The Lung Cancer Predictive Accuracy for Non-Smokers Using Classification and HFPS Algorithm

B.Mohamed Faize Basha[a], Dr.M.Mohamed Surputheen[b]

[a]Research Scholar, Department of Computer Science, Jamal Mohamed College (Autonomous), [Affiliated to Bharathidasan University], Tiruchirappalli, India.
[b]Associate Professor, Department of Computer Science, Jamal Mohamed College (Autonomous), [Affiliated to Bharathidasan University], Tiruchirappalli, India.

Abstract: Lung cancer identification and projection at the earliest point of comparison may be really helpful to boost the patient's survival rate. But cancer detection and prediction remains one of radiologists' toughest obstacles. An intelligent, computer-assisted diagnostic device may be very helpful for radiologists in identifying, forecasting and diagnosing lung cancer. This paper proposed a Hybrid algorithm as Firefly–Particle Swarm Algorithm (HFPS) for feature selection process. In this study, we conduct the simulation by taking the LIDC-IBRI dataset. Initially the raw data is pre-processed by using Laplacian filters. The segmentation of images by using the modified KFCM techniques. Then extraction of feature by using transformation techniques as DWT. After completion of feature selection and extraction then we classified the data as cancer or normal. By this classification process is successfully done with the help of deep learning techniques as recurrent neural networks, Auto Encoder and LSTM algorithms are used. In this study we compare the analysis by using three different classifier with input of feature selection method and without that. In our model, we achieved the successful classification rate of HFPS Feature Selection with RNN classifier achieved 97.43% of accuracy. In, these methods used to detect the lung cancer in early stage to reduce the mortality rate.

Keywords: Auto Encoder, Deep Neural Networks, Hybrid Firefly–Particle Swarm Algorithm and LIDC-IBRI dataset.

1.Introduction

Lung cancer is the highest source of cancer mortality worldwide(**American Lung 2014)** and prohibits productive therapies of individuals that have been diagnosed with lung cancer in early stages of the disease. Earlier cancer diagnosis increases the longevity of patients by taking appropriate care without actually prolonging their lives. About $9.6 billion is being invested on lung cancer care in the United States annually. While they have health care, this is a huge financial strain on the people. With the rise in modern technology and therapies, cancer prevention budgets may be more quickly increased than total care costs.**(Bethesda 2017)** These facts place pressures on budgetary cancer control and prevention systems, such as computerised pulmonary cancer screening programmes. The National Lung Screening Trial indicates that Low Dose Compute Tomography (LDCTs) screening lowers death rates from lung cancer **(Nasim F 2019)** Recommendations for successful conduct of the lung cancer screening programme are provided by the American College of Chest Physicians.**(Wiener RS 2015)** Lung cancer screening with LDCT is suggested for adults aged 55 and 80. Routine screening with CT imagery is advised for early cancer diagnosis in high-risk patients. However, greater consideration is required when repeating LDCT checks, as radiation exposure accumulates. The newly released realistic recommendations of the American College of Chest Physicians also proposed lengthy periods for CT scans.**(Detterbeck FC 2013)** In contrast with the cutting rate of death related to cancer, the US Preventive Programs Task Force has recorded a marginal effect on radiation exposure.

Screening of LDCT is helpful in diagnosing lung cancer, and will contribute to a greater likelihood of long life if lung cancer is earlier before it progresses to other species. False-positive (FP) findings will however bring individuals to a more advanced radiation examination that is likely to harm their usual wellbeing. Cautious evaluation and screening are also highly necessary. The machine-learning group has recently established computerised diagnostic resources and learning models that demonstrate clinically appropriate efficiency. Currently, the Food and Drug Administration has provided premarket clearance in two CAD areas, for example, mammogram diagnosis for breast cancer and chest diagnosis for lung cancer (Rao RB 2007). In order to reach a greater precision, the segmentation efficiency needs to be enhanced due to lung-node variability in CT pictures. Deep Learning is a significant region of machine learning that applies neurons to learn tasks including mathematical formats. In the past, in the several decades**(Fukushima 1980)** of the decision-making process, the Neural Networks have been influential in allowing researchers to win and release favour. Deep learning and neural networks was merged to meet the problems of classification termed deep neural networks (DNNs),.

2.Related works

Nima et al. **(Torbati 2014)** presented a modern medical image segmentation neural network method. They used the updated SOM network, dubbed Moving average SOM (MA-SOM), used for medical picture segmentation. A new method of fusion was developed to bind image artefacts of a joint cluster together after initial segmentation. Their method is even better than the ISNN and SOM (incremental controlled neural networks) and the MR (Computerized Tomography) approach. Medical photos are

(Drozdzal 2018) Michal et al. implemented a clear, but efficient pipeline that combines completely convolutionary networks with fully convergent residual networks for the medical imaging segmentation process. They also explored a design that gains from recent developments in both convolutionary neural network and resnet understanding. You also concentrated on the most significant pre-processing by using FC-ResNets and shown that an FCN model with a low ability can be a pre-processor for the normalisation of medical details. During their segmentation process, the neural networks are used to obtain normed pictures which are optimised iteratively for the design of a segmentation prediction model.

Wencheng et al. **(Zhao 2018)** have developed a modern and updated version of the medical picture active contour model combining local and global medical picture strength details with local fitting and global word. In image segmentation with reduced iteration during medical picture handling, they have achieved better efficiency.

For edge detection of biomedtical pictures, Vardhana et al.**(Vardhana, M 2018)** proposed low-power architecture. In order to diagnose different diseases by using convolutional image recognition, the output of their edge detection method was provided as the output to their model. They also used the new algorithms Sobel, Prewitt and VLSI, which have been used to detect the edges and incorporate the architecture's interactive IC style. These algorithms are used with a co-simulation with MATLAB and Modelsim for edge detection.

Jiaqing et al.**(Miao, J 2018)** used a cosine function to convey the fit of the data energy of a standard active contours, developed a new model centred, in line with the local energy-fitting active contours used for medical and synthetic photos, on the sectional picture recovery. **(Miao, J 2018)** In addition, the devices were designed to manage the images with equivalent or faster noise, ruggedness and computational speed in contrast with current models in that direction. In addition, they defined the model in a distinct form that is simpler to apply to the segmentation a normal word. Finally, they revealed how effective and productive their segmentation findings were for different forms of medical pictures.

By utilising the discrete cosine transform, Chandan and Anu **(Singh, C 2018)** suggested a different method. The DCT-based local and non-local FCM(DCT-LNLFCM) is referred to as the FCM distance function as the sum of the local and non-local distances which are in themselves the Euclidean distance weighted values used in the FCM.

3.Proposed system

In below figure 1 shows that the prediction and detection of lung cancer by our proposed model flow diagram. In this work flow, we conduct the simulation analysis by using the dataset. In these dataset description is briefed in following section and all these raw data's have noisy contents so planned to pre-process the raw data by using the HLOG filter. Then the modified KFCM clustering technique is used to segment the CT images. In these clustering method, basically the centroid problem is occurred during the selection, to avoid this problem by using the fire fly based centroid technique is applied, then transformations of images is significant thing in image processing, so we conduct the DWT transformation technique is used to extracts the images. Then major contribution of this study is feature selection, in these process we used the Hybrid Firefly–Particle Swarm Algorithm (HFPS). By the classification of lung cancer, the deep learning based algorithms are used. The algorithms explanations are breifiely explained in following section.

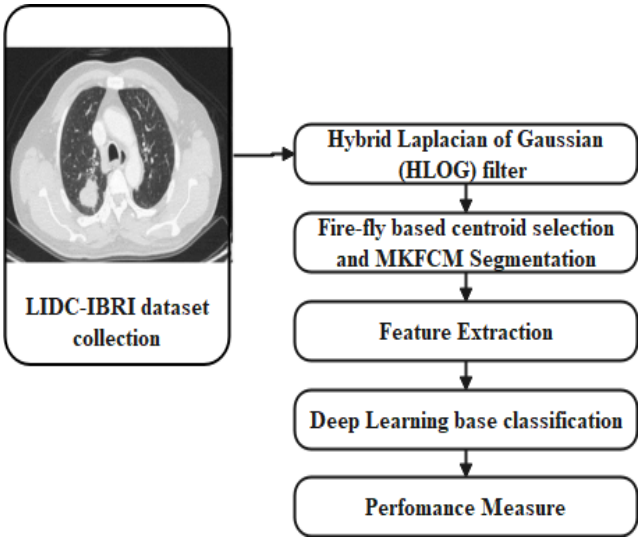

Fig.1. Proposed system.

3.1.LIDC-IBRI dataset collection

The Lung Image Database (LIDC-IDRI) consortium comprises of demonstrative thoracic registered tomography (CT) screens with expanded damage. It is a worldwide web-open asset for the development, planning and review of the PC, which has been used for demonic detection and conclusion of lung malignancy strategies. This knowledge set, which comprises 1018 cases, was rendered possible by seven academically based and eight restauratory imaging organisations. Every topic consolidates images from a clinical thoracic CT review and an accompanying XML database recording the eventual findings of a technique of two-set images of four accomplished thoracic radiologists. 240 images are educated and 90 images checked in the suggested method in LIDC-IBRI dataset.

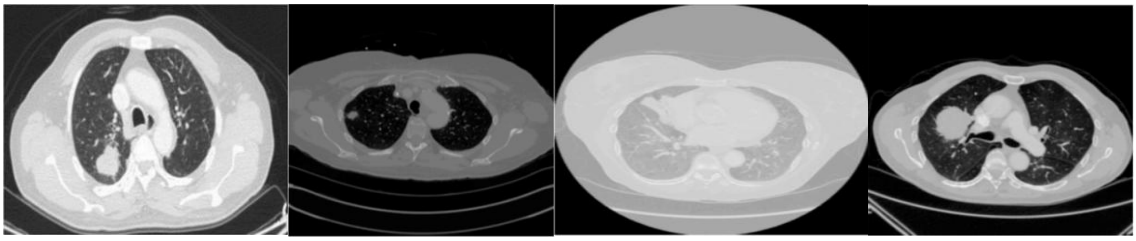

Fig.2. LIDC-IBRI dataset Sample images.

3.2.Pre-processing

The laplacian filters are usually derivative filters and are used to identify the edge areas in pictures. These derivative filters are very susceptible to noise and are used to smooth out the picture using Gaussian filters. Prior to Laplacian filtering the Gaussian filtration is carried out. The Gaussian filtering is added to smooth the pictures again after these two methods. The LOG filter is indicated in Equation (1) and the Gaussian LOG filter is equated (2).

The representation of LoG scale-space is

$$\nabla^2 G(x,y) = \frac{x^2+y^2-2\sigma^2}{\pi\sigma^4} exp\left(\frac{-x^2+y^2}{2\sigma^2}\right) \quad (1)$$

$$G(x,y;\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(\frac{-x^2+y^2}{2\sigma^2}\right) \quad (2)$$

After the LOG process, the output image is given to segmentation process.

3.3.Fire-fly based centroid selection and Segmentation

FF-based MKFCM technology attempts to define the bunch focus, which restricts the (objective) potential of MKFCM. By iteratively refreshing the emphasis of the community and the participation ranking for each date point, MKFCM shifts the focus group "to one side" within a compilation of details. Whatever it is, seeking an optimal arrangement in an ideal time is impractical. The functioning of the MKFCM depends on the centroids, so in the MKFCM it is most necessary to decide the centroids. The FF-dependent MKFCM is an automated centroid collection based on MKFCM based FF for this resolution. The suggested protocol consists of two components, the clustering module FF and the clustering module MKFCM. The grouping module FF is done for a short time in the underlying process for the programmed bunching, system circular or close round type knowledge packets. As the underlying seed of the MKFCM module the FF grouping moducle findings are used. Contour by Geometric Dynamic Structure Framework are the improved parts (GDSF). Here, the goodness and greatest health love are fixed in the middle of our appreciation. The exhibition is eventually dealt with using the wellbeing value of the MKFCM Division.

A.Firefly (FF) based centroid selection

The author (Yang 2008, Yang 2009) developed the FF algorithm, and romatized action depends on fireflies' gleaming quality. For convenience, the three following criteria may be used as those glimmering characteristics. One FF is taken in by multiple burners who pay no regard to their sex as the fireflies effect is unisexual.

Their attraction is social, the less beautiful one would switch to the more glorious one in this process with any two flaming fireflies. The quality of the engagement equals the luminosity and the two fall as their separation rises. If no one is lighter than a particular FF, it is haphazard;

A scene of the destination to be progressed is affected or determined by the brilliance or light power of the FF. The splendour can be linked to the mission for an expansion problem. In these respects the health work of genetic or bacterial scrounge measurements (BFA) can be described by various forms of glamour (Gazi and Passino 2004).

There are two critical problems in the FA: the light force variation and the attraction concept. We may practically cause the appeal of an FF to be arranged by the magnificence or light power of the encoded goals. The brilliance I of the FF in a specified field x can be chosen as I(x) for the most easy-to-use streamlining problem (x). In every case the appeal β is comparatively, it can be identified by the audience or by many fireflies. In this way, the distinction "rij" between FF I and FF "j" could fluctuate. Because light power decreases with the right approaches and light is expended in the media, we can encourage the appeal to adjust with the degree of assimilation.
In the basic method of light intensity $I(r)$ different with the distance r monotonically and exponentially. That is
$I = I0e - \gamma r$ (3)
The light retention coefficient is where I0 is the first light force and μ. As an FF appeal is relative to the light force seen by nearby fireflies, we can now distinguish the allure β of an FF
$\beta = \beta 0e - \gamma r2$ (4)
Where β0 is the call in r = 0. It is worth noting that different skill, for instance $\gamma r2$ when m> 0, will substitute for γr sort.

B.Post process outcomes and visualization

Two fireflies are happy for the Cartesian separation rij = ||xi - xj||2 or the '2 norm between the separation from I and j at Xi and XJ. The time delay or other reasonable arrangement for various processes, for example booking, may be time delay, not really cartesian division. The production of an FF I is eventually attracted by another (glorified) FF driven by,

$$xi = xi + \beta 0e - \gamma rij\, 2\,(xj - xi) + \alpha \epsilon i, \alpha i, \quad (5)$$

Whereas the fascination for the following term, while the third term involves randomization with the vector of irregular factors pi taken out of a Gaussian transport. We will take β0 = 1, α to [0,1] and β = 1. for the better cases of our use. Moreover, when the scales change intriguingly in changed measurements, e.g. −105 to 105 from one measure, but are −10−3 to 103 along others, it is intended to change α by α Sk, whereby in d measurements the scale parameters Sk(k=1,...d) should not be stopped by unmistakable intrigue sizes. The β parameter basically shows the distinction between an intrigue and its value is vital in determining the union speed and how the FA is calculated. In principle, the Sk trademark/mean Sk of the frame to be expanded, however, in practise, 5-0 | > 1 is controlled by the trademarks. The allure is constant β = β0 in one frightening when μ is around 0. This equals that the light force in a glorified sky would not decrease. The FF burning room

can be viewed everywhere. In this way an ideal is undoubtedly solitary (usually worldwide). This concerns a remarkable case of swarm molecules rationalisation (PSO). It can be said in truth, if the inner circle for j is dismantled and the present worldwide best case is replaced by Ij, FA is mainly converted into the standard PSO, and, in these respects, the productivity of this extraordinary case equals PSO. Then again when μl is a β(r) to μmlm(r) is a β(lm) work. This implies that when you see different fireflies, the engaging quality is almost zero or the fireflies are nonsense. This is proportional to the fact that the fireflies haphazardly fly in a foggy area. There can be no different fires and every firefly meanders completely irregularly. This concerns the utterly random strategy of hunting in certain lines. Thus α tests how the calculation functions. It is also necessary to adjust β in order to locate different optimas at the same time.

C. MKFCM Algorithm

FCM system is dividing the dataset $\{x_k\}_{k=1}^{N}$ into c sum of clusters based on the subsequent objective function Eq. (6).

$$J_{\text{MKFCM}} = 2[\sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^{p}(1 - K(x_k, v_i))] \quad (6)$$

"P," in which the actual number is indicated and the method of quality control is indicated for the fluffiness of the resulting group, belongs to the cluster and is the pixel of the image. The middle of the cluster is to be pleased and noticed. In the previous equation, the total number of clusters is indicated and the data points count is indicated. The FCM distinguishes by updating the affiliation values and the cluster centres. Each data's membership value is also extracted from each centroid after each centroids update. And from the Eq below it can be derived. **(Torbati 2014)**.

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{\|x_k - v_i\|^2}{\|x_k - v_j\|^2}\right)^{\frac{1}{p-1}}} \quad (7)$$

The bunch centres, based on the separation between the knowledge point and the group middle, is refreshed. Eq. (8).

$$v_k = \frac{\sum_{k=1}^{N} x_k U_{ik}^{p}}{\sum_{k=1}^{N} u_{ik}^{p}} \quad v_i = \frac{\sum_{j=1}^{N} u_{ij}^{m} x_j}{\sum_{j=1}^{N} u_{ij}^{m}} \quad (8)$$

The objective work quantifies the weighted whole outcomes between the emphasis of the Category and knowledge provided in the fluffy categories. The videos, which have no clamour, are best separated by MKFCM. These images are provided to the intersection by extraction systems following the division process.

3.4.Feature Extraction

DWT takes into account the rectangular function to calculate the entrance coefficient of the input images. The DWT system has excellent compression of energy, short filters, low calculations and proper redundancy-free reconstruction. DWT follows an ambiguous noise system that provides better directional selectivity and changes in subbands and lower frequencies. The actual image form is calculated through magnification and subtraction during the multi-resolution process. DWT, which takes into account rectangular function, is usually used to calculate the input coefficient of input images. The system has excellent energy compression, short filter support, low calculation and proper redundancy reconstruction. DWT follows an ambiguous noise procedure, providing improved directional selectivity and changing subbands and a lower frequency. The actual composition of the image is calculated by magnification and subtraction in the multiresolution process. In general, the images in the different resolutions are divided into multiple sub-images in order to keep data low and high. DWT helps you to extract textural information from images. The integrated squared function f(u) and wave transformation w is the domestic product f and the value = ni/u is the real value function. In equages, the wave function is divided into several sub-images at different resolution stages, which keep information low and high frequency. The DWT property helps you to extract textures information from pictures. The integral squared

function f(u) and the transformation w of the wave represent the inner product f and the value ompletion(u) is the actual value function. The wave function in eq. (9). is given

$$w[f(s,\tau)] = \left(f, \psi_{s,t}^{k}\right) = \int_{-\infty}^{\infty} f(u)\psi_{s,t}^{k}(u)du \quad (9)$$

$$Where, \psi_{s,t}^{k}(u) = \left(\frac{1}{\sqrt{s\psi_{s,t}^{k}}}/s\right) denotes \quad (10)$$

Wave family, s μz is the size, ć is translation, and k {h,v,d}is the parameters of orientation. H, V and ddenote respectively have vertical, horizontal and diagonal orientation. Durations of <2^j and μ=2^j, n,j,nαz is obtained by DWT. The dynamic wavelet breakdown is a scalable DWT trial accompanied by a geometric ratio 2 series. Dyadic waves are deployed for the subsequent breakup of the wavelet using perfect re-construction filter banks.

3.5.Hybrid Firefly–Particle Swarm Algorithm (HFPS)Feature Selection

The optimization of model parameters is a difficult problem in using ensemble dependent learning when forecasting flash fluxes, and metaheuristic optimisation algorithms have shown superiority[14-15] in this regard. This study used a firefly optimization algorithm (FA) incorporation and PSO (Particular Swarm Optimization) suggested by Aydilek**(Aydilek, I.B 2018)**. This selection is attributed to the assumption that the HFPS algorithm inherits the quick PSO calculations and FA's solidity to construct a versatile new algorithm. Therefore, in different engineering issues, the HFPS algorithm outperforms benchmark algorithms**(Aydilek, I.B 2018)**. In general, the HFPS protocol can be defined briefly as follows.

(1) Specify the swarm population, venue, velocity and the total number of iterations used by each particle;

(2) Determine the health of a particle regarded as the best particle, then evaluate all pbests to get the best in the world (pbest). (2) (gbest).

(3) Determine and change the position (Pos) and velocity (Vel) for each iteration for all particles in the swarm with the use of Equations (11) and (12). If the physical state is not changed, Pos and Vel will be revised for each practise with Equations (13) and (14).).

$$Pos_i(t+1) = Pos_i(t) + B_0 e^{-\gamma r^2 ij}(Pos_i(t) - gbest^t) + a \quad (11)$$

$$Vel_i(t+1) = Pos_i(t+1) - Pos_i(tp) \quad (12)$$

$$Vel_i(t+1) = \omega Vel_i(t) + C_1 r_1(pbest_i(t) - Pos_i(t) + C_2 r_2(gbest(t) - Pos_i(t)) \quad (13)$$

$$Pos_i(t+1) = Pos_i(t) - Vel_i(t+1) \quad (14)$$

(1) Where Posi and Veli are position and speed of the firefly of i-th or i-th particles, rij is the gap between the fireflies, β is the fireflies' light absorption coefficient, B0 is the PSO's weight of attractivity, tp is the temporal position, r_1 and r_2 are the random parameters 2 [0, 1], c_1 and c_2 are the acceleration coefficients of a

(2) (4) In each iteration, compute the best gbest and then delete with this gbest the particle coordinates. The coordinate values for the flash flood ensemble model are called the configured parameters.

3.6.Classification

In the proposed system, various classification techniques are separately analysed to verify the proposed performance of the proposed classification techniques. The classification techniques such as recurrent neural network, Auto Encoder and LSTM.

A.Auto Encoder Classification

The parts in question are classified using the Auto Encoder DNN Classifier based on the selected features. In this research, the automated encoder can be a suitable solution for the proper selection classification process of the DNN when there is no prior knowledge of the distribution data. Auto encoder DNN is normally an FFN and is an unmonitored, greedy-layered pre-training technique. Data flow is obtained without looping in DNN from the input layer to the output layer. The main advantage of the Auto Encoder DNN Classification is that the chance of losing value is too small.

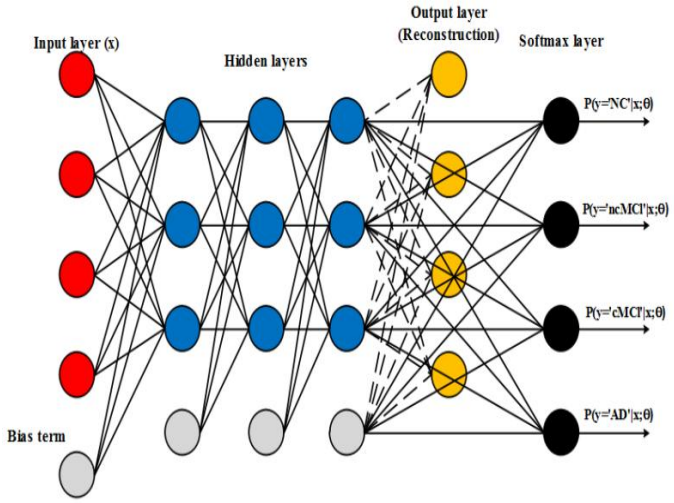

Figure.3. Structure of deep leaning.

An automatic encoder is a coding structure composed of a multilayer neural network as shown in Figure 3. Input-encoding neurons represent the true input value. Each hidden layer can be seen as a high-level representation of its former fold, however, the true significance of each layer is very small. The cover of the insertion layer is the smallest representation of the insertion layer of the same size as the insertion layer. Activation signals are transferred across the network using Eq. (15). Until the accrual reaches the exit rate. Neuronal activity can be measured using EQ of each overlay. (16).

$$\begin{cases} a_{(i)}^{(i)} = x^{(i)}, l = 1 \\ a_{(i)}^{(i)} = \sigma(W^l a + b), l = 1 \end{cases} \quad (15)$$

$$h(W, b, x) = a^{(N)} \quad (16)$$

Where x represents unlabeled data $\{x^{(i)}\}_{i=1}^{m}, w = 1$,represents weight matrix controls the activation effect between neurons on the neighbor layer, b signifies bias term, σsignifies activation function, it can be set to hyperbolic target function to delivernon-linearity for network to model difficultrelationship, and $h(W, b, x)$signifies input data as well as beginning output layer. To train this unsupervised ideal, the loss of production is used as an objective function for optimization as given in Eq. (17).

$$L(W, b, x, z) = \min_{W,b} E(W, b, x, z) + \gamma||W||_2^2 + \beta K(W, b, x) \quad (17)$$

Where, $E(W, b, x, z) + \gamma||W||_2^2$signifies the demonstration loss with squared error. The second term reduces the weight leading to the lower weight, while the third term regulates the sparsity penalty regimeβ with an activation of the targetρ near target 0, which imposes a poor representation of the objective function using the Kull-back-Leibler deviation in all training trials. . This is defined in Eq. (18).

$$K(W, b, x) = \sum_j^n ID_{KL}(\rho||\sum_{i=1}^m h_i(x^{(i)}; W, b)) \quad (18)$$

This research simultaneously trains the hidden layer of the encoder and creates a complete NN by eliminating the temporal output layer. A soft-Max output layer has been added to the top of the trained self-coding stack, which contains only the previously hidden layer. The soft-max layer uses another activation function, which can be linear, different from the one applied to the previous layer. The soft-max activation function is Eq. (19).

$$h_i^l = \frac{e^{w_i^l h^{l-1} + b_i^l}}{\sum_j w_i^l h^{l-1} + b_i^l} \quad (19)$$

Where w_i^l is i^{th} row of $W^l$and $b_i^l$is i^{th}ending layer bias term. This research can employ h_i^las an estimator of$P(Y = i|x)$. Where Y is the connected label of input data vector$x$.In this case, the 4 output neurons in the softmax layer can be interpreted as an AD determinant.

B.LSTM based Classification

Some model succession, e.g. intermittent neural organisations, are entirely reasonable for assessment as they can demonstrate long-term dependence and use the relevant data in such a manner. However, intermittent neural systems are fundamentally concerned with the deterioration and explosion of tendencies **(Bengio et al., 1994)**. This problem can be managed by LSTM **(Hochreiter and Schmidhuber, 1997)**. LSTM also has a repetitive layer consisting of memory squares, as other intermittent neural systems. There's a memory cell unit that can store memory status data in every memory obstacle and a few doors to control the memory status difference. The more formally,

$$ft = \sigma(Wfzzt-1 + Wfxxt + bf + \text{Wfxxt} + \text{bf} \quad (20)$$

$$it = \sigma(Wizzt-1 + Wixxt + bi) \quad (21)$$

$$C\sim t = \varphi(Wczzt-1 + Wcxxt + bc\text{t} = \varphi(\text{Wczzt-1} + \text{Wcxxt} + \text{bc} \quad (22)$$

$$Ct = it \odot C\sim t + ft \odot Ct-1 \quad (23)$$

$$ot = \sigma(Wozzt-1 + Woxxt + bo) \quad (24)$$

$$zt = ot\, \varphi(Ct) \quad (25)$$

In this case, ft and ot indicate the entry to the details, individually, overlooking the door and door. xt is a vector of knowledge and zt is a wrapped graphic. Wand b is separately the weight bar and the word predisposition. Sigmoid is¨and tan is¨h . The oscillation of the part is shrewd. Component - Replication shrewd. As seen, the entries restrict the details, the yield and the condition of the cells and the LSTM may choose to retrieve or ignore data on the intermittent layer along these lines. This gives this template the minimal learning time that is beneficial to our company.

C.Recurrent neural networks based Classification

It is evident that the RNN model is trained in two parts - forward propagation and reverse propagation. The forward propagation calculates the output values and the reverse propagation is responsible for passing through residues collected to update weights that are not fundamentally different from normal neural network training.

The Standard RNN is formalised as follows: given training samples for x poui (i=1,2,...,m), hidden stateshi (i=1,2,...,m), and predictions for I (i= 1,2,...,m). The standard RNN standard is formalised as follows. Whx is a weight matrix input to hidden weight, Whh is the weight matrix hidden to hidden weight matrix, Whyh is a weight matrix hidden from output, and the vectors are bh and through. The e activation function is a sigmoid function and the g classification function is SoftMax.

A function associated with RNNs is defined as f((=f (a), y i), and the distance feature L is a distance feature that measures the predictions y friv the actual label(s) y i. The objective function associated with RNNs for a single training pair (x (i,)yi) is determined as f(f) =L. Let¨be the rate of learning and k be the current number of iterations. Due to the labelsyi sequence I = 1,2,...,m).

4.Results and discussion

The Proposed system is experimented using MATLAB (version 2018a) with 3.0 GHz Intel i3 processor, 1TB hard disc and 8 GB RAM. The validations are carried out to test the effectiveness of each objective are described in the following sub-section of parameter evaluation used in segmentation and classification and performance analysis based on qualitative and quantitative.

4.1.Parameter Evaluation:

The challenge evaluation metrics is used for evaluating the both segmentation and classification performance of our method. For the Classification, the evaluation criteria includeinclude sensitivity (SE), specificity (SP), accuracy (AC), Recall (R) and Precision (P).. The performance criteria are defined is as:

$$SE = \frac{tp}{tp+fn} \quad (26)$$

$$SP = \frac{tn}{tn+fp} \quad (27)$$

$$\mathrm{AC} = \frac{tp+tn}{tp+fp+tn+fn} \quad (28)$$

$$R = \frac{tp}{tp+fn} \quad (29)$$

$$P = \frac{tp+tn}{tp+tn+fp+fn} \quad (30)$$

Where tp, tn, fp and fn denote the number of a true positive, true negative, false positive and false negative.

4.2. Performances analysis

The proposed system performance has validated in numerous ways such, Qualitative and Quantitative which are conversed in the under section.

A.Qualitative Analysis

Table.1. Comparative analysis of the proposed system

Method	SE	SP	R	P	ACC
Fire fly Feature Selection LSTM	88.3333	85.1667	86.2719	88.3333	96.1667
Fire Fly Feature Selection AutoEncoder	79.8750	88.5000	88.5084	79.8750	96.5000
Fire Fly Feature Selection RNN	99.5000	92	92.8509	99.5000	96.7500
HFPS Feature Selection LSTM	99.7500	93.7500	94.4620	99.7500	96.5000
HFPS Feature Selection AutoEncoder	99.5000	92.8333	93.5643	99.5000	**97.2500**
HFPS Feature Selection RNN	100	93	93.6371	100	**97.4375**

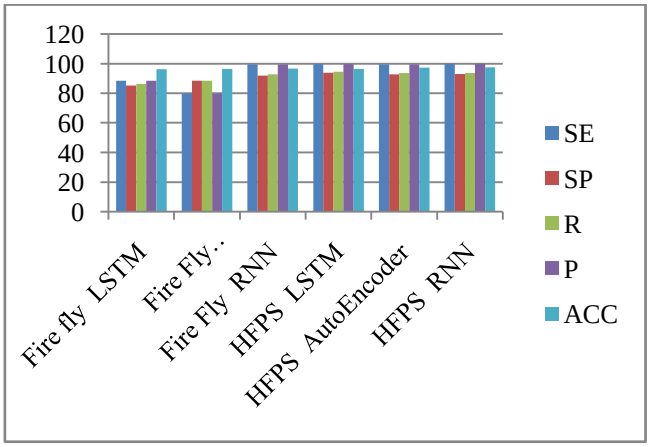

Fig.4. Comparative analysis of the proposed system

In table 1 and figure 4 explained that the comparative analysis of proposed model in different sorts. There are unique and hybrid feature selections are used, also compared with three classifiers. In these result demonstrate that the unique feature selection grade is better than the hybrid feature selection methods. And classification model as RNN achieved the better results than the two classifier models as results of 97.43% accuracy.

B.Quantitative analysis

In this division, the proposed system has compared with various existing systems. The analyses have distinct in table.3.

Table.2. Comparative study of the proposed scheme with various existing systems.

Author	Year	Dataset	Classifier	Accuracy(%)
El-Baz. *et.al.* **(El-Baz 2013)**	2013	LIDC IDRI	CNN	84.32
Aggarwal, Furquan, and Kalra **(Aggarwal 2015)**	2015	--	LDA	97.14
Moffy Vas et al. [19]	2017	Manipal hospital	ANN	92
SurenMakaju et al. **(Makaju 2018)**	2018	LIDC IDRI	SVM	92
Roy, Sirohi, and Patle **(Roy 2015)**	2015	lola11	Fuzzy	94.12
SurenMakajua et al. **(A. Asuntha 2016)**	2016	LIDC IDRI	SVM	87.5
Sangamithraa, and Govindaraju **(Sangamithraa, P 2016)**	2016	--	BPN	90.7
HFPS Feature Selection LSTM	**2020**	**LIDC IDRI**	**LSTM**	96.5000
HFPS Feature Selection AutoEncoder	**2020**		**AutoEncoder**	**97.2500**
HFPS Feature Selection RNN	**2020**		**RNN**	**97.4375**

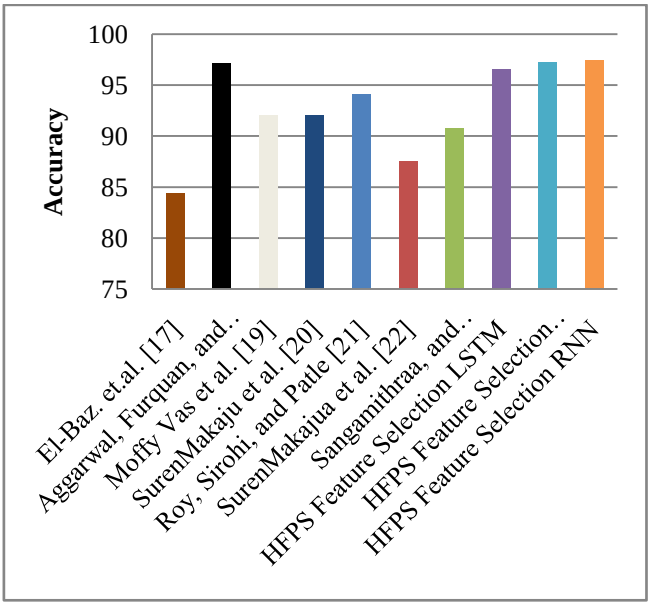

Fig.5. Comparative study of the proposed scheme with various existing systems.

In table 2 and figure 5 signifies the quantitative analysis of proposed model with existing model results. In this study we conduct the seven recent existing techniques to compare with our proposed model. In most of existing studies are conducted the LIDC-IDRI dataset. In El-Baz. et.al. **(El-Baz 2013)** used the classification model of CNN technique to achieve the accuracy of 84.32% by conducted the LIDC-IDRI dataset. In same dataset, SurenMakaju et al. **(Makaju 2018)** andSurenMakajua et al. **(A. Asuntha 2016)** are used, in these author **(Makaju 2018)** used the SVM model to achieve 92% of accuracy and author **(A. Asuntha 2016)** used classification model of ANN to achieved 87.5% of accuracy.However our proposed model, HFPS Feature Selection with LSTM classifier attained the accuracy of 95.50%, then HFPS Feature Selection AutoEncoder

achieved the accuracy of 97.25% and finally HFPS Feature Selection with RNN classifier achevied the accuracy of 97.43%. by conclude these comparision study prove that the our proposed model achieved better results than other existing model and also peak accuracy is attained by LSTM classifier.

5.Conclusion

In this paper, we conducted the simulation study of prediction of lung cancer in earlier stage by introduced the better simulation model. In mostly feature selection is the trustworthy thing in machine learning and deep learning concept. Proper feature selection can provide the better classification results and minimize the computational time period of the model. So we introduced the HFPS Feature Selection method. Another important thing in our proposed model is classification, improper classification can give high mortality rate due to cancer. So we used deep learning concept classification models to analysis the lung cancer more accurately. The proposed classification techniques such as recurrent neural network, Auto Encoder and LSTM. In this study we conduct to analysis the system performance has validated in numerous ways such, Qualitative and Quantitative study, in these study our proposed model can detect and predict more accurately than other model. Further, in these technique is applicable to control the mortality rate of human beings from lung cancer.

Reference

1. American Lung. Association and others, Lung Cancer Fact Sheet. [Last accessed on 2014 Aug 12]. Available from: https://wwwlungorg/lung-health-diseases/lung-disease-lookup/lung-cancer/resource-library/lung-cancer-fact-sheet .
2. National Cancer Institute, NIH, DHHS, Bethesda, MD. Cancer Trends Progress Report. [Last accessed on 2017 Jan]. Available from: https://progressreportcancergov/
3. Nasim F, Sabath BF, Eapen GA. Lung Cancer. Med Clin North Am. 2019;103:463–73.
4. Wiener RS, Gould MK, Arenberg DA, Au DH, Fennig K, Lamb CR, et al. An official American Thoracic
5. Society/American College of Chest Physicians policy statement: Implementation of low-dose computed tomography lung cancer screening programs in clinical practice. Am J RespirCrit Care Med. 2015;192:881–91.
6. Detterbeck FC, Mazzone PJ, Naidich DP, Bach PB. Screening for lung cancer: Diagnosis and management of lung cancer. American College of Chest Physicians evidence-based clinical practice guidelines. Chest. (3rd ed) 2013;143:e78S–92S.
7. Rao RB, Bi J, Fung G, Salganicoff M, Obuchowski N, Naidich D. LungCAD: A Clinically Approved, Machine Learning System for Lung Cancer Detection, in Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2007:1033–7.
8. Fukushima, K., Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol. Cybern. 36(4):193–202, 1980. 9
9. Torbati, N., Ayatollahi, A., and Kermani, A., An efficient neural network based method for medical image segmentation. Comput. Biol. Med. 44:76–87, 2014.
10. Drozdzal, M., Chartrand, G., Vorontsov, E., Shakeri, M., Di Jorio, L., Adrian, A. T., Yoshu, R., Pal, B. C., and Kadoury, S., Learning normalized inputs for iterative estimation in medical image segmentation. Med. Image Anal. 44:1–13, 2018.
11. Zhao, W., Xu, X., Zhu, Y., and Xu, F., Active contour model based on local and global Gaussian fitting energy for medical image segmentation. Optik 158:1160–1169, 2018.
12. Vardhana, M., Arunkumar, N., Lasrado, S., Abdulhay, E., and Ramirez-Gonzalez, G., Convolutional neural network for biomedical image segmentation with hardware acceleration. Cogn. Syst. Res. 50:10–14, 2018.
13. Miao, J., Huang, T.-Z., Zhou, X., Wang, Y., and Liu, J., Image segmentation based on an active contour model of partial image restoration with local cosine fitting energy. Inf. Sci. 447:52–71, 2018.
14. Singh, C., and Bala, A., A DCT-based local and non-local fuzzy Cmeans algorithm for segmentation of brain magnetic resonance images. Appl. Soft Comput. 68:447–457, 2018.
15. Bui, D.T.; Ngo, P.-T.T.; Pham, T.D.; Jaafari, A.; Minh, N.Q.; Hoa, P.V.; Samui, P. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. Catena 2019, 179, 184–196.
16. Ngo, P.-T.T.; Hoang, N.-D.; Pradhan, B.; Nguyen, Q.K.; Tran, X.T.; Nguyen, Q.M.; Nguyen, V.N.; Samui, P.; Tien Bui, D. A novel hybrid swarm optimized multilayer neural network for spatial prediction of flash floods in tropical areas using sentinel-1 sar imagery and geospatial data. Sensors 2018, 18, 3704. [CrossRef]

17. Aydilek, I.B. A hybrid firefly and particle swarm optimization algorithm for computationally expensive ˙numerical problems. Appl. Soft Comput. 2018, 66, 232–249.
18. El-Baz, A., Beache, G.M., Gimel'farb, G., Suzuki, K., Okada, K., Elnakib, A., Soliman, A. and Abdollahi, B., 2013. Computer-aided diagnosis systems for lung cancer: challenges and methodologies. International journal of biomedical imaging, 2013.
19. Aggarwal, Taruna, AsnaFurqan, and KunalKalra. "Feature extraction and LDA based classification of lung nodules in chest CT scan images." In 2015 International Conference on Advances in Computing, Communications, and Informatics (ICACCI), pp. 1189-1193. IEEE, 2015.
20. Vas, Moffy, and AmitaDessai. "Lung cancer detection system using lung CT image processing." In 2017 International Conference on Computing, Communication, Control, and Automation, pp. 1-5. IEEE, 2017.
21. Makaju, Suren, P. W. C. Prasad, AbeerAlsadoon, A. K. Singh, and A. Elchouemi. "Lung cancer detection using CT scan images." Procedia Computer Science 125 (2018): 107-114.
22. Roy, Tanushree Sinha, NeerajSirohi, and ArtiPatle. "Classification of lung image and nodule detection using fuzzy inference system." International Conference on Computing, Communication & Automation, pp. 1204-1207. IEEE, 2015.
    A. Asuntha, A. Brindha, S. Indirani, and A. Srinivasan, “Lung cancer detection using SVM algorithm and optimization techniques,” J. Chem. Pharm. Sci., vol. 9, no. 4, pp. 3198–3203, 2016.
23. Sangamithraa, P. B., and S. Govindaraju. "Lung tumour detection and classification using EK-Mean clustering." In 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), pp. 2201-2206. IEEE, 2016.