Indonesian Journal of Electrical Engineering and Informatics (IJEEI)
Vol. 11, No. 1, March 2023, pp. 248~266
ISSN: 2089-3272, DOI: 10.52549/ijeei.v11i1.4394

VGG19+CNN: Deep Learning-Based Lung Cancer Classification With Meta-Heuristic Feature Selection Methodology

Bhagya Lakshmi Nandipati[1], Nagaraju Devarakonda[2]
[1,2]School of Computer Science and Engineering, VIT-AP University, Andhra Pradesh, India

Article Info

Article history:

Received Dec 28, 2022
Revised Feb 9, 2023
Accepted Mar 12, 2023

Keyword:

Lung cancer
Classification
Optimization
Deep neural network

ABSTRACT

Lung diseases are lung-affecting diseases that harm the respiratory mechanism. Lung cancer is one of the major causes of death in humans internationally. Early diagnosis could optimise survivability amongst humans. This remains feasible to systematise or reinforce the radiologist for cancer prognosis. PET and CT scanned images can be used for lung cancer detection. On the whole, the CT scan exhibits importance on the whole and functions as a comprehensive operation in former cancer prognosis. Thus, to subdue specific faults in choosing the feature and optimise classification, this study employs a new revolutionary algorithm called the Accelerated Wrapper-based Binary Artificial Bee Colony algorithm (AWBABCA) for effectual feature selection and VGG19+CNN for classifying cancer phases. The morphological features will be extracted out of the pre-processed image; next, the feature or nodule related to the lung that possesses a significant impact on incurring cancer will be chosen, and for this intention, herein AWBABCA has been employed. The chosen features will be utilised for cancer classification, facilitating a great level of strength and precision. Using the lung dataset to do an experimental evaluation shows that the proposed classifier got the best accuracy, precision, recall, and f1-score.

Corresponding Author:

Bhagya lakshmi Nandipati,
School of Computer Science and Engineering,
VIT-AP University,
Andhra Pradesh, India.
Email: bhagyalakshmi.nandipati@gmail.com

1. INTRODUCTION

Lung diseases' escalation in the contemporary industrial world increases the requirement for the latest methodologies for precise and advanced prognosis. Amidst Lung diseases, lung cancer yet remains identified as one of the very perilous cancers. Cancer refers to the atypical development and multiplication of cells within the body. Entire cancers possess an unrestrained development form and a characteristic to separate out of the source and metastasize [1]. A typical lung cell might turn into a lung cancer cell for nil distinct cause, yet in maximal cases, the transition remains a consequence of continual exposure to carcinogens like tobacco and alcohol. The emergence and functioning of cancer cells remain disparate from typical cells. A mutation or modification in the DNA or cell's genetic material happens [2]. DNA remains accountable for managing the semblance and functioning of cells. If a cell's DNA modifies, that cell distinguishes from the healthful cells subsequent to it, and nevermore does the body's typical cells. This modified cell detaches out of its nearby cells and in no way understands when it must cease developing and perish [3]. That is to say, the modified cell in no way ensures the inward orders and indicates that the rest of the cells remain accountable and functions indiscriminately rather than cooperating with the rest of the cells [4].

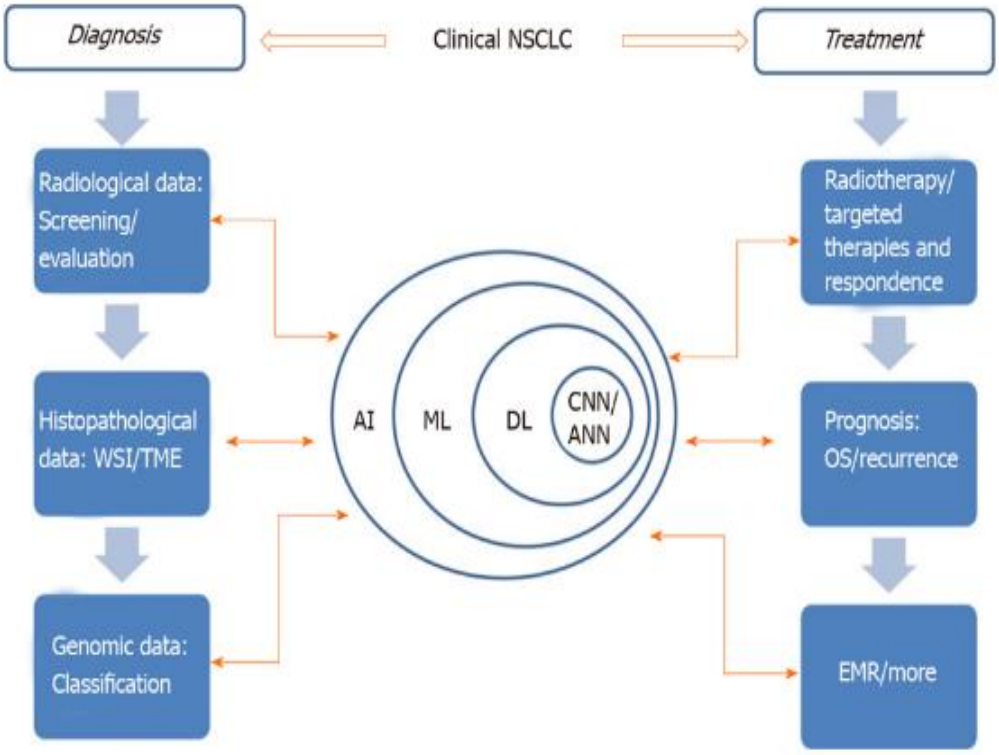

Figure 1. The application of artificial intelligence involved in clinical non-small cell lung cancer

AI can improve patients' treatment results, ameliorate patients' treatment processes, and even mend medical management. Because of the increasing application of AI in lung cancer treatment (Figure 1), this paper reviews AI applications being developed for NSCLC detection and treatment as well as the challenge facing clinical adaptability.

One-third of entire cancer demises occur because of lung cancer. Around eighty percent of sick persons possess 5 years remaining in their finest scenario subsequent to being detected with this cancer kind [5]. Centered on a survey by the American Cancer Society, Lung Cancer in males and females remains the second most common cancer in the USA [6]. About 228,820 novel cases and about 135,720 LC demises, centered on ACS stats, would happen [7]. Air pollution because of the industrialization of urban areas, tobacco usage, and genetic factors remain the chief reasons for such illnesses [8]. Lung disease's early prognosis would possess a chief influence upon the feasibility of the illness' definite medicament.

Chief prognostic methodologies for lung cancer incorporate radiographic imaging and computed tomography, biopsy, bronchoscopy, and testing of cells in the sputum. Meantime, the CT scan imaging methodology remains vastly employed as a dominant prognostic methodology. In this methodology, the clinician investigates feasible nodules in the images. A pulmonary nodule remains a little, circular, opaque mass, which creates within the lung tissue [9]. That is to say, nodules remain spheroid radiographic opacities below 30 millimeters in diameter.

Previously, lung disease has been detected centered upon the assistance of specialists' eye capability with nil employment in computer science. Nevertheless, lately, centered upon the disparate imaging approaches centered upon computer science and artificial intelligence, the prognosis could be additionally accurate. In the majority of such methodologies, subsequent to catching the images out of the sick person, disparate image processing methodologies are executed to detect tumors [10].

Feature selection remains the chief approach for lessening the magnitude decrement, which possesses a general aim, i.e., to lessen the quantity of dangerous, unnecessary, and noisy features in a dataset for easy and quick data processing intentions. The general methodology of lessening the data's magnitude to be assessed remains to lessen the features' or variables' quantity to a very governable quantity. Feature Selection or variable selection comprises lessening the existing features to a set, which remains optimal or sub-optimal and is able of generating outcomes that remain equivalent or finer to that of the initial set [11].

Lessening the feature set measures down the data's magnitude which consequently lessens the training duration of the algorithm choosing and computative price, enhances the classification paradigm's precision, and turns the data mining outcomes simpler to comprehend and very practical. The execution of maximum classification Algorithms could be enhanced by lessening the feature set, particularly for KNN Algorithm; this

might as well reduce the decision trees' accuracy. Decision trees possess the ability to lessen the initial feature set in the tree-constructing procedure, starting this procedure with lesser features that might influence the last execution. Administered learning will be employed in most of the classification issues in which the fundamental class probabilities remain unfamiliar, and every instance will be related to a class label. An impertinent feature influences the learning procedure and an unnecessary feature in no way includes anything novel to the objective concept. In several classification issues, it remains hard to learn fine classifiers prior to eliminating such redundant features because of the data's big dimension [12]. Data augmentation has resolved this issue in the neural networks to some extent, but this increases the size of the dataset that will again need more storage capacity and computation time. Keeping these issues in view, there is still a demand for handcrafted feature extraction methods. Deep learning extracts low-level features that help to acquire the best results, but these are not enough for image classifcation. Therefore, the system uses a fusion of features extracted using a pre-trained model of deep learning, i.e., VGG19, and various handcrafted feature extraction algorithms, i.e., SIFT, SURF, ORB, and ShiTomasi corner detector for image classification [21]. The learning algorithm's functioning duration could be efficiently lessened by lessening the impertinent/unnecessary features' quantity that produces an additional common classifier. This aids in obtaining a finer perception of the fundamental notion of actual-life classification issues. Feature selection chiefly impacts the classification's training stage. Subsequent to producing features, rather than processing data with the entire features to the learning algorithm straightly, Feature Selection for classification would initially execute Feature selection to choose a feature subset, and, later, process the data with the chosen features to the learning algorithm. The FS stage may remain individual of the learning algorithm, such as filter paradigms, or this might repetitively employ the Las execution for analyzing the chosen features' quality, such as wrapper paradigms. With the lastly chosen features, a classifier will be instigated for the anticipation stage.

The following are the contributions of this study:

- The cancer classification will be executed by employing the VGG19+CNN, which will be calibrated ideally by employing the proffered Accelerated Wrapper-based binary Artificial Bee Colony algorithm (AWBABCA) that is the amalgamation of the wrapper within the Ant bee colony algorithm.
- Pre-processing will be performed by implementing adaptive filters having resolution enhancement. Next, the pre-processed data's magnitude will be lessened by employing a Spatial Pyramid Coupling network in order that the size lessened and vital features will be supplied to the classifier for additional procedure.

This study has been organized as: Segment 2 emphasizes a few existing research works, Segment 3 presents the proffered approach and methodologies, Segment 4 highlights the experimental outcomes and discussion, and, finally, Segment 5 concludes with a conclusion and prospective study.

2. RELATED WORKS

This segment addresses similarly associated literature and conveys in what way these remain disparate out of this proffered methodology. Normally, similar to this study, many prevailing works to identify lung cancer employ deep-learning methodologies. The review encourages scholars in applying a novel paradigm for cancer classification. Lastly, the adversities of this study have been listed comprehensively.

The study [13] introduces a new hybrid binary version of the Enhanced Chaotic Crow Search and Particle Swarm Optimization algorithm (ECCS-PSOA) for resolving feature selection issues. In this, for guiding the feature space, the authors hybridized the Cuckoo search Algorithm optimized variant that possesses a finer search strategy and PSO that remains able of advancing into the finest global solution in the search space. The authors additionally embed an opposition-related learning approach in the hybrid induction algorithms' local search.

The study [14] employs a wrapper-related FS paradigm for the same intention. Over the past few years, the Grasshopper Optimization algorithm (GOA) has confirmed its supremacy over the rest of the optimization algorithms in disparate study disciplines. This study proffers an enhanced variant of GOA named LAGOA that employs Learning Automata for calibrating GOA's criteria in an adaptable manner and 2-stage mutation for optimizing the algorithm manipulation ability. Learning Automata has been employed for calibrating the criterion value of every grasshopper in the populace separately. In 2-stage mutation, the initial stage lessens the chosen features' quantity when sustaining improved classification accuracy. The next stage includes pertinent features that enhance classification accuracy.

The study [15] proffers a technique in which the image classes have been trained out f scratch with the DarkNet-19 paradigm that remains one of the DL paradigms. In the FSt excerpted out of the DarkNet-19 paradigm, choosing the ineffectual features has been executed by employing the Equilibrium and Manta Ray Foraging optimization algorithm. Next, an array comprising the ineffectual features has been differentiated from the other FSts generating an effectual feature set (corresponding rule insets). The effectual features

acquired by the dual-employed optimization algorithm have been joined and classified with the Support Vector Machine methodology.

The study [16] provides a new enhancement technique for finer FS in clinical data classification in this study. This technique is named a centroid mutation-related and Rescue optimization algorithm (CMSROA) centered upon a K-Nearest Neighbour (KNN) classifier for illness classification. The employment of CMSROA in feature selection remains to seek the optimum FTs, which exhibit robust separation betwixt 2 classes resolving early convergence and enhancing SAR AG's LS capability. The authors employ fuzzy logic as a logical system that remains the multi-valued logic's expansion for producing a fuzzy set and implement a centroid mutation operant upon this.

The study [17] presents Harris Hawk Optimization (HHO) for resolving disparate requiring augmentation works as a meta-heuristic tool. The first HHO remains for dealing with augmentation issues in a continual atmosphere, yet FS remains an augmentation work in binary space. Hence, this applies a Multi-Objective Quadratic Binary HHO (MOQB-HHO) approach alongside the KNN methodology as a wrapper classifier to excerpt the optimum FSSs. Lastly, this work employs the crowding distance value as a tertiary parameter for selecting the finest out of the non-dominating resolutions. In this, for assessing the proffered technique's execution, 12 benchmark datasets have been regarded.

The study [18] proffers a 2-phase improved grey wolf optimization (IGWO) algorithm for feature selection on large-sized data. In the initial phase, a multilayer perceptron (MLP) network alongside group lasso regularization terms (GLRT) has been initially trained for building an integer optimization issue by employing the proffered algorithm for pre-feature selection and hidden layer framework's augmentation. The dataset has been condensed by employing the FSS attained in the initial phase. In the next phase, an MLP NL alongside GLRT has been retrained by employing the condensed dataset, and the proffered algorithm has been used for building the discrete optimization issues for feature selection. Meantime, a swift assessment scheme has been built to alleviate the assessment price and enhance the assessment effectiveness in the feature selection procedure.

The study [19] contemplates the KNN approach wherefore a genetic algorithm has been implemented for the effectual feature selection for lessening the dataset sizes and optimizing the classifier speed that has been used to prognosis the phase of a sick person's illness. For enhancing the proffered algorithm's accuracy, the finest value for K has been discerned by employing an experimental process.

In the study [20], the Pre-Processing phase has been executed by employing the fast non-local means filter. For the segmentation procedure, the Masi entropy-based multilevel thresholding by employing the salp swarm algorithm has been employed for segmenting the cancer nodule out of the lung images. By employing the grey-level run length matrix, disparate features have been mined in the feature extraction. The binary grasshopper optimization algorithm has been implemented for choosing the optimal features for the feature selection procedure. Next, the chosen features have been classified by employing the hybrid classifier (HC) called deep neural network (DNN) with adaptive sine cosine crow search algorithm (DNNASCCSA). The proffered hybrid classifier precisely identifies lung cancer. This proffered DNNASCCSA has been applied by MATLAB by employing the Lung Image Database (DB) Consortium and Image DB Resource Initiative datasets.

The adversities of the methodologies proffered in the literature include the following:

Although several enhancements have been discussed in lung cancer classification, until now few adversities related to machine learning (ML) methodologies prevail. The primary adversity related to ML remains about the large-sized data, which has produced a myriad couple of hundred thousand features to train the classifier. The secondary adversity remains about the dimension of the publically prevailing dataset having samples of lesser quantity. The tertiary adversity remains about the unlabeled data (UD) existence. The labeled data comprises only a fraction of the entire dataset, and, simultaneously, such UD's enormous values comprise copious data to classify cancers in order that acquiring this data needs deeper excerption stages. Such adversities turn Cancer classification into frenzied work.

3. PROFFERED METHODOLOGY

The comprehensive system framework for classifying lung cancer has been illustrated in Figure 2. At first, PET and CT lung cancer images have been supplied as input ensued by Pre-processing by employing adaptive filters having resolution enhancement. Next, the pre-processed images' output has been supplied to the Spatial Pyramid Coupling network for Feature Extraction. The excerpted features have been merged by employing feature maps (FMs) ensued by the Accelerated Wrapper-based binary Artificial Bee Colony algorithm (AWBABCA). Lastly, VGG19+CNN has been employed for classification.

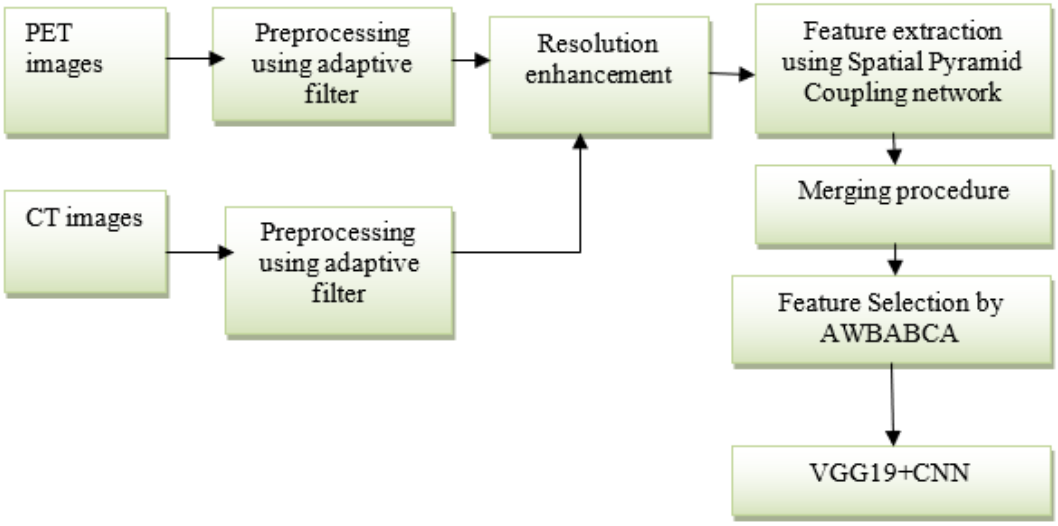

Figure 2. Comprehensive system framework for LC classification

4. DATASET DESCRIPTION

This dataset consists of PET-CTSs for sick persons having biopsy-proven NSCLC. As the objective remains to assess the thorax, the imaging expert selected portrayal subjects, which incorporate sick persons having solitary lung primary tumors alongside or devoid of hilar (Phase 2) and mediastinal (Phase 3) nodal participation and these were the tumors included in the mediastinum and the chest wall. Such subjects were chosen over subsequent subjects over 3 months from the imaging archive of the Department of Molecular Imaging at the Royal Prince Alfred Hospital in Sydney, Australia. A Biograph 128-slice mCT has been used to obtain the entirety of research projects (PET-CT scanner; Siemens Healthineers, Hoffman Estates, I1, USA). This CT is still a high-resolution tomography with time-of-flight, HD rebuilding, and flow motion characteristics.

A CT volume and a PET volume are both required for every study; the CT resolution has been 512 512 pels at 0.98 mm x 0.98 mm, and the PET resolution has been 200 200 pels at 4.07 mm x 4.07 mm with a slice thickness and an interslice distance of 3 mm. With a similar number of slices, the two volumes have been recreated. Between one to several tumours, including those in the thorax, have been the subject of studies. Diverse lung lobes, the mediastinum, and hilar nodes are all integrated by the tumour placements. Every piece of data has been re-detected. The images have been resized to a resolution of 256 x 256 pels (x-y axes), resizing the PET and CT volumes, and allocating a similar coordinate space, which continues to be a standard procedure for PET-CT data assessment.

The PET images have been normalized by a transition to standard uptake values (SUVs). This SUV remains a semi-quantitative value of FDG's degree obtained by the tumor's locations associated with isotope dose and sick person's weight. The thorax subvolume (an array of two-dimensional thorax slices) of every study has been physically detected (mean: 86.5 slices, standard deviation: 6.8 slices). For ensuring a balanced class dispensation for CNN training, the authors employed the actuality for choosing just the axial thorax slices comprising entire ROIs (lungs, mediastinum, and tumors) generating the last DS of 855 PETCT slice couples (855 CT and 855 PET). The sampling remains a conventional scheme to learn from unbalanced data. The actuality has been inferred out of the prognostic imaging report that elaborated the primary tumor's positions and whatsoever incorporated thoracic lymph nodes.

A single, highly qualified image expert who has read more than 80,000 PET and PET-CTS scans have completed entire reports. We used the findings from the report to guide a semi-automated process for ROI labelling. To remove the lung actuality from the CT volume, a commonly used adaptive thresholding algorithm has been implemented. Likewise, thresholding has been linked to roughly discerning the mediastinum. The tumor's ground truth has been excerpted by employing forty peak SUV linked thresholding for detecting the 'hot sports' that are detected in the prognostic reports that remain a normal disconnection, that remains comprehensively employed when scrutinizing tumors in PET images. Minimum physical alterations have been executed for supporting initiating excerption of the real ROIs (for instance, avoiding the left and right lung areas out of remaining linked together with the edge of mediastinum). Figure-3 shows sample images from the dataset.

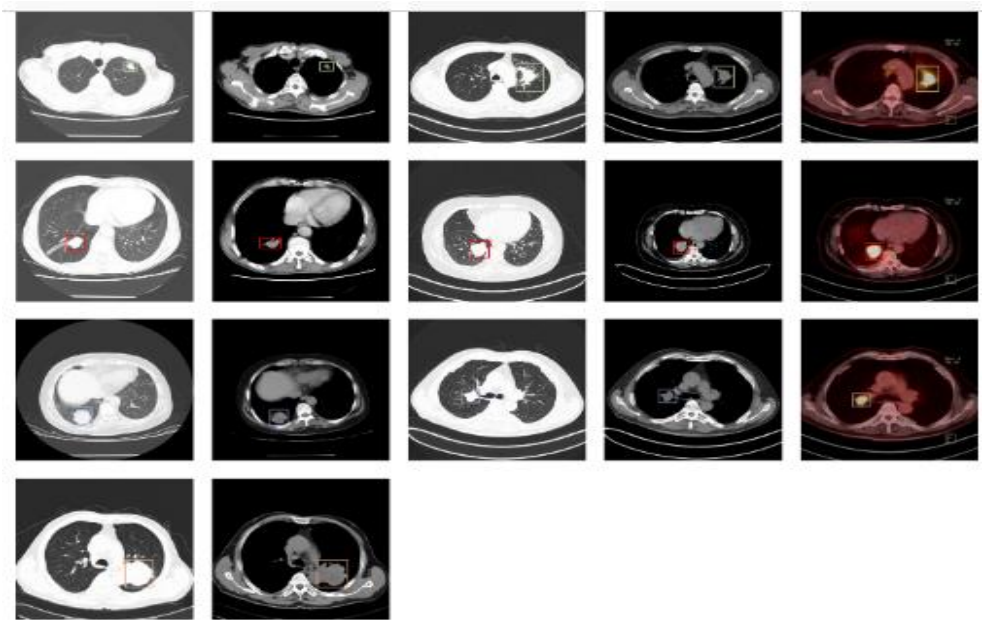

Figure 3. sample images from the dataset

5. PRE-PROCESSING BY EMPLOYING ADAPTIVE FILTER

The PP phase remains a vital component in the image assessment scheme. This could optimize the initial image and lessen noise or redundant particulars.

1. Histogram equalization (HE) – The methodology expands the pixel's intensity range from the initial range to 0 to 225. Hence, the optimized image possesses a broader range of intensity and somewhat greater contrast.
2. HE + Gaussian blur (GB) – The filter lessens a few noises and redundant particulars, which could remain ambiguous for the neural network (NN); this filter's kernel dimension has been experientially fixed to 5 × 5.
3. HE + bilateral filter – The filter as well lessens a few noises and redundant particulars, which could remain ambiguous for the NN, yet its chief attribute remains in sustaining edges; the filter's experimental setup criteria are *diameter* = 5, $\sigma_{color} = \sigma_{space} = 75$.
4. Adaptive masking (AM) – In the proffered PP methodology, initially, the maximal (max) and minimal (min) pels' intensity are noticed; later, the binary thresholding employing the threshold conveyed in the equation is implemented. The subsequent phase employs morphologic closing. It generates the adaptive mask, which, subsequent to the bitwise procedure, eliminates the diaphragm out of the source image.
5. AM + HE + GB – The methodology combines AM with HE and GB.

$$threshold = min + 0.9.(max - \min)$$

6. RESOLUTION ENHANCEMENT

The terminology super-resolution (SR) indicates an approach where many low-resolution (LR) points of views (POVs) have been joined repetitively for acquiring a high-resolution image (HRI). In the Irani and Peleg formulation of an SR AG, the original anticipation of the HRI $f^{(0)}$ could be centered upon the mean of the upsampled acquisitions transformed to a common reference frame (CRF):

$$f^{(0)} = \frac{1}{K}\sum_{k=1}^{K} T_k^{-1}(gk \uparrow s) \quad (1)$$

in which gkindicates one of the *K* acquisitions, T_k^{-1}indicates the geometric transition (GT) to a CRF, and $\uparrow s$ indicates the upsampling operant from LR to the HR portrayal. The LR-calculated data gkcan be acquired out of the "true" image f when the acquisition system has been sufficiently designed. This procedure will incorporate transferring the image f towards the *k*th POV, blurring to rationalize finite system resolution, downsampling (DSp) to the system's sampling rate, and including noise. For a provided anticipation of the image $f^{(n)}$, the LR data (LRD) can be designed:

$$g_k^{(n)} = (T_k(f^{(n)}) * \mathrm{h}) \downarrow \mathrm{s} \quad (2)$$

in which $* h$ denotes the blurring procedure having the kernel h, and $s \downarrow$ denotes the DSp operant that makes a mean of the pels to the LR. The noise equation has been removed. The kth acquisition's initial GT out of the CRF remains T_k. It remains normally the manual transfer betwixt the object and the image out of the initial location. For acquiring finer anticipation of the image f, the former anticipation of HR image $f^{(n)}$ has been rectified by the variation betwixt the LRD g_k and the expression g˜(n)k, which portrays how the LRD could remain if the anticipation $f^{(n)}$ remains right. The subsequent iteration $f^{(n+1)}$ of an HR anticipation remains:

$$f^{(n+1)} = f^{(n)} + \frac{1}{k} \sum_{k=1}^{k} T_k^{-1}(((g^k - g_k^{(n)}))) \tag{3}$$

In this, the variations betwixt g_k and g˜(n)k have been upsampled, $\uparrow s$ attains the lesser SR pel dimension shifted to a CRF $T^{-1}{}_k$ and made mean above K acquisition. The sign $*p$ remains a sharpening kernel.

7. FEATURE EXTRACTION AND FUSION PROCEDURE

Subsequent to image PP for removing noise within the image, features have been excerpted out of these images for detecting and scoring probable cancers. FE remains one of the vital phases in analyzing CT images (CTIs). The feature might be excerpted at CTIs' cell level or tissue level for optimized estimations. For finely obtaining the morphological data, the present study takes into account area, brightness, diameter, perimeter, elongation, and others as features for imitating nuclei unevenness within CTIs. Cellular feature level points upon quantifying independent cells' characteristics devoid of calculating spatial dependency amidst these. In CTIs, figure and morphological attributes have been taken into account for extraction.

SPC module (SPCM) chiefly employs the spatial association betwixt pels within the image for producing spatial FMs (SFMs). As the convolutional layer (CL) just has an object's translation in the image invariance, and the object's scale transition could not be processed, hence, while excerpting the SFM, this study includes the spatial pyramid (SP) transition for excerpting the multiple-scale (MS) SFM. The SPCM remains equal in executing feature convolution out of the bottom up upon the image's FM and, later, merging MSs' FMs.

The SPCM employs the FM acquired by the channel coupling module as the IP FM. Initially, the convolution procedure within the SPCM separates the IP features and, later, employs SP pooling (SPP) for executing abstract features' MS transition for acquiring 4 features of disparate dimensions. The disparate scales FMs have been acquired via MS convolution subsequent to SPCM. Lastly, the 4 convolution feature outcomes have been executed via the channel concatenation procedure for acquiring the last SP FM. The SPCM's computation procedure remains:

$$F' = Cs(F) = Concat(Conv\left(SPP\left(Conv(F')\right)\right)) \tag{4}$$

where SPP () portrays the SPP procedure. Particularly, the SPP procedure remains to separate the initial IP features by employing 4 disparate pooling scales(16, 4, 2, 1) or the merging of PET and CT features – the easiest fusion procedure (FnP) remains to linearly add these. Nevertheless, this procedure disregards the relationship betwixt neighboring pels within the area and remains without the expression of the image's comprehensive data.

For finely optimizing the merged image features'(IFs) display capability devoid of missing the initial data in PET and CT, the present study employs the cross-correlation layer (CCL) for merging the features excerpted by the encoder. Provided the Ifs F_{ct} and $F_{pet,}$ the merging outcomes would be computed as per the CCL. The particular formulation remains:

$$F_{pc} = Cor(F_{ct}, F_{pet}) \tag{5}$$

in which Cor() portrays the CCL, and the CCL remains a specialized CL, which employs a cross-correlation (C-C) procedure. Disparate out of the common convolution function, the C-C function remains a procedure betwixt 2 data. If the data has been transitioned within the NW, the weight required to be trained remains unneeded. The computation expression remains:

$$Cor(x_1, x_2) = \sum_m \sum_n x_1(m,n) * x_2(m+o, n+o) \tag{6}$$

in which $\boldsymbol{x_1}, \boldsymbol{x_2}$ portrays the feature patch upon the 2 FMs $\boldsymbol{F_{ct}}$ and $\boldsymbol{F_{pet}}$ accordingly, m, n portrays the patch's dimension, and o portrays the patch stride – this study fixed o as three. Every patch upon the FM $\boldsymbol{F_{ct}}$ should

be cross-correlated with entire patches upon the rest of the FM$\boldsymbol{F_{pet}}$. For acquiring the OP of a similar dimension as the initial image, the padding pattern chooses the 'SAME' while executing C-C.

By executing a three-dimensional convolution devoid of padding the modality size, for a provided channel c, an FM having a singleton 3D (S3D) is acquired in which the value at the position (x, y) has been discerned out of the neighborhood of $F_{CT}(x, y)$ as well as $F_{PET}(x, y)$

$$(W_{multi} * X'_{multi})(x, y) = \sum_i \sum_j \sum_l (W_{multi}(i, j, l).X'_{multi}(x - i, y - j, l) \quad (7)$$

Next, the S3D has been squeezed for attaining an OP FM ($W_{multi} * X'_{multi}$of dimension$w \times h \times 2c$, an equal breadth and length such as the 2 modality-specific (MSp) IP FMs $F_{CT}(x, y)$ and $F_{PET}(x, y)$, and twice the channels' quantity that remains vital for the weighing of MSp FMs by the co-learnt fusion maps (CLFMs) as explained under. The objective remains that the CLFM manages the rate of significance provided to the data out of every modality at every position, opposing the global fusion proportion in PET-CT pel fusing. Hence, the CLFMs straightly influence the IP dispensation of the learnable layers, which instantly ensues the co-learning module. The FnP incorporates the MSp FMs as per the values (coefficients) within the multiple-modality fusion map (MMFM) by,

$$\boldsymbol{F_{fused}} = \boldsymbol{F_{fusion}} * (F_{CT} + F_{PET}) \quad (8)$$

in which $\boldsymbol{F_{fused}}$ denotes the fused co-learntFM, + denotes the stacking procedure, and $*$ denotes a component-wise multiplication. The procedure fuses the 2 MSp FMs F_{CT} and F^ F_{PET} and weighs these by the co-learnt MMFM F_{fusion} just like pel fusing. For the NW's loss function (LF), this study chooses the structural similarity (StS) loss (6, 27) as LF. Subsequent to acquiring the merged PET-CT, this would initially compute the StS alongside the initial PET and CT accordingly as,

$$SSIM(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} . \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (9)$$

where, x and y portray the fused image and the initial image accordingly, μ portrays image average, σ_x, σ_y portray image variance, and $\boldsymbol{\sigma_{xy}}$ portrays images' co-variance. $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$ portray the constants (evade the denominator remaining zero), and the computing formulation remains $\boldsymbol{C_1 = k_1 L^2, C_2 = k_1 L^2}$ in which L indicates the image's grayscale modification. As this study executes FE, the image has been normalized. Thus, L remains one. *K1* and *K2* portray 2 constants having the default value of 0.01 and 0.03 respectively. The NW's loss could be described by,

$$L_{ssim}(x, y) = 1 - SSIM(x, y) \quad (10)$$

While employing SSIM, there would be an issue with edge noise flaws. Hence, this study presents an *L1* regularization term (L1RT) in the initial LF. The L1RT kωk could be employed for anticipating the variation betwixt the target value x and the anticipated value y that could efficiently lessen the noise within the image, possessing a particular rate of strength, and could as well evade the NN out of over-fitting (OFt) while doing training.

8. FEATURE SELECTION BY EMPLOYING AWBABCA

In MH ABC AG, bees travel within SS inside the actual and continual domain. Bees are of 3 types – onlooker bees (OBs), scout bees (SBs), and employed bees (EBs) that search for a food source (FdS). As the FS procedure addresses tracking for the binary domain, this remains requisite to embrace discrete space when revising neighborhood and candidate solution (CdS). Consider 'n' indicates features' quantity associated with 'm' objects. When feature j has been selected for object i, ten has been allotted; otherwise, 00 has been allotted. The encoding portrays FdS as the dimension of matrix m, n. Hence, AWBABCA includes the following stages:

1. Fix principal values for a total of OBs, a total of EBs, maximal iterations quantity (IQ), and deterministic IQ for SBs.
2. Enable FdS' populace (probable FSt) randomly allot this to EBs by binary solution.
3. Compute the original FdSs' populace by ensuing the objective function (ObF):

   SBs: Leave FdS of EB, its current FdS fitness (FdSF) has not optimized inside the deterministic IQ and produces the correlating for SB and generates novel FdS (haphazardly). Study the best solution obtained till now and reiterate from phase three until a deterministic ending scheme has been fulfilled.

EBs: Build a novel CdS generated in the existing solutions' neighborhood by exchanging 2 selected objects haphazardly (total column within the matrix) within current solutions (for entire EBs). The novel subset (SSt) has been produced by employing the following equation:

$$V_{ij} = X_{ij} + \phi_{ij}\left(X_{ij} - X_{kj}\right) \tag{11}$$

Update EBs: For every EB, when (FdS) CdS remains supreme to its present one; later, replace this with CdS.

OBs: For OBs, select one of the EBs' FdSs probably centered upon the next pattern. This generates a novel solution for each OB by exchanging whatsoever 2 haphazardly selected columns within the selected FdSs. When novel FdSF remains above the existing FdS, modify EB's position to the new FdS. Otherwise, by no means alter EB's FdS.

- **Wrapper subset analysis (WSA)**

WSA remains normally called Wrapper that remains a type of FS methodology incorporated with ABABC AG for generating AWBABCA. The methodology has been employed for analyzing the SSt value produced by the EB within the framework of the analysis function (AsF). The AsF of each feature would provide anticipated quality for shaping LA. Hence, this leads to the best accuracy calculation. In this, the k-folds cross-validation procedure has been executed by employing this methodology for generating supreme accuracy.

- **Pertinence parameters**

The candidate feature's integrity has been analyzed by the analysis action $(classifier, Data^{(SwagU(C))}$, that trains and substantiates the classifier by employing a fivefold CV (FCV) above the DS. The data estimated above $S_{new}U\ (c)$(C remains the target class). Instead of employing FCV's mean accuracy (MA) and performing a t-test above the FCV outcomes proffered formerly, this study embraces the ensuing parameters:(i) an FCV has been utilized for determining, in any case, a novel feature (NF) has been included to the chosen FSS S, and(ii) the NF f has been added solely when the FCV's MA above $Data^{(S\,U\,f\,U\,c)}$remains finer than that of the FCV above $Data^{(S\,U\,f\,)}$and not less than MinFoldersBetter (mf) of the 5-folds functions well. MinFoldersBetter (mf) indeed remains a counter for registering the number of times the 4 classification accuracies have been acquired out of the FCV above $Data^{(S\,U\,f\,U\,c)}$that remains finer than the FCV's MA above$Data^{(S\,U\,f\,U\,c)}$. This technique evades the evaluation of employing an analytical test with a little sample dimension. For finer noise management and Oft in the FS, the suggested empirical values for mf remains two or three. For the wrapper-related SFS methodology, the parameter remains that the NF f has been added solely when the FCV's MA above $Data^{(SUfUc)}$remains finer than that of the FCV above $Data^{(SUfUc)}$. This study employs 2 ObFs - $fit1$ and $fit2$. The initial ObF F1 has been employed to seek the feature's quantity (that is, the quantity of 1's). This proffered initial ObF remains:

$$fit1_{(v)} = \frac{N - O_v}{N} \tag{12}$$

The next ObF F2 determines the range to which the feature could identify amidst the object pairs (ObP).

$$fit2_{(v)} = \frac{R_v}{c1 * c2} \tag{13}$$

in which *v* indicates the selected FSSs, O_v indicates the quantity of 1's in *v*, *C1* and *C2* indicate the objects' quantity in every class, and R_v indicates the ObP' quantity (that is, rows within the distinction table [DTb]) where *v* could determine betwixt. The ObF *Fit1* provides the candidate credit for consisting of low features' quantity in *v*, and *Fit2* discerns the range to which the candidates could determine amidst ObP in the DTb. Figure 4 shows the flow diagram of Binary Artificial Bee Colony Algorithm

Input: Population size SN, Maximum number of iterations NMAX.
Output: The optimal individual xbest, the best fitness value f(xbest).
Initialize the population by using Equation (1).
Evaluate the fitness value of each individual.
For it = 1 to NMAX do For i = 1 to SN do
Select a different food source xk at random.
Produce a new food source according to Equation (2) and map it to discrete values by Equation (4).

Evaluate the fitness value of each food source. Update xi according to greedy selection.
End
For i = 1 to SN do Update the position using operators and map it to discrete values by Equation (4).
Evaluate the fitness value of each individual.
End
End
Output xbest and f(xbest).

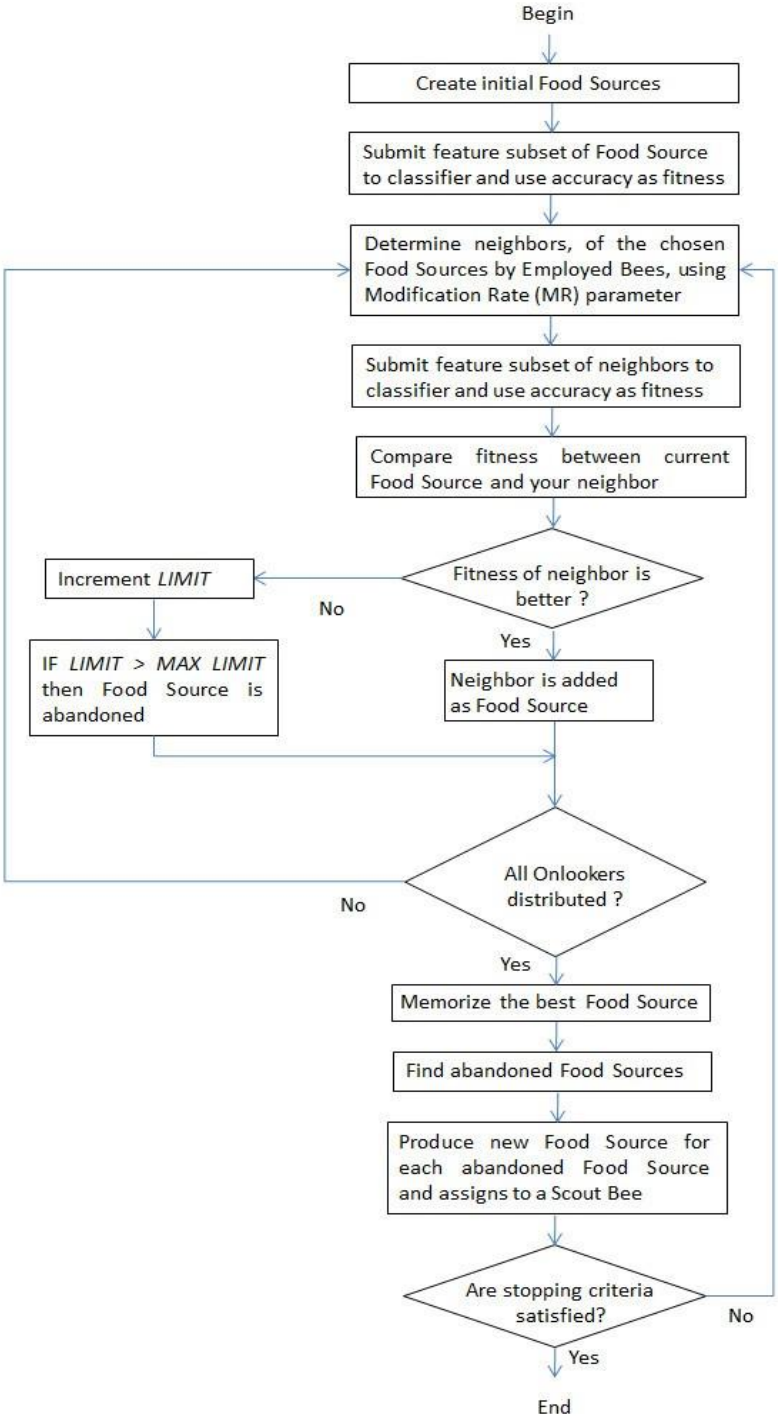

Figure 4. Flow Diagram of Binary Artificial Bee Colony Algorithm

9. CLASSIFICATION PROCEDURE

The proffered VGG19+CNN classifier is modeled on the pre-trained paradigm VGG19 ensued by CNN's NW for classifying the multi-chest illness by employing chest radiographs (RGs). Normally, CNN

remains centered upon 3 different layer types – CL, pooling layer, and completely linked layer. The VGG19+CNN paradigm particulars have been illustrated in Figure 5.

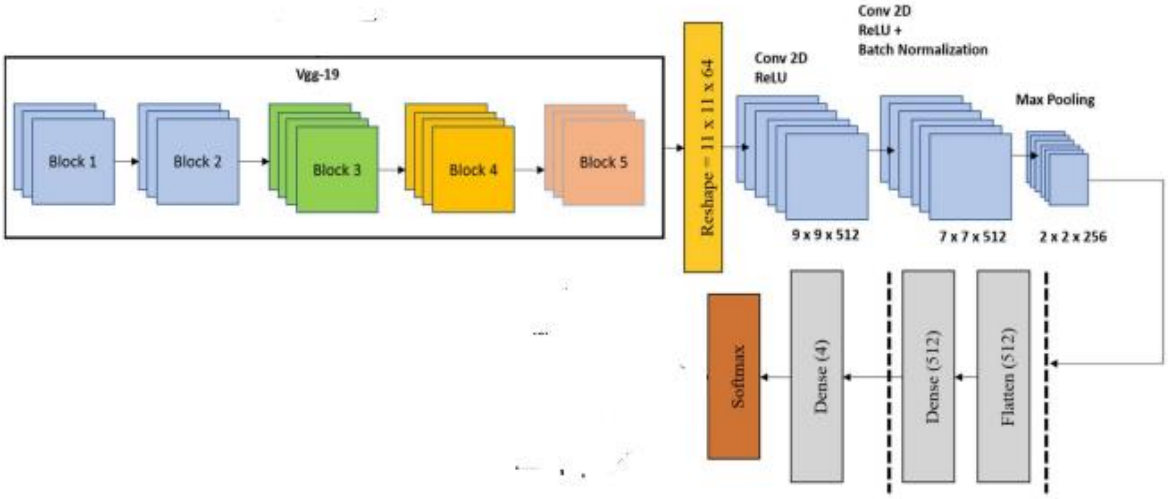

Figure 5. VGG19+CNN classifier's framework

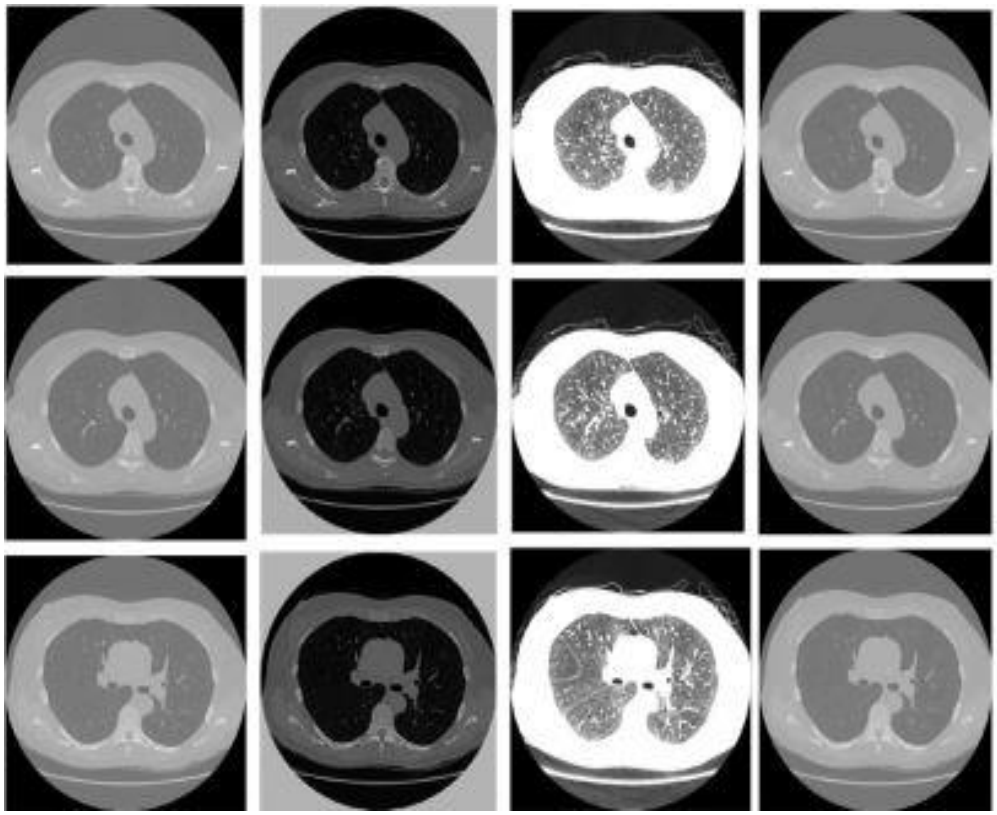

Figure 6. Input image

Figure 6 shows RGB channels remain in the input image (II). This paradigm's initial layer remains CL. This layer begins the procedure by employing filters that are as well called kernels. The kernel dimension remains centered upon 2 values:

$$filtersize(fs) = f_w * f_h \quad (14)$$

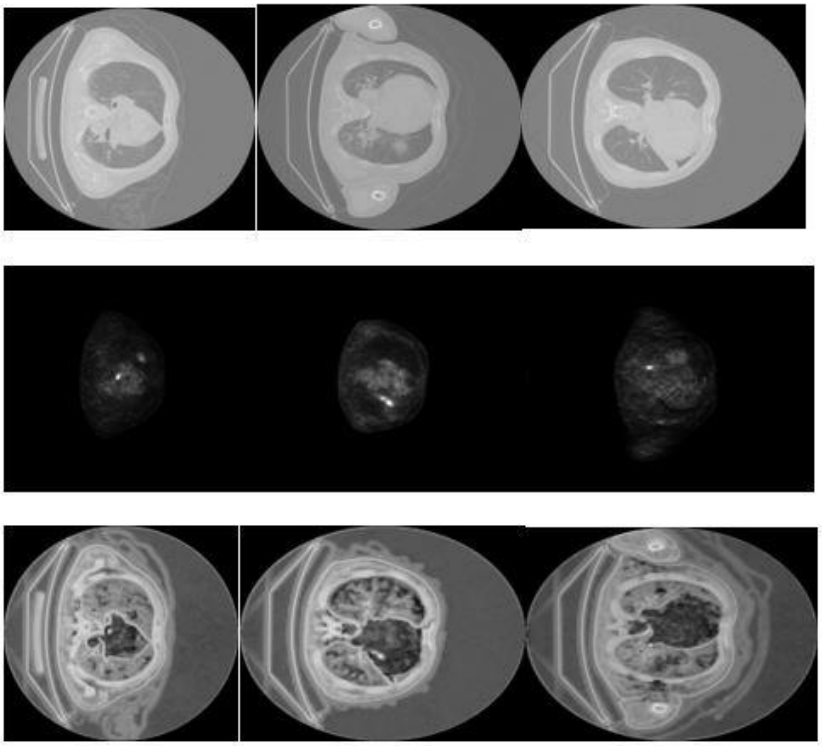

Figure 7. Correlating CT and PET slices out of disparate kinds of lung cancer. (i) NSCLC (ii) LUADC (iii) SCLC

in which f_w portrays the filter breadth, and f_h portrays the filter length. This study sets the filter's dimension to the value of three, and the expression turns f_s =3×3. Such filters assist to acquire images' less-range features like edges and curves and are named feature detectors. The additional 3 CLs are included in the paradigm for excerpting the deep features out of the RGs and, hence, giving images' overall patterns to the paradigm. The filters as well begin the convolutional procedures with the RGs' sub-area (SA). The cause for implementing the convolutional procedure is to multiply and add the RGs' filters and pel values.

The datasets comprise data specimens of 397 sick persons overall, 91 sick persons having SCLC, 103 sick persons having LUSC, and 203 sick persons having LUAD. Instance lesion slices for 3 disparate kinds of lung cancer out of CT and PET analyses are illustrated in Figure 7. The figure is the fused images of the threshold limit fused images and it is the fused images of the optimal Generative network

The RGs' SA can be as well known as the receptive field (RF). Next, additional CLs have been included in the paradigm stage by stage for enhancing the paradigm's capability for excerpting the feature components. The filter values have been named weights, and the weights should be learnt by the paradigm during the training. The filter performs the convolution procedure from RGs' beginning and continuously shifts over the whole RG. The methodology of executing filter convolution stops while the entire image has been processed. The value of shifting the filter has been limited by another criterion called stride. This work fixes the stride's value (SV) to one for entire CLs. By choosing the stride's greater values, it lessens the spatial dimension (SD) and generates few problems; for example, the RF could move outward of the IP dimension. Thus, for subduing such issues, a zero-padding (ZP) approach has been implemented that pads the surrounding IP RGs' border for sustaining the OP dimension as similar to the IP. The SV has been acquired by employing the ZP approach,

$$P = \frac{h-1}{2} \tag{15}$$

in which P portrays the ZPs, and h comprises the filter length or breadth since the filter's length as well as breadth remain similar in this work. CLs employ disparate filters' quantity to excerpt several predominant

features. This paradigm's initial CL comprises sixteen filters. The filters quantity of the rest of the CLs has been modified from 16 to 512. The OP volume (as well as called the activation map) has been computed accordingly.

$$V_w = \frac{I_w - f_w + 2P}{D} + 1$$
$$V_h = \frac{I_h - f_h + 2P}{D} + 1 \quad (16)$$
$$O_i = F_n$$

in which I_w and I_h denote the length and breadth of the II, and f_w and f_w denote the filter dimension concerning length and breadth. P denotes the ZP, and D remains the stride, whereas F_n comprises the filter's quantity implemented to the CLs. The initial CL of this proffered paradigm comprises I_w =300, I_h=300, f_w=3, f_h=3, D=1, P=0, and F_n=16.

The activation function (AFn) rectified linear unit (RLU) has been implemented upon the convolutional Ops. The RLU activation procedure remains to modify the entire negative outcomes by 0. The intention to employ the RLU function in this proffered CNN paradigm remains that this enhances the nonlinearity and aids to enhance the computative duration devoid of impacting the paradigm's accuracy. The max-pooling layer (MPL) has been included subsequent to the CLs that assists in lessening the IP RGs' spatial size. The filter dimension is set as 2×2 and as well the stride value is fixed to two for MPLs. The filter convolves throughout the IP volume and generates the greatest value of the RGs' SA. It is noticed that MPLs comprises a specific feature's location regarding the rest of the features. This as well lessens the computative price by reducing the weight's quantity and evades the paradigm by OFt.

Next, the dropout (DP) layer has been included. DP's disparate values are implemented to this paradigm like 0.05, 0.20, and 0.25, yet just the suitable and important threshold of 0.20 has been stated in the present work. The cause of implementing disparate DP values is to evade this proffered paradigm from OFt. This layer's functioning is to drop out the activation haphazardly and as well validate the paradigm for anticipating the real label. This paradigm's flattened layer has been implemented to transform the 2D FM into a 1D feature vector (FV). The data produced by the FL have been supplied to the FCL. The FCL executes the classification procedure out of the acquired 1D FV. For the present work, FCL comprises 512 neurons. The FCL's first dense layer (DsL) forwards the activation outcome to the rest of the DsLs. The paradigm's final OP has evolved out of a DsL having 4 neurons and a softmax AFn that classifies the OP image.

10. PERFORMANCE ANALYSIS

For the intention of assessment, accuracy, precision, recall, and F1-score have been chosen as the criteria. The proffered AWBABCA enhanced VGG19+CNN can be called OVGG19+CNN. It has been correlated with 3 conventional methodologies – DNNASCCSA and enhanced chaotic crow search and particle swarm optimization algorithm (ECCSAPSOA) centered upon these criteria.

Accuracy provides the capability of the comprehensive anticipation generated by the paradigm. True positive (TP) and true negative (TN) give the ability to anticipate cancer's existence and non-existence. False positive (FP) and false negative (FN) give false anticipations done by the employed paradigm.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (17)$$

Table 1. Accuracy correlation

No. of epochs	DNNASCCSA	ECCSAPSOA	OVGG19+CNN
10	85.1	90.1	97.9
20	86	89.5	97.6
40	85.3	89.7	98
60	85.2	89.5	97.4
100	86	89.2	97

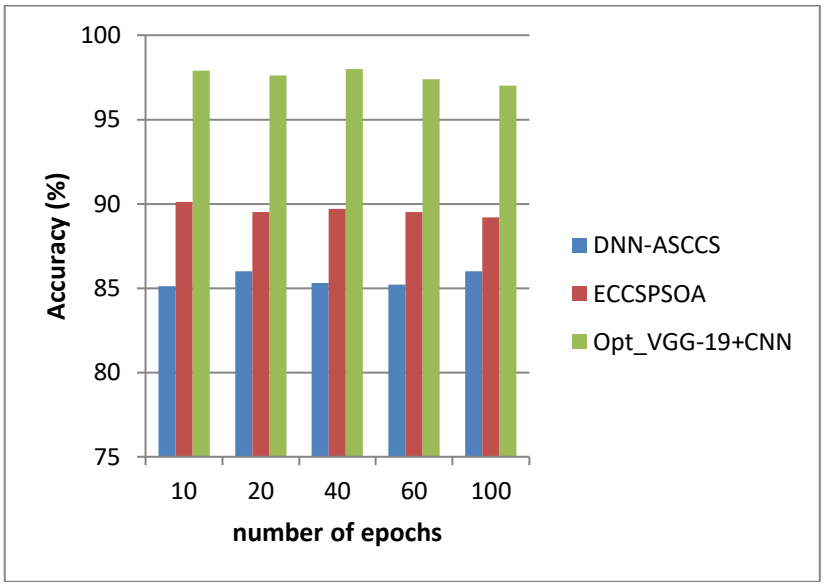

Figure 8. Accuracy correlation

Figure 8 exhibits the accuracy correlation betwixt the prevailing DNNASCCSA, ECCSAPSOA, and the proffered OVGG19+CNN methodology, wherein the X-axis portrays the epochs' quantity employed for the assessment, and Y-axis portrays the accuracy values acquired in percentage. While correlated, DNNASCCSA and ECCSAPSOA methodologies attained an accuracy of 86.3% and 90.2%. The proffered OVGG19+CNN methodology attained 98.34% accuracy which remains 12.04% finer than DNNASCCSA and 8.14% finer than ECCSAPSOA.

Precision remains the proportion of the positive sample number. Instead, it portrays the ratio of the anticipation paradigms if cancer really exists.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{18}$$

Table 2. Precision correlation

No. of epochs	DNNASCCSA	ECCSAPSOA	OVGG19+CNN
10	78.1	84.1	97
20	77.5	83.9	96.9
40	77.6	83.7	96.8
60	77.4	84	96.1
100	77.9	83.8	97.5

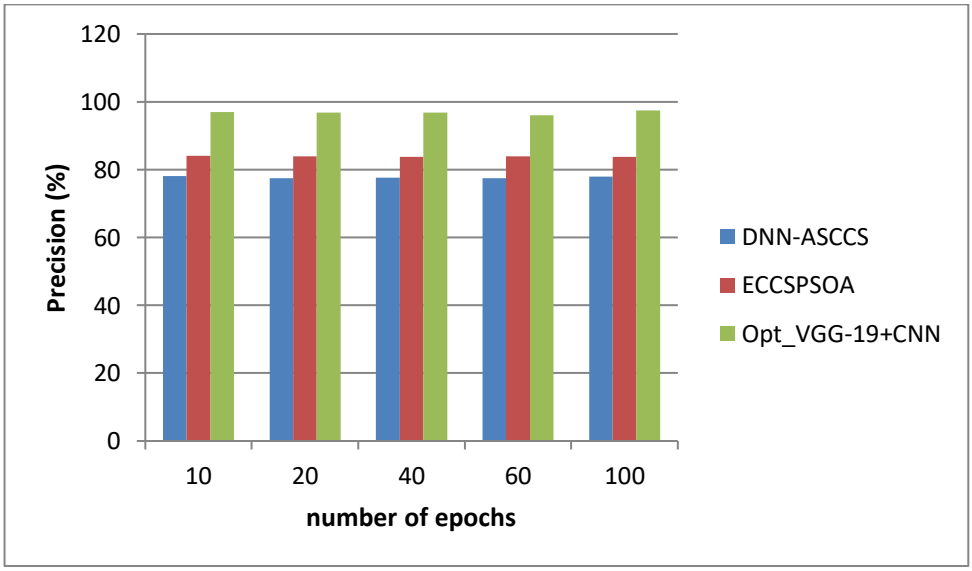

Figure 9. Precision correlation

Figure 9 exhibits the precision correlation betwixt the prevailing DNNASCCSA, ECCSAPSOA, and the proffered OVGG19+CNN methodology, wherein the X-axis portrays the epochs' quantity employed for the assessment, and Y-axis portrays the precision values acquired in percentage. While correlated, DNNASCCSA and ECCSAPSOA methodologies attained a precision of 78.3% and 84.5%. The proffered OVGG19+CNN methodology attained 97.96% precision which remains 19.63% finer than DNNASCCSA and 13.41% finer than ECCSAPSOA.

The datasets comprise data specimens of 397 sick persons overall, 91 sick persons having SCLC, 103 sick persons having LUSC, and 203 sick persons having LUAD. Instance lesion slices for 3 disparate kinds of lung cancer out of CT and PET analyses are illustrated in Figure 9.

The above Figure 9 is the fused images of the threshold limit fused images and it is the fused images of the optimal Generative network

Recall indicates the identification capability to precisely identify cancer in the DS; the sensitivity computation in no way considers undetermined test outcomes as a test could not be reiterated, and undetermined samples must entirely be excepted out of the assessment.

$$\text{recall} = \frac{TP}{TP+FN} \tag{19}$$

Table 3. Recall correlation

No. of epochs	DNNASCCSA	ECCSAPSOA	OVGG19+CNN
10	89.9	96	98.4
20	89.6	95.7	98.4
40	89.7	95.3	98.4
60	90	96.1	98.7
100	90.5	96	98.3

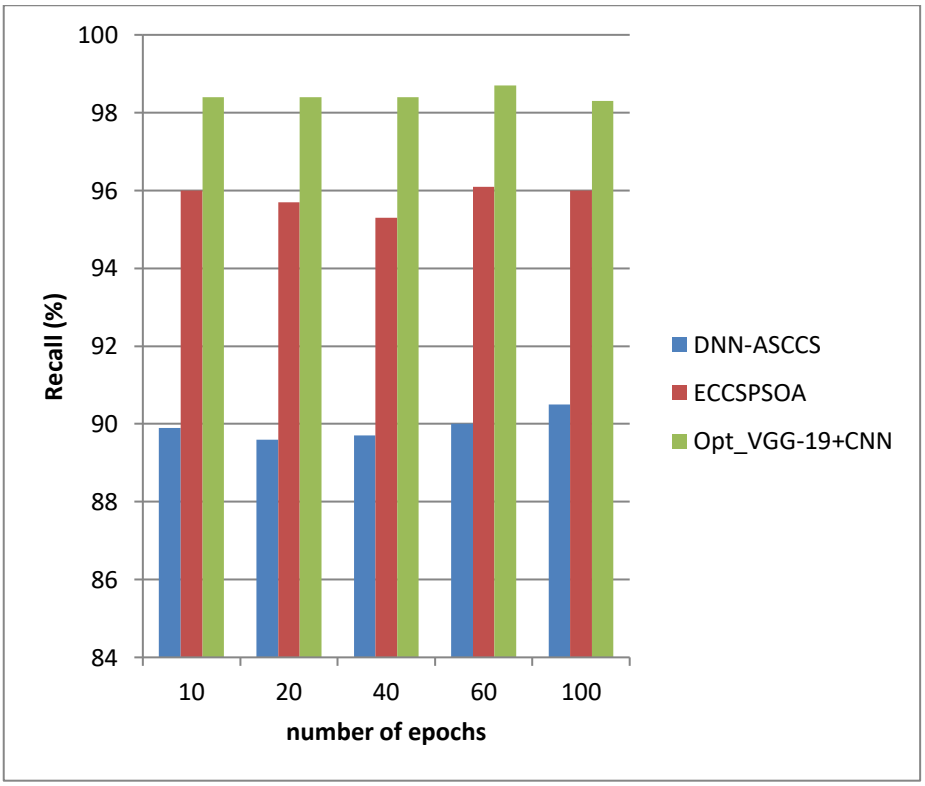

Figure 10. Recall correlation

Figure 10 exhibits the recall correlation betwixt the prevailing DNNASCCSA, ECCSAPSOA, and the proffered OVGG19+CNN methodology, wherein the X-axis portrays the epochs' quantity employed for the assessment, and Y-axis portrays the recall values acquired in percentage. While correlated, DNNASCCSA and

ECCSAPSOA methodologies attained a recall of 90% and 92.6%. The proffered OVGG19+CNN methodology attained 98.72% recall which remains 8.72% finer than DNNASCCSA and 4.12% finer than ECCSAPSOA.

F1-score can be employed for determining the anticipation execution. This remains the weighted mean of precision and recall. The value of one discerns the finest whereas zero remains the poorest. F1-score in no way regards TNs and is computed as,

$$f1 - score = \frac{2 * P * R}{P + R} \tag{20}$$

Table 4. F1-score correlation

No. of epochs	DNNASCCSA	ECCSAPSOA	OVGG19+CNN
10	78.5	84	98.4
20	79	83.5	98.4
40	78.4	83.9	98.4
60	78.7	83.6	98.7
100	78	83.1	98.3

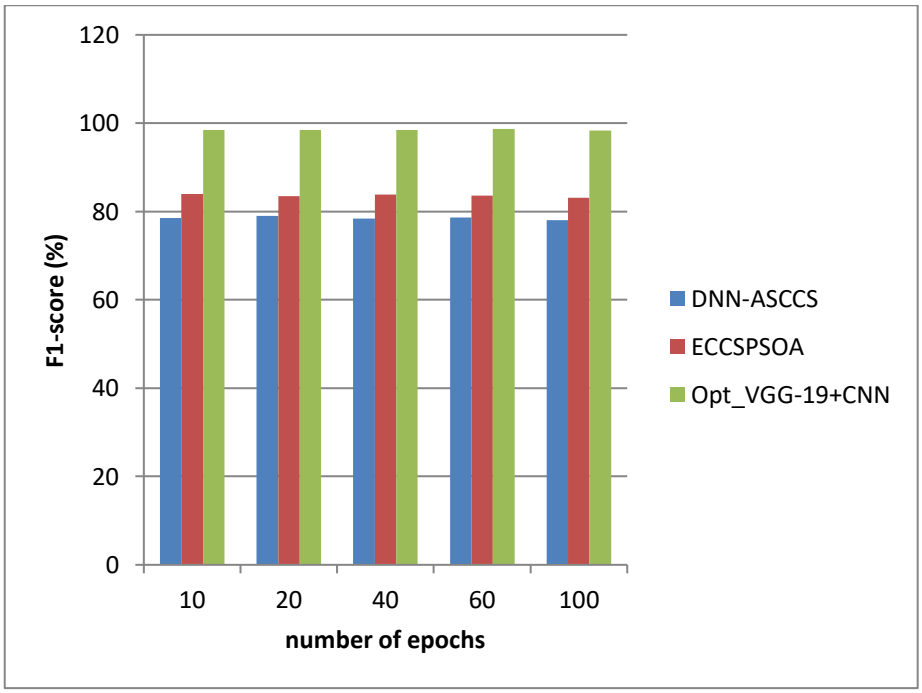

Figure 11. F1-score correlation

Figure 11 exhibits the F1-score correlation betwixt the prevailing DNNASCCSA, ECCSAPSOA, and the proffered OVGG19+CNN methodology, wherein the X-axis portrays the epochs' quantity employed for the assessment, and Y-axis portrays the F1-score values acquired in percentage. While correlated, DNNASCCSA and ECCSAPSOA methodologies attained an F1-score of 79.1% and 84.4%. The proffered OVGG19+CNN methodology attained 98.32% F1-score which remains 8.72% finer than DNNASCCSA and 4.12% finer than ECCSAPSOA.

Table 5. computation time correlation

No. of epochs	DNNASCCSA	ECCSAPSOA	OVGG19+CNN
10	57	71	34.3
20	57.4	72.5	33
40	54	73	33.3
60	58.3	71.9	35
100	58	72	34

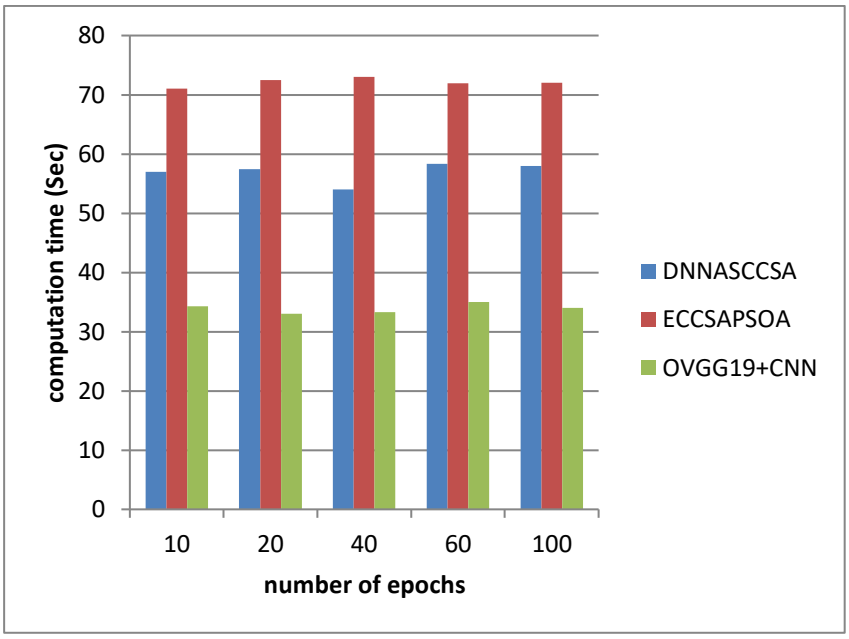

Figure 12. computation time

Figure 12 exhibits the computation time correlation betwixt the prevailing DNNASCCSA, ECCSAPSOA, and the proffered OVGG19+CNN methodology, wherein the X-axis portrays the epochs' quantity employed for the assessment, and Y-axis portrays the computation time values acquired in seconds. While correlated, DNNASCCSA and ECCSAPSOA methodologies attained an computation time of 57.4sec and 72sec. The proffered OVGG19+CNN methodology attained 33.4sec computation time which remains 23sec lesser than DNNASCCSA and 39sec lesser than ECCSAPSOA.

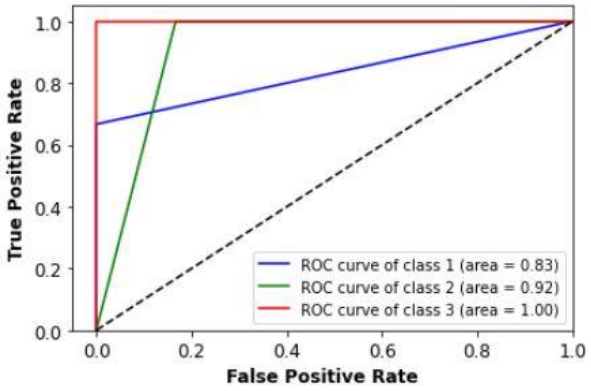

Figure 13. ROC curve for the testing process

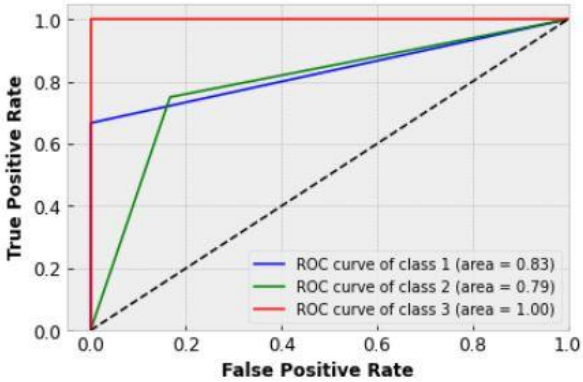

Figure 14. ROC curve for the training process

Figure 13 shows the ROC curve of the testing process and Figure 14 describes the ROC curve of the training process. Table 5 exhibits the comprehensive correlative assessment between the prevailing methodologies like DNNASCCSA and ECCSAPSOA with the proffered OVGG19+CNN methodology.

Table 5. Comprehensive correlative assessment

Methodologies	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Computation time (sec)
DNNASCCSA	86.3	78.3	90	79.1	57.4
ECCSAPSOA	90.2	84.5	92.6	84.4	72
OVGG19+CNN	98.34	97.96	98.72	98.32	33.4

Computational Complexity:
The computational complexity of the proposed algorithm is O(mn2). Where n is the population size and m is the no. of selected features. Compared to the other architectures they take less execution time. The time complexity will be reduced.

11. CONCLUSION

DL's highest advantage over the rest of the AGs on ML remains its capability in performing feature engineering by itself. It analyzes the data to search for related features and integrates these to give for swift learning. This makes use of spatial coherence in the IP. The images' training and testing have been performed in which images have been pre-processed, and images' FS and FE have been performed. When the training and testing portion has been performed victoriously, the VGG19+CNN Algorithm classifies the IP lung image as typical or atypical, and the OP would be exhibited. The proffered CC methodology remains low competitively intensive and provides greater accuracy. Therefore, this identifies cancer rightly that results in sick persons' greater survivability. By identifying cancer in the initial stage, the cancer operations' hit ratio will be raised. Nevertheless, the proffered system's execution exhibits slight enhancement to the prevailing methodologies. In the prospective study, the proffered system's execution would be additionally enhanced by implementing hybrid FS methodologies.

REFERENCES

[1] R. Navid, M. Ashourian, M. Karimifard et al., "Computer-aided diagnosis of skin cancer: a review," *Current Medical Imaging*, vol. 16, no. 7, pp. 781–793, 2020.
[2] S. C. Satapathy, N. Sri Madhava Raja, V. Rajinikanth, A. S. Ashour, and N. Dey, "Multi-level image thresholding using Otsu and chaotic bat algorithm," *Neural Computing and Applications*, vol. 29, no. 12, pp. 1285–1307, 2018.
[3] Q. Liu, Z. Liu, S. Yong, K. Jia, and N. Razmjooy, "Computer-aided breast cancer diagnosis based on image segmentation and interval analysis," *Automatika*, vol. 61, no. 3, pp. 496–506, 2020.
[4] A. Hu and R. Navid, "Brain tumor diagnosis based on metaheuristics and deep learning," *International Journal of Imaging Systems and Technology*, vol. 31, no. 2, pp. 657–669, 2020.
[5] R. Wender, E. T. H. Fontham, E. Barrera et al., "American Cancer Society lung cancer screening guidelines," *CA: A Cancer Journal for Clinicians*, vol. 63, no. 2, pp. 106–117, 2013.
[6] N. Ghadimi, "An adaptive neuro-fuzzy inference system for islanding detection in wind turbine as distributed generation," *Complexity*, vol. 21, no. 1, pp. 10–20, 2015.
[7] N. Razmjooy, V. V. Estrela, and H. J. Loschi, "Entropy-based breast cancer detection in digital mammograms using world cup optimization algorithm," *International Journal of Swarm Intelligence Research*, vol. 11, no. 3, pp. 1–18, 2020.
[8] N. Dey, V. Rajinikanth, A. Ashour, and J. M. Tavares, "Social group optimization supported segmentation and evaluation of skin melanoma images," *Symmetry*, vol. 10, no. 2, p. 51, 2018.
[9] V. Rajinikanth and S. C. Satapathy, "Segmentation of ischemic stroke lesion in brain MRI based on social group optimization and Fuzzy-Tsallis entropy," *Arabian Journal for Science and Engineering*, vol. 43, no. 8, pp. 4365–4378, 2018.
[10] Xing Liu, Lin Shang "A Fast Wrapper Feature Subset Selection Method based On Binary Particle Swarm Optimization" 2013 IEEE Congress on Evolutionary Computation.
[11] D. Asir Antony Gnana Singh, S. Appavu alias Balamurugan, E. Jebamalar Leavline "Literature Review on Feature Selection Methods for High-Dimensional Data", International Journal of Computer Applications (0975 – 8887) February 2016
[12] Suresh Dara, Haider Banka" A Binary PSO Feature Selection Algorithm for Gene Expression Data, International Conference on Advances in Communication and Computing Technologies,2014
[13] Adamu, A., Abdullahi, M., Junaidu, S. B., & Hassan, I. H. (2021). A hybrid particle swarm optimization with a crow search algorithm for feature selection. *Machine Learning with Applications*, *6*, 100108.
[14] Dey, C., Bose, R., Ghosh, K. K., Malakar, S., & Sarkar, R. (2022). LAGOA: Learning automata-based grasshopper optimization algorithm for feature selection in disease datasets. *Journal of Ambient Intelligence and Humanized Computing*, *13*(6), 3175-3194.

[15] Toğaçar, M. (2021). Disease type detection in lung and colon cancer images using the complement approach of inefficient sets. *Computers in Biology and Medicine*, *137*, 104827.
[16] Houssein, Essam H., Eman Saber, Abdelmgeid A. Ali, and Yaser M. Wazery. "Centroid mutation-based Search and Rescue optimization algorithm for feature selection and classification." *Expert Systems with Applications* 191 (2022): 116235.
[17] Piri, J., & Mohapatra, P. (2021). An analytical study of modified multi-objective Harris Hawk Optimizer towards medical data feature selection. *Computers in Biology and Medicine*, *135*, 104558.
[18] Shen, C., & Zhang, K. (2022). Two-stage improved Grey Wolf optimization algorithm for feature selection on high-dimensional classification. *Complex & Intelligent Systems*, *8*(4), 2769-2789.
[19] Maleki, N., Zeinali, Y., & Niaki, S. T. A. (2021). A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection. *Expert Systems with Applications*, *164*, 113981.
[20] Surendar, P. (2021). Diagnosis of lung cancer using a hybrid deep neural network with adaptive sine cosine crow search algorithm. *Journal of Computational Science*, *53*, 101374.
[21] Bansal, M., Kumar, M., Sachdeva, M., & Mittal, A. (2021). Transfer learning for image classification using VGG19: Caltech-101 image data set. *Journal of ambient intelligence and humanized computing*, 1-12.