Journal of Science and Technology
ISSN: 2456-5660 Volume 6, Issue 06 (Nov- Dec 2021)
www.jst.org.in *DOI:https://doi.org/10.46243/jst.2021.v6.i06.pp269- 281*

Leveraging Cloud Technology for Convolutional Neural Networks in Cancer Detection and Diagnosis

[1]**Vijai Anand Ramar**
Delta Dental Insurance Company, Georgia, USA
vijaianandramar@gmail.com

[2]**R. Hemnath**
Nandha Arts and Science College, Erode, India.
hemnathrmca@gmail.com

To Cite this Article

Vijai Anand Ramar, R. Hemnath, "**Leveraging Cloud Technology for Convolutional Neural Networks in Cancer Detection and Diagnosis**" *Journal of Science and Technology, Vol. 6, Issue 06-Dec 2021, pp269-281*

Article Info

Received: 30-09-2021 Revised: 06-12-2021 Accepted: 17-12-2021 Published:27 -12-2021

Abstract

Lung cancer is one of the areas in which detection still poses some challenges for the healthcare arena. The classical way is to analyse histopathological images by human intervention; this is a time-consuming process that is somewhat on the error-prone side. A cloud-based convolutional neural network model has been proposed for the diagnosis of lung cancer, specifically targeting lung adenocarcinoma via histopathological images from TCGA_LUAD datasets. For cleaner input data, sophisticated preprocessing methods are applied like denoising and augmentation. The Firefly Optimization Algorithm has been implemented for feature selection that consumes the least computation. Cloud computing helps with scaling, real-time predictions, and access for healthcare professionals, thus eliminating some constraints on local hardware resources. Notably, cancer detection with the proposed CNN model brings efficiency to the process, increases diagnostic accuracies, and reduces the workload of pathologists for treatment decisions. The system is designed to scale and grow with a security infrastructure, so it will adapt and incorporate future data and advances in cancer research to obtain more timely and accurate clinical diagnostics.

Keywords

Cloud-Based Model, Convolutional Neural Networks, Lung Cancer Detection, Histopathological Images, TCGA_LUAD Dataset, Data Preprocessing, Denoising Techniques, Data Augmentation

1. INTRODUCTION

One of the most dangerous diseases these days is cancer. Early detection of cancers has been associated with better survival outcomes[1].The classical methods of lung cancer detection have depended on manual histopathological image interpretation, being a lengthy process and having easy access to errors[2]. The complexities and volume of imaging data increase the importance of quick, reliable automated diagnostic systems[3]. Convolutional neural networks show a scepticism in this respect, as they specialize in the automation of classification tasks, as applied to medical images, and provide very significant improvements in terms of accuracy and speed[4].This research thus focuses on developing a CNN model for the cloud for cancer detection, with lung cancer (adenocarcinoma) involved, based on histopathological images from the TCGA_LUAD dataset[5]. This endeavor aims to provide an enhanced workflow in diagnostics, delivering quick and accurate results while combined with cloud computing[6]. This model uses advanced preprocessing techniques, comprising denoising and augmentation, to enhance input quality and increase model generalization[7]. Further, the Firefly Optimization Algorithm is utilized to maximize accuracy while minimizing computational cost by selecting the most important image features[8]. This Intuition aims to provide a reliable, efficient, and scalable tool for the automated diagnosis of cancer, improving upon clinical decision-making and patient outcome[9]. The designed CNN model based on a cloud system for cancer detection also addresses scalability and accessibility challenges in healthcare systems[10]. It is capable of handling big data sets and ensuring proper processing and storage without local hardware limitations by leveraging cloud computing[11]. Therefore, this is an important feature in clinical environments where real-time diagnostics are required[12]. The model is designed to be directly accessed simple to health professionals through a user-friendly interface prompting upload of histopathological images and obtaining results immediately[13]. Continuous updating and improvement are possible for such a model through integration with cloud storage since this system can evolve with future data and advances in cancer research[14]. The model, therefore, should not only cater to the time frames of supplying timely, automated, and accurate diagnostic support for improved cancer detection but also lessen workload requirements on pathologists thereby facilitating quicker treatment decisions, which improve the outcome for patients[15].

1.1 PROBLEM STATEMENT

Accelerated identification and diagnosis of cancer, particularly lung cancer, pose a significant challenge in the medical field, with traditional approaches being time-consuming and always prone to interpreter error[16]. Histopathological image analysis wherein pathologists discriminate between cancerous and non-cancerous cells is an important practice for tissue diagnosis. However, given the huge volume and complexity of medical imaging data and the need for accurate and quick analysis, traditional means of diagnosis seem to falter under the pressure[17]. A delay in the diagnosis might lead to negative repercussions on the patients themselves, particularly for aggressive cancers like lung adenocarcinoma[18]. Since timely interventions are crucial, there is a pressing need for automated systems that provide accurate and timely diagnostic support for the pathologist[19]. Against this backdrop, this research seeks a cloud-based convolutional neural network (CNN) model for detecting and diagnosing cancer (lung cancer in particular) with histopathological images[20]. Using TCGA_LUAD dataset and via deep learning techniques, this model will be capable of automatically helping classify images into normal, benign, and malignant categories[21]. The cloud computing integration would provide the huge prevalence of high-performance processing solutions for large datasets at quick-turnaround to clinical environments where timely analysis is critical[22]. The cloud-based model will facilitate real-time prediction, thus making it possible for the professionals to obtain reliable results without being constrained by the availability of hardware resources, thereby enhancing overall accuracy in diagnosis while speeding up the actual decision-making process [23].

1.1.2 Objective

- In essence, the design and development of convolutional models in this study were on the realization of an automated detection classification of lung cancer histopathological images using TCGA_LUAD dataset, together with several other lesser tasks, actually putting the very practical cancer detection algorithm into existence.
- Data Preprocessing: Considerable quality improvement is done on the input data using denoising methods like Gaussian filtering supported with various data augmentation methods including rotation, flipping, and zooming, and normalization in the context of CNN for maximum efficiency.
- Improvement in Feature Selection Using the Firefly Algorithm: The Firefly optimization algorithm aims to filter out only the most relevant features for the given high-dimensional image data, so as to minimize computational costs and maximize model accuracy.
- User-friendly interface for an on-demand lung cancer diagnosis prediction model. Open up the cloud to healthcare professionals for instant predictions based on histopathological images, thus improving the efficiency of clinical workflows.
- The architecture is designed to allow scalability for being integrated in the future with larger datasets and compliance with various healthcare regulatory acts considering secure and reliable cloud-based diagnostics.

2. LITERATURE SURVEY

In recent decades, the evolution and rapid deployment of deep learning methodologies have significantly transformed the landscape of medical diagnostics, particularly in the realm of histopathological image analysis for cancer detection [24]. Histopathology, which involves the microscopic examination of tissue samples, is a gold standard in cancer diagnosis. However, traditional techniques reliant on human expertise often suffer from inter-observer variability, subjectivity, fatigue-induced errors, and a requirement for high-level domain experience [25]. To circumvent these limitations, Convolutional Neural Networks (CNNs) have emerged as a powerful computational approach capable of automating complex image classification tasks with remarkable accuracy and consistency [26]. CNNs function by learning intricate hierarchical representations from raw image data, progressively abstracting low-level features such as edges and textures to high-level semantic patterns that differentiate normal from malignant tissues [27]. These capabilities have been extensively validated in literature, where CNNs demonstrate the ability to outperform classical machine learning methods and even match the performance of experienced pathologists in certain diagnostic tasks [28]. This transition towards AI-driven diagnostics is particularly compelling in the case of lung cancer—an aggressive and often late-diagnosed malignancy characterized by highly complex tissue morphologies. The accurate interpretation of histopathological slides in lung cancer requires not only a high level of precision but also the ability to handle heterogeneous patterns that vary across patients and stages. Consequently, reanalyzing histopathological images through deep learning frameworks like CNNs has become an area of prime research interest [29]. Moreover, the integration of cloud computing platforms has revolutionized how medical image analysis systems are deployed and scaled. Cloud-based infrastructures facilitate the processing of large-scale histopathological datasets by providing virtually unlimited computational resources, seamless scalability, and ubiquitous accessibility [30]. These platforms allow pre-trained CNN models to be deployed and accessed in real-time from anywhere in the world, thereby reducing the dependency on local hardware setups and enabling healthcare professionals in resource-constrained settings to benefit from advanced diagnostic systems [31], Additionally, cloud solutions enhance collaborative research by offering centralized repositories for data storage, model sharing, and joint validation. To further improve the performance and computational efficiency of CNN-based models, recent advancements have introduced the integration of data augmentation techniques and feature selection strategies using metaheuristic optimization algorithms [32]. Among these, the Firefly Optimization Algorithm (FOA) has shown promising results. Inspired by the luminescent signaling behavior of fireflies, FOA operates on the principle of attraction, where brighter fireflies attract others, thereby guiding the search process towards optimal solutions [33], This bio-inspired mechanism proves especially useful in selecting the most informative features from high-dimensional image data, which not only reduces the computational load but also enhances the classification performance of the model [34]. By employing FOA for feature selection, CNN models become more robust and interpretable. The algorithm strategically filters out redundant or non-informative features, ensuring that only the most diagnostically relevant patterns are retained [35], This leads to faster convergence during training, improved model generalizability, and enhanced diagnostic precision [36]. When coupled with CNNs, FOA serves as a synergistic approach that enables the development of lightweight, yet highly accurate, diagnostic tools suited for real-time analysis of extensive medical image datasets [37], The growing trend towards cloud integration further complements this framework by offering a scalable and cost-effective solution for storing, managing, and analyzing massive volumes of medical data Cloud services, such as those provided by AWS, Azure, or Google Cloud [38], not only support the deployment of AI models but also enable continuous model updates, real-time feedback loops, and integration with electronic health record (EHR) systems [39], This convergence of cloud computing and AI significantly accelerates the adoption of intelligent diagnostic systems across a wide spectrum of healthcare institutions, ranging from advanced research hospitals to rural clinics [40].

Building upon these technological advancements, there is an increasing movement toward designing end-to-end automated diagnostic pipelines that integrate deep learning models with feature selection algorithms and cloud infrastructure to deliver intelligent, responsive, and scalable cancer detection systems [41]. In such frameworks, the diagnostic journey begins with the acquisition of histopathological images, which are then preprocessed and normalized to ensure uniformity across samples [42]. These images undergo augmentation techniques such as rotation, flipping, zooming, and color jittering to artificially increase dataset diversity and prevent overfitting during training [43]. Subsequently, CNNs are employed to extract high-dimensional feature maps from the input images, capturing both spatial and textural nuances critical for distinguishing cancerous from non-cancerous tissues [44]. To optimize the learning process and eliminate redundancy, the Firefly Optimization Algorithm (FOA) is deployed to select the most salient features that contribute directly to classification accuracy [45]. This step not only reduces computational complexity but also enhances the model's ability to generalize across diverse patient populations [46]. The selected features are then passed through fully connected layers or decision layers that generate the final prediction outcomes—categorizing samples into malignant or benign classes [47]. Once trained, the entire model is hosted on a cloud platform where it can be accessed by clinicians and healthcare practitioners globally [48]. Real-time image uploads, model inference, and feedback reporting are managed seamlessly through cloud APIs, allowing for rapid diagnostic support in clinical settings [49]. Moreover, cloud platforms enable model retraining and version control, ensuring that the system evolves with new data and maintains high diagnostic fidelity [50]. This integrative paradigm embodies the essence of modern AI-powered healthcare: a confluence of data-driven intelligence, evolutionary computation, and ubiquitous accessibility [51]. As the field progresses, future directions may include the integration of explainable AI (XAI) methods for interpretability, federated learning for privacy-preserving model training, and multimodal learning for combining histopathological data with genomic, clinical, or radiological information to deliver more holistic and personalized cancer diagnostics.

3. PROPOSED METHDOLOGY

The proposed CNN model that is cloud-based for the detection of lung cancer is shown in figure 1. This input starts from data acquisition from the database TCGA_LUAD and comes through a preprocessing step that includes denoising and augmenting the input data to make it more qualitative and varied. After preprocessing, classification of lung cancer with the CNN model is performed. The next step is deploying it on the cloud for scalable processing and real-time predictions. Also, with the selected features, the Firefly Optimization Algorithm is executed, which would help to optimize the performance of the model by selecting the relevant features while keeping the computational cost down. This interoperability framework shows how data preprocessing, cloud coupling, CNN-based classification, and optimization techniques work in tandem to give automation and improvement in cancer detection.

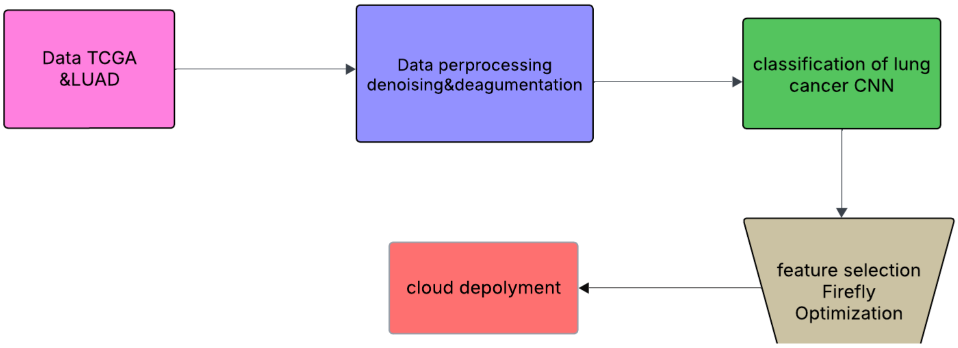

Figure1: *Proposed Architecture*

3.1 Data collection

The present research employed the TCGA_LUAD resource (Lung Adenocarcinoma) data available on The Cancer Genome Atlas (TCGA) via hosting by the Genomic Data Commons (GDC). This dataset is based on Whole Slide Images (WSIs) of lung biopsies stained with Hematoxylin and Eosin (H&E) and scanned at different magnifications, ranging from 5x to 40x. These high-resolution digital slides areannotated by certified pathologists, whose relevant metadata includes tumor classification (normal, benign, malignant), tumor stage (I–IV), mutation types, patient age, and survival outcomes. GDC Data Portal and GDC API give access to the data, both of which start a journey toward secure and structured downloading. In the present study, WSIs were downloaded in .svs format and screened for quality, annotation completeness, and relevance to lung adenocarcinoma. Each image shares a connection with structured clinical metadata, making it adequate in supervised learning. This collected data is then moved to a secure cloud storage environment for large-scale parallel preprocessing, training, and deployment of CNN models while maintaining compliance with data governance protocols.

https://www.kaggle.com/datasets/nahin333/tcgaluad

3.2 Data Preprocessing

The preprocessing phase becomes essential for the classification of histopathological images coming from the TCGA_LUAD dataset with relatively fast and accurate convolutional neural networks. The raw Whole Slide Images (WSIs) are often contaminated by visual noise from artifacts of the scanners, uneven staining, or background irregularities. To remove these undesired effects, denoising methods such as Gaussian blurring and median filtering are used to smooth out the irrelevant noise while retaining local structural features that are imperative for making diagnostic inferences. Once the cleaning process is done, data augmentation is further performed to enhance the robustness of models and prevent overfitting. Intensification techniques include random rotation, flips, zooming, cropping, contrast enhancement, and brightness variations to imitate real-world variations in slide preparation and tissue orientation. Post-augmentation, the images are resized to a standard dimension (e.g., 224×224 pixels) and normalized to allow the pixel intensity ranges to remain consistent for all images. The proposed preprocessing pipeline simultaneously boosts feature quality, increases data diversity, and secures that good representative and high-fidelity input data train the CNN model.

3.2.1 Data Denoising

Histopathological imaging analysis becomes a vital prior step in any data denoising process for possible artifact removal or removal of pixel-level variations that might obscure important features in diagnosis. The whole slide images of the TCGA_LUAD dataset generally contain noises stemming from inconsistencies in tissue staining, scanner artifacts, or compression losses. Importance of histopathological image analysis in preprocessing the data denoised for possible removal of artifacts and pixel-level variations that overshadow important diagnostic features. Presence of noise in the whole slide images of the dataset TCGA_LUAD due to the tissue the method applies Gaussian filtering to smooth the images by averaging the pixel value of the pixel in question with those of its neighbours, using a weighted kernel. This enables the reduction of high-frequency noise while retaining cellular regions' edges and structures essential in cancer classification. The Gaussian filter is defined by the equation:

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

where $G(x,y)$ is the Gaussian function, and σ is the standard deviation controlling the degree of smoothing. Applying this kernel across the image reduces random variations and enhances the clarity of morphological patterns for CNN-based feature extraction.

3.2.2Data augmentation

Data augmentation remains an instrumental process to enlarge training data artificially while also enhancing data variety, significantly in medical image analysis, where annotated data are in short supply. It aids TCGA_LUAD histopathological images in achieving a greater generalization of the models during augmentation by simulating potential variations in tissue orientation and position and overall appearance that can occur in real clinical settings. Common transformations are random rotations and flips, along with zooming, brightness adjustments, and cropping. These operations result in the generation of new training examples that keep the essence of pathology, further enabling the CNN to learn robust features. Rotation is a common augmentation process mathematically described by a 2D rotation matrix applied to each pixel coordinate

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{2}$$

where (x,y) are the original pixel coordinates, (x',y') are the rotated coordinates, and θ is the rotation angle. This controlled augmentation ensures the CNN becomes invariant to spatial orientations and better detects cancerous patterns across varied image contexts.

3.3 Classification of lung Cancer using CNN

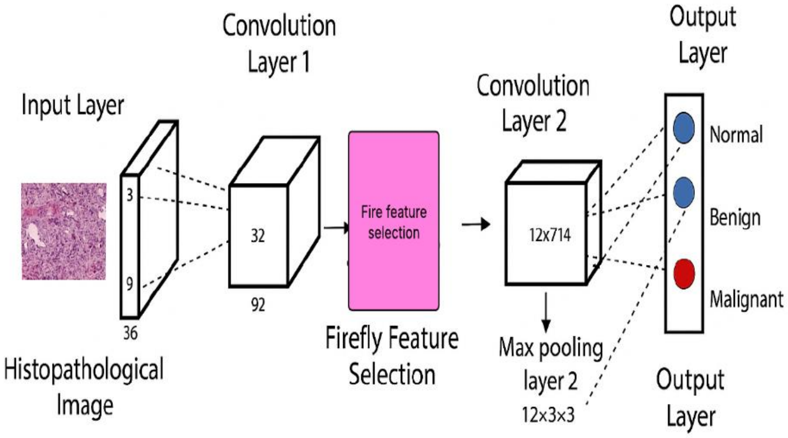

Figure2: *CNN-Based Lung Cancer Detection with Firefly Feature Selection*

Convolution is the primary operation in CNNs that extracts spatial characteristics from input images by having learnable filters slide over the input matrix is shown in figure 2. Each filter captures local patterns like edges and textures, while deeper-layer filters capture more abstract feature representations. It reduces the images' spatial dimensions but maintains the crucial information necessary for classification.

3.3.1Convolution of operation

The most important operation in Convolutional Neural Networks (CNNs) is convolution, which uses learnable filters to slide over input matrices to create spatial features from input images. An individual filter extracts patterns from the original input, such as edges and textures, and later might capture more abstract patterns in the deeper layers. This reduces the spatial dimension, but retains sufficient information for classification decisions.

$$h_{i,j}^{(k)} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x_{i+m,j+n} \cdot w_{m,n}^{(k)} + b^{(k)} \tag{3}$$

Where $h_{i,j}^{(k)}$ is the output of the k-th feature map, x is the input, w is the kernel, and b is the bias.

3.3.2 Activation

After convolution, an activation function brings non-linear operations into the network and lets the CNN learn complex patterns. The most commonly applied activation is ReLU (or Rectified Linear Unit), which lets only positive values pass but turns negative outputs into zeroes.

$$f(x) = \max(0, x) \tag{4}$$

This plain but effective operation acts to improve the convergence of the model and helps avoid vanishing gradients during backpropagation.

3.3.3 Max Pooling

Max pooling is a down sampling technique that reduces spatial dimensions by selecting the maximum value within a pooling window. It helps in reducing computation, controlling overfitting, and preserving the most significant features from the feature maps.

$$y = \max\{x_1, x_2, \ldots, x_n\} \tag{5}$$

Where y is the pooled value and $x_1, x_2, \ldots, x_n$ are the values in the pooling region.

3.3.4 Fully Connected & Classification

The fully connected (dense) layer flattens the pooled feature maps and passes them to a set of neurons to compute the final classification. Each neuron in the output layer represents a class, and the neuron with the highest activation score determines the prediction label.

$$z_i = \sum_{j=1}^{n} w_{ij} a_j + b_i \tag{6}$$

Where z_i is the input to the softmax layer from neuron i, a_j are activations from the previous layer, w_{ij} is the weight, and b_i is the bias.

3.3.5Classification Loss (Cross-Entropy)

Using cross-entropy loss to measure the model's prediction error, this system can compute the difference between expected probabilities and class labels whereby training minimizes it via backpropagation.

$$\mathcal{L} = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{7}$$

Where y_i is the actual label (one-hot encoded), $\hat{y}_i$ is the predicted probability for class i, and C is the number of classes.

3.4 Feature Selection Optimization using Firefly

The Firefly Optimization Algorithm for feature selection is an intelligent metaheuristic method inspired by the flashing behaviour of fireflies, where each firefly corresponds to a probable solution-in this particular case, a subset of image features. Therefore, using this method for classifying TCGA_LUAD histopathological images will yield the most informative features such as texture, shape, or intensity descriptors-from the highly-dimensional feature space extracted by CNN layers. The algorithm evaluated each feature subset according to a fitness function, often represented by the classification accuracy obtained from such a criterion. Brighter fireflies i.e. those with better performing feature sets attract the rest towards the swarm in the direction of optimal or near-optimal solutions. The balance between exploration and exploitation is incorporated in this iterative movement of fireflies such that it discards redundant, noisy or non-informative features. Therefore, the selected features increase model accuracy, lessen computation time and improve generalization at cancer detection tasks

3.4.1 Initialization:

A population of fireflies is initialized, where each firefly represents a possible subset of features extracted from the CNN. These subsets can be encoded as binary vectors indicating whether a specific feature is selected

3.4.2 Fitness Evaluation:

Each firefly is evaluated using a fitness function that measures the performance of a classifier (e.g., CNN or SVM) using only the selected features. A common fitness metric is classification accuracy, calculated as:

$$\text{Fitness} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{8}$$

Where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives respectively.

3.4.3 Brightness Comparison:

Brighter fireflies (higher fitness) attract dimmer ones. Each firefly compares its fitness value with others to determine movement direction.

3.4.4Attractiveness Computation:

The attractiveness β of a firefly decreases with distance and is computed as:

$$\beta = \beta_0 e^{-\gamma r^2} \tag{9}$$

where:

- β_0 : initial attractiveness

- γ : light absorption coefficient
- r : Euclidean distance between two fireflies

3.4.5 Movement Update:

A firefly i is moved toward a more attractive firefly j using the rule:

$$x_i = x_i + \beta(x_j - x_i) + \alpha \cdot \text{ rand} \tag{10}$$

where α is the randomization parameter and rand are a random vector in $[0,1]$.

3.4.6 Iteration and Convergence:

Steps 2-5 are repeated for a fixed number of iterations or until convergence. The firefly with the best fitness value at the end represents the optimal feature subset.

3.5 Cloud Storage & Deployment

Cloud storage and deployment are needed to manage and scale machine learning applications in practice. A cloud-based CNN model: datasets and trained models for it will be stored inside a secure cloud environment such as AWS S3, Google Cloud Storage, or Azure Blob Storage, and they will be made available for access, scalability, and data protection. This cloud platform essentially provides seamless integration with the computational resources of GPUs or TPUs for effective training or inference. Once the model has been trained, it is deployed using the cloud services such as AWS Lambda or Google Cloud Functions that allow users to send their data to receive predictions to use it during RESTful APIs end-points. With respect to flexibility in scaling, low latency and high availability, cloud resource-sharing also ensures that all data safety regulations (e.g., HIPAA, GDPR) A key equation for cloud-based model deployment involves the data transfer rate R, which is the speed at which data is transferred between the user and the cloud server:

$$R = \frac{D}{T} \tag{11}$$

Where D is the data size (in bytes) and T is the time taken for the data transfer (in seconds). This equation ensures that data transmission is optimized for fast and efficient cloud model deployment.

4. Results And discussion

This three-dimensional graphic involves related visualizations indicating how the training time relates to the feature selection, data augmentation, and ultimately model accuracy. The first one shows the linear growth in the number of the selected features with respect to training time. More extended training periods allow the models to capture many relevant features that contribute to their performance enhancement. The second plot shows how data augmentation affects model accuracy quite evidently, where accuracy increases with an increase in training time, implying that these tricks of data augmentation, such as rotating, zooming, or flipping images, help in making the models robust in practice with longer training sessions. The plot, on the contrary, represents how training hours are enhanced by augmentations. It illustrates the gradual improvement of accuracy due to augmentations and reaffirms the importance of this technique for extensive training hours.

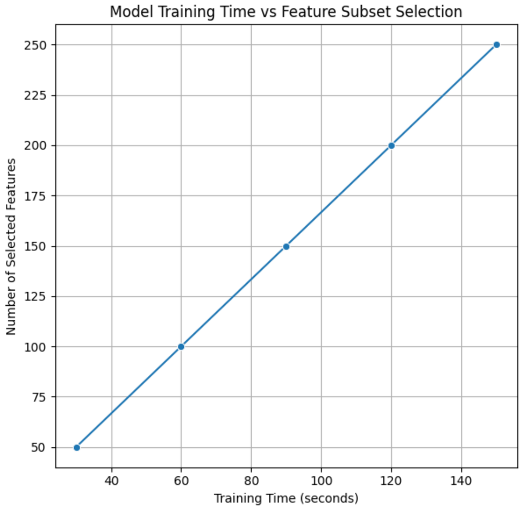

Figure 3: *Model Training Time vs Feature Subset Selection*

The graph shown on figure 3 depicts a linear relationship between training time of the model (x-axis) and number of features selected (y-axis). This, in turn, can be validated using the Firefly Optimization Algorithm. From this, it can be inferred that increasing time results in more features being selected, which translates into more training time allowing maximal numbers of features based on all possibly relevant attributes. The increased training duration might thereby enable the optimization algorithm to assess a wider portion of the feature set, allowing for a further distinction of attributes by models. Such a pattern suggests that time and computational resources investments matter in the development of more precise feature subsets and, thus, better models for specific tasks such as cancer detection.

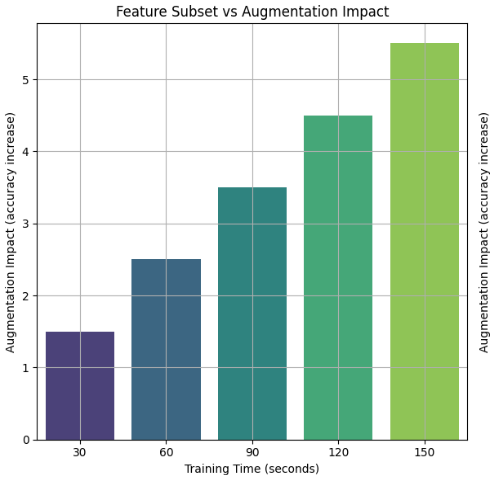

Figure 4: *Model Training Time vs Feature Subset Selection*

As figure 4 the training time extends, the effect of the augmentation becomes apparent on accuracy. For shorter training times (e.g., 30 seconds), augmentation impacts very little, while for increased training times (up to 150 seconds), the augmentation effect increases remarkably, proving that the model benefits more from augmentation when training runs longer. This implies that longer training gives the model a greater chance to exploit the augmented data, thus promoting the generalization ability and reducing overfitting, thereby improving the accuracy towards the cancer detection task.

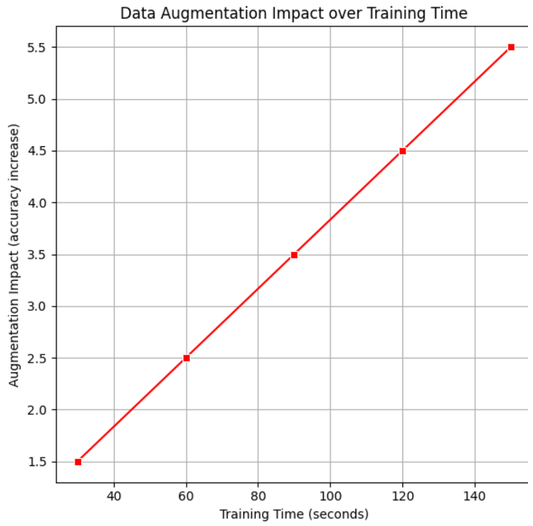

Figure 5: Data Augmentation Impact over Training Time

Thus figure 5 the plot reveals a linear relationship between training time and the data augmentation effect on model accuracy. As training time progresses, the impact of data augmentation on the model's accuracy also increases steadily. This trend indicates that given increased training time, the model maximizes the use of the augmented data-by, say, rotating it, flipping it, or zooming into it-for generalizing and making more accurate predictions. The model progressively becomes better at handling and taking advantage of the augmented data. Thus, longer training sessions would enable better-but more importantly, effective-learning from augmented variations, thus helping combat overfitting and improving the model performance on cancer detection tasks.

5. Conclusion and future enhancements

This is the research outcome of a cloud-based model of convolutional neural network (CNN) that should detect and diagnose lung cancer. With application of most advanced preprocessing techniques, such as data denoising and augmentation, it will increase input data quality and generalization, thence achieving correct classification of histopathological images. The Firefly Optimization Algorithm is integrated for selection of features, which beautifies the model additionally because it selects the relevant features to be considered, thus, reducing computational complexity and increasing the accuracy of classification. Also, since the system is deployed on the cloud, it enables scalability in handling the huge datasets-as well as allowing access by health care professionals to the model on real-time basis to fast and accurate diagnosis of cancer. Thus this is an important milestone in the direction of totally automating diagnosis in cancers where clinical workflow gets improved, and ultimately patient outcome through timely and accurate diagnostic support can be expected.

References

[1] Jayaprakasam, B. S., & Padmavathy, R. (2020). Autoencoder-based cloud framework for digital banking: A deep learning approach to fraud detection, risk analysis, and data security. International Research Journal of Education and Technology, 03(12).

[2] Masood, A., Yang, P., Sheng, B., Li, H., Li, P., Qin, J., ... & Feng, D. D. (2019). Cloud-based automated clinical decision support system for detection and diagnosis of lung cancer in chest CT. IEEE journal of translational engineering in health and medicine, 8, 1-13.

[3] Mandala, R. R., & Kumar, V. K. R. (2020). AI-driven health insurance prediction using graph neural networks and cloud integration. International Research Journal of Education and Technology, 03(10).

[4] Ganesan, M., Sivakumar, N., & Thirumaran, M. (2020). Internet of medical things with cloud-based e-health services for brain tumour detection model using deep convolution neural network. Electronic Government, an International Journal, 16(1-2), 69-83.

[5] Ubagaram, C., & Kurunthachalam, A. (2020). Bayesian-enhanced LSTM-GRU hybrid model for cloud-based stroke detection and early intervention. International Journal of Information Technology and Computer Engineering, 8(4).

[6] Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V., & Madabhushi, A. (2019). Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. Nature reviews Clinical oncology, 16(11), 703-715.

[7] Ganesan, S., & Hemnath, R. (2020). Blockchain-enhanced cloud and big data systems for trustworthy clinical decision-making. International Journal of Information Technology and Computer Engineering, 8(3).

[8] Shaukat, Z., Ali, S., Farooq, Q. U. A., Xiao, C., Sahiba, S., & Ditta, A. (2020). Cloud-based efficient scheme for handwritten digit recognition. Multimedia Tools and Applications, 79, 29537-29549.

[9] Musam, V. S., & Purandhar, N. (2020). Enhancing agile software testing: A hybrid approach with TDD and AI-driven self-healing tests. International Journal of Information Technology and Computer Engineering, 8(2).

[10] Abbas, S., & Mahmoud, A. M. (2020). Deep Learning-aided Brain Tumor Detection: An Initial Experience based Cloud Framework. Indonesian Journal of Electrical Engineering and Informatics (IJEEI), 8(4), 770-780.

[11] Musham, N. K., & Bharathidasan, S. (2020). Lightweight deep learning for efficient test case prioritization in software testing using MobileNet & TinyBERT. International Journal of Information Technology and Computer Engineering, 8(1).

[12] Ramar, V. A., & Rathna, S. (2018). AI-Driven Cloud-Based Deep Learning for Predictive Healthcare Analytics: Enhancing Disease Diagnosis with CNNs in Medical Imaging. Chinese Traditional Medicine Journal, 1(5), 12-18.

[13] Radhakrishnan, P., & Kumar, R. L. (2020). Advancing neurological disease prediction and management through AI and machine learning for improved patient outcomes. International Journal of Multidisciplinary Research and Growth Evaluation, 1(4), 59–66.

[14] Ganesan, M., Sivakumar, N., Thirumaran, M., & Saravanan, R. (2020). IOT And Cloud Based Brain Tumor Detection and Classification Model Using Optimal Densely Connected Convolutional Networks. IIOAB J, 11(5), 1-7.

[15] Gollavilli, V. S. B. H., & Pushpakumar, R. (2020). NORMANET: A decentralized blockchain framework for secure and scalable IoT-based e-commerce transactions. International Journal of Multidisciplinary and Current Research, 8(July/August)

[16] El Adoui, M., Mahmoudi, S. A., Larhmam, M. A., & Benjelloun, M. (2019). MRI breast tumor segmentation using different encoder and decoder CNN architectures. Computers, 8(3), 52.

[17] Induru, V., & Pushpakumar, R. (2020). Combining LSTM and GRU for efficient intrusion detection and alert correlation in cloud networks. International Journal of Multidisciplinary Research and Growth Evaluation, 1(1), 154–160.

[18] Khamparia, A., Gupta, D., de Albuquerque, V. H. C., Sangaiah, A. K., & Jhaveri, R. H. (2020). Internet of health things-driven deep learning system for detection and classification of cervical cells using transfer learning. The Journal of Supercomputing, 76, 8590-8608.

[19] Basani, D. K. R., & Kurunthachalam, A. (2020). Energy efficient signal processing for IoT-enabled robotic systems with challenges and solutions. International Journal of Multidisciplinary Research and Growth Evaluation, 1(2), 64–70

[20] Ghazvinian Zanjani, F., Zinger, S., Piepers, B., Mahmoudpour, S., Schelkens, P., & de With, P. H. (2019). Impact of JPEG 2000 compression on deep convolutional neural networks for metastatic cancer detection in histopathological images. Journal of Medical Imaging, 6(2), 027501-027501.

[21] Deevi, D. P., & Padmavathy, R. (2019). A hybrid random forest and GRU-based model for heart disease prediction using private cloud-hosted health data. International Journal of Applied Science Engineering and Management, 13(2).

[22] Chang, W. J., Chen, L. B., Chen, M. C., Chiu, Y. C., & Lin, J. Y. (2020). ScalpEye: A deep learning-based scalp hair inspection and diagnosis system for scalp health. Ieee Access, 8, 134826-134837.

[23] Ganesan, S., & Mekala, R. (2019). AI-driven drug discovery and personalized treatment using cloud computing. International Journal of Applied Science Engineering and Management, 13(3).

[24] Wang, H., Ding, S., Wu, D., Zhang, Y., & Yang, S. (2019). Smart connected electronic gastroscope system for gastric cancer screening using multi-column convolutional neural networks. International Journal of Production Research, 57(21), 6795-6806.

[25] Nagarajan, H., & Pushpakumar, R. (2019). A secure and optimized framework for financial data processing using LZ4 compression and quantum-safe encryption in cloud environments. Journal of Current Science, 7(1).

[26] Xie, L., Yang, S., Squirrell, D., & Vaghefi, E. (2020). Towards implementation of AI in New Zealand national diabetic screening program: Cloud-based, robust, and bespoke. Plos one, 15(4), e0225015.

[27] Jayaprakasam, B. S., & Jayanthi, S. (2019). Cloud-based real-time fraud detection using RNN and continuous model optimization for banking applications. Journal of Current Science, 7(2).

[28] Masood, A., Sheng, B., Yang, P., Li, P., Li, H., Kim, J., & Feng, D. D. (2020). Automated decision support system for lung cancer detection and classification via enhanced RFCN with multilayer fusion RPN. IEEE Transactions on Industrial Informatics, 16(12), 7791-7801.

[29] Ubagaram, C., & Bharathidasan. (2019). AI-driven cloud security framework for cyber threat detection and classification in banking systems. Journal of Current Science, 7(3).

[30] Trägårdh, E., Borrelli, P., Kaboteh, R., Gillberg, T., Ulén, J., Enqvist, O., & Edenbrandt, L. (2020). RECOMIA—a cloud-based platform for artificial intelligence research in nuclear medicine and radiology. EJNMMI physics, 7, 1-12.

[31] Vallu, V. R., & Arulkumaran, G. (2019). Enhancing compliance and security in cloud-based healthcare: A regulatory perspective using blockchain and RSA encryption. Journal of Current Science, 7(4).

[32] QIN, J., LANFRANCHI, V., & KIM, J. (2019). Cloud-Based Automated Clinical Decision Support System for Detection and Diagnosis of Lung Cancer in Chest CT.

[33] Garikipati, V., & Pushpakumar, R. (2019). Integrating cloud computing with predictive AI models for efficient fault detection in robotic software. International Journal of Engineering Science and Advanced Technology (IJESAT), 19(5).

[34] Mintz, Y., & Brodie, R. (2019). Introduction to artificial intelligence in medicine. Minimally Invasive Therapy & Allied Technologies, 28(2), 73-81.

[35] Ayyadurai, R., & Kurunthachalam, A. (2019). Enhancing financial security and fraud detection using AI. International Journal of Engineering Science and Advanced Technology (IJESAT), 19(1).

[36] Ameli, A., Hooshyar, A., El-Saadany, E. F., & Youssef, A. M. (2018). Attack detection and identification for automatic generation control systems. IEEE Transactions on Power Systems, 33(5), 4760-4774.

[37] Basani, D. K. R., & Bharathidasan, S. (2019). IoT-driven adaptive soil monitoring using hybrid hexagonal grid mapping and kriging-based terrain estimation for smart farming robots. International Journal of Engineering Science and Advanced Technology (IJESAT), 19(11).

[38] Ghoneim, A., Muhammad, G., & Hossain, M. S. (2020). Cervical cancer classification using convolutional neural networks and extreme learning machines. Future Generation Computer Systems, 102, 643-649.

[39] Kodadi, S., & Purandhar, N. (2019). Optimizing secure multi-party computation for healthcare data protection in the cloud using hybrid garbled circuits. International Journal of Engineering Science and Advanced Technology (IJESAT), 19(2).

[40] Hollon, T. C., Pandian, B., Adapa, A. R., Urias, E., Save, A. V., Khalsa, S. S. S., ... & Orringer, D. A. (2020). Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. Nature medicine, 26(1), 52-58.

[41] Gudivaka, R. K., & Rathna, S. (2018). Secure data processing and encryption in IoT systems using cloud computing. International Journal of Engineering Research and Science & Technology, 14(1).

[42] Uthoff, R. D., Song, B., Sunny, S., Patrick, S., Suresh, A., Kolur, T., ... & Liang, R. (2019). Small form factor, flexible, dual-modality handheld probe for smartphone-based, point-of-care oral and oropharyngeal cancer screening. Journal of biomedical optics, 24(10), 106003-106003.

[43] Panga, N. K. R. (2018). ENHANCING CUSTOMER PERSONALIZATION IN HEALTH INSURANCE PLANS USING VAE-LSTM AND PREDICTIVE ANALYTICS. International Journal of HRM and Organizational Behavior, 6(4), 12-19.

[44] Kurt, M. N., Ogundijo, O., Li, C., & Wang, X. (2018). Online cyber-attack detection in smart grid: A reinforcement learning approach. IEEE Transactions on Smart Grid, 10(5), 5174-5185.

[45] Peddi, S., & Aiswarya, RS. (2018). Securing healthcare in cloud-based storage for protecting sensitive patient data. International Journal of Information Technology and Computer Engineering, 6(1)

[46] Song, B., Sunny, S., Uthoff, R. D., Patrick, S., Suresh, A., Kolur, T., ... & Liang, R. (2018). Automatic classification of dual-modalilty, smartphone-based oral dysplasia and malignancy images using deep learning. Biomedical optics express, 9(11), 5318-5329.

[47] Kodadi, S., & Kumar, V. (2018). Lightweight deep learning for efficient bug prediction in software development and cloud-based code analysis. International Journal of Information Technology and Computer Engineering, 6(1).

[48] Massat, M. B. (2018). A Promising future for AI in breast cancer screening. Appl Radiol, 47(9), 22-25.

[49] Narla, S., & Kumar, R. L. (2018). Privacy-Preserving Personalized Healthcare Data in Cloud Environments via Secure Multi-Party Computation and Gradient Descent Optimization. Chinese Traditional Medicine Journal, 1(2), 13-19.

[50] Hoque, N., Kashyap, H., & Bhattacharyya, D. K. (2017). Real-time DDoS attack detection using FPGA. Computer Communications, 110, 48-58.

[51] Zhang, R., Cheng, C., Zhao, X., & Li, X. (2019). Multiscale mask R-CNN–based lung tumor detection using PET imaging. Molecular imaging, 18, 1536012119863531.