed Tomography (CT) Scans", 2016 2nd International Conference on Contemporary Computing and Informatics (ic3i), pp. 489-493, 2016.

[16] Divya Chauhan, Varun Jaiswal, "An Efficient Data Mining Classification Approach for Detecting Lung Cancer Disease", Communication and Electronics Systems (ICCES), International Conference, pp. 102-110, 2016.

[17] Sruthi Ignatious, Robin Joseph, "Computer Aided Lung Cancer Detection System", Proceedings of 2015 Global Conference on Communication Technologies(GCCT 2015), pp. 555-558, 2015.

[18] Fatma Taher, Naoufel Werghi, Hussain Al-Ahmad, "Computer Aided Diagnosis System for Early Lung Cancer Detection", Systems, Signals and Image Processing (IWSSIP), 2015 International Conference, pp. 5-8, 2015.

[19] Md. Badrul Alam Miah, Mohammad Abu Yousuf, "Detection of Lung Cancer from CT Image Using Image Processing and Neural Network", 2nd Int'l Conf on Electrical Engineering and Information & Communication Technology (ICEEICT), pp. 132-137, 2015.