INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

A Novel Meta Heuristic Approach with Optimal Deep Learning Neural Network Based Oral Cancer Detection Model

A Subbulakshmi[1]*, Sridevanai Nagarajan[2]

1 Assistant Professor, Sri Kanyaka Arts & Science college for women, Madras University, Chennai, Tamil Nadu, India
2 Assistant Professor, Hindustan Institute of Science and Technology, Tamil Nadu, India

OPEN ACCESS

Received: 24-08-2023
Accepted: 24-10-2023
Published: 05-12-2023

Citation: Subbulakshmi A, Nagarajan S (2023) A Novel Meta Heuristic Approach with Optimal Deep Learning Neural Network Based Oral Cancer Detection Model. Indian Journal of Science and Technology 16(45): 4164-4176. https://doi.org/10.17485/IJST/v16i45.2154

* **Corresponding author**.

asubbulaklshmi6@gmail.com, y2subu@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: This study proposes a new approach to improve oral cancer detection in medical images by utilizing a Deep Convolutional Neural Network (DCNN) and an optimized Long Short-Term Memory (LSTM) technique. **Methods:** First, the input oral squamous cell carcinoma images are pre-processed using median filtering as well as CLAHE. Next, feature extraction is performed using the Local Tetra Pattern (LTrP) to extract different features. The HHHLO algorithm is then applied to select the optimal features for the subsequent feature selection process. Finally, the selected features are classified using a hybrid classifier called DCNN-LSTM, which predicts the diagnosis of patients with oral cancer. The investigation of the DCNN-LSTM model involves conducting experiments on a commonly used biomedical image dataset that is readily accessible through the Kaggle repository. **Findings:** The proposed method was implemented on the MATLAB platform, and its performance was evaluated using various metrics. The results demonstrated the superiority of the DCNN-LSTM model over existing methods, achieving a maximum accuracy of 0.975. **Novelty:** Oral cancer is a common and formidable type of cancer associated with a significant mortality rate.

Keywords: Oral cancer; Improved Squirrel Search Algorithm (ISSA); Deep Convolutional Neural Network (DCNN); Contrast Limited Adaptive Histogram Equalization (CLAHE); Hybrid Horse herd lion optimization (HHHLO)

1 Introduction

Oral cancer constitutes a grave health concern, characterized by malignant growths affecting various anatomical structures within the oral cavity, including the lip, buccal mucosa, hard palate, floor of the mouth, gingiva, and the anterior two-thirds of the tongue. Globally, oral cancer ranks as the sixth most common cancer, presenting formidable challenges in terms of both mortality and morbidity. As of 2020, a staggering 377,713 new cases of oral cancer were reported, resulting in 177,757 global deaths [(1)]. Despite the availability of diverse treatment modalities, the overall mortality rate remains strikingly high, with five-year survival rates ranging from a mere 28% to a relatively optimistic 67%. Early-stage oral cancer (Stage I) does exhibit more favorable

survival rates, hovering at approximately 81%. However, advanced-stage cases (Stage 2 and 3) present considerably grimmer survival prospects, with rates plummeting to less than 20%. It is important to note that over 90% of oral cancer cases fall into the category of oral squamous cell carcinoma (OSCC), often originating from oral potentially malignant disorders (OPMD) such as leukoplakia and erythroplakia. Detecting these OPMD, which bear a risk of progressing to malignancy, emerges as a pivotal strategy for reducing morbidity and mortality associated with oral cancer. In response to this pressing concern, screening programs have prioritized the early identification of these potentially malignant disorders. Nevertheless, the primary method employed in these programs, visual examinations, encounters substantial challenges when applied in real-world healthcare settings. Visual examinations are typically conducted by primary care healthcare professionals who, unfortunately, may lack extensive training and experience in recognizing these lesions[(2)]. Furthermore, the wide variation in the appearance of oral lesions further complicates their identification, leading to significant delays in referrals to oral cancer specialists. Adding to the complexity, early-stage OSCC lesions and OPMD often remain asymptomatic and may appear as innocuous, benign lesions, contributing to delayed diagnoses. It is here that the importance of early detection assumes a pivotal role in improving the survival prospects of individuals at risk[(3)]. Early identification, particularly of those lesions showing histological evidence of dysplasia, when coupled with essential lifestyle changes such as smoking cessation, quitting tobacco use, reducing alcohol consumption, and implementing appropriate surgical management, can significantly reduce the risk of malignant transformation.

Oral cancer maintains its reputation for notably low survival rates, standing at just 30% presently. Even with timely detection, the improvement in survival rates remains modest, showing only a 10% to 15% enhancement. It's crucial to emphasize that squamous cell carcinomas are responsible for over 90% of oral cancer cases. Early-stage oral cancer is characterized by symptoms such as whitish or reddish patches in the oral cavity, persistent wounds, sores, blisters, or abnormal growths that do not heal within two weeks, as well as discomfort during swallowing. Tissue changes associated with oral cancer include the loss of cellular polarity, irregular maturation of cells from basal to squamous cells, heightened cellular density, premature keratinization, and the presence of keratin pearls in the deep epithelial layer[(4)]. Timely identification of cancerous lesions can lead to a lower disease stage and reduced mortality rates.

Given these circumstances, the importance of early identification of OSCC is paramount, not only for improved diagnosis and treatment but also for overall survival[(5)]. Late diagnosis hampers the progress of precision medicine, despite advancements in our understanding of the molecular mechanisms of cancer. This is where the promising realm of machine learning (ML) and deep learning (DL) models enters the scene. ML techniques, including shallow learning, have demonstrated considerable potential for improved prognostication in oral cancer compared to conventional statistical analysis. These techniques offer promising outcomes as they can capture complex relationships within the dataset[(6)]. The application of ML approaches in cancer prognostication has gained considerable interest in recent decades, as it aids clinicians in making informed decisions and improving patient management. Furthermore, the evolution of shallow ML into deep learning (DL) has shown significant promise for enhancing cancer management. It is within this broader context that various methods have been developed by researchers to predict the overall survival (OS) of oral cancer. For instance, Bansal et al., [(7)] introduced a novel approach to early prediction and diagnosis of various oral cancer types, combining deep transfer learning techniques with hybrid optimization methods. Their method encompassed image preprocessing to eliminate noise, feature extraction through morphological operations, and the utilization of deep transfer learning models optimized for superior performance. Huang et al., [(8)] proposed an optimal deep learning neural network for oral cancer diagnosis, employing the Improved Squirrel Search Algorithm (ISSA) to select network weights that enhance accuracy. Myriam et al., [(9)] investigated oral cancer detection using a combination of convolutional neural networks (CNN) and an optimized deep belief network (DBN). They optimized the CNN and DBN design parameters using a hybrid optimization algorithm that integrated Particle Swarm Optimization (PSO) and Al-Biruni Earth Radius (BER) Optimization. In a different study, Ding et al., [(10)] explored a modified Locust Swarm optimizer for oral cancer diagnosis. Their methodology involved image segmentation based on reinforcement learning, image feature extraction using the Gabor wavelet transform, and final classification using an RBF-kernel-based SVM. Additionally, they incorporated a modified metaheuristic technique known as Modified Locust Swarm Optimization (MLSO) for feature selection. These studies highlight the diverse range of approaches, including deep learning, transfer learning, optimization algorithms, and image processing techniques, utilized to improve the prediction and diagnosis of oral cancer. This paper introduces two novel approaches based on deep learning models for segmenting oral cancer and classifying different types of oral cancer.

1.1 Contributions of our work

(i) The main objective of this study is to achieve accurate prediction of overall survival (OS) in individual patients with oral cancer using the DCNN-LSTM approach.

(ii) This paper introduces a pre-processing mechanism to enhance the efficiency of oral cancer detection and classification.

(iii) By employing the Hybrid Horse herd lion optimization (HHHLO) approach, the attribute selection process aims to reduce calculation time.

(iv) An efficient unsupervised deep learning algorithm called DCNN is generated for the classification of oral cancer.

(v) The developed system aims to provide predictions for the overall survival (OS) rate in the process of oral cancer prediction.

The proposed research methodology is structured as follows: in section 1, the existing research methodologies related to oral cancer based overall survival analysis are explained; in section 2, the proposed methodology has been explained; in section 3, the experimental analysis of the proposed methodology is given; and in section 4, the proposed research is concluded with the possibility of future improvement.

2 Methodology

The research methodology introduces a novel approach for oral cancer detection using DCNN with improved accuracy. The process begins by collecting input data from patients. The collected data undergoes pre-processing to address missing values and redundant data. Missing values are imputed using similar attribute values from other patients, and redundant data is eliminated. Additionally, the data is normalized to reduce prediction errors. Following the pre-processing stage, the oral cancer classification is performed using a DCNN-LSTM model. This model predicts the presence or absence of oral cancer based on the processed data. The performance of the proposed research methodology is then analyzed, evaluating its accuracy and effectiveness in oral cancer detection. The methodology utilizes a block diagram, as shown in Figure Figure 1 , to illustrate the different stages involved in the research process.

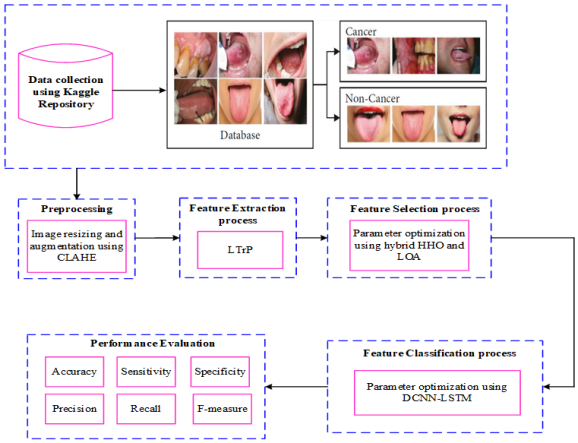

Fig 1. Block diagram for the proposed research methodology

2.1 Input image

The classification process in this study focuses on utilizing an oral image dataset, as mentioned by[11].

$$Q = (S_1, S_2, \cdots, S_x \ldots S_t) \tag{1}$$

In the dataset, $S_t$represents the total number of images, and $S_x$denotes the specific xth image index that is processed during the pre-processing phase.

2.2 Pre-processing

During this stage, we perform preprocessing on the provided image. Initially, the image is cropped, and resized ensuring that all images have consistent dimensions, which is essential for feeding them into the deep learning model, and the pixel values were adjusted to a standard range (i.e., 0 to 1) to enhance model stability during training, subsequently, we employ the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique to enhance the image's contrast level. Increased contrast in the

image provides additional information about noise, gamma correction, and data augmentation, thereby enhancing the quality of the original photos. However, adding noise to the input picture during improvement is possible. This study approach sets the limit based on an average estimate of the intensity of the picture pixels to address that issue. The CLAHE is the name given to the suggested research approach.

Consider an image *S*with discrete gray levels that range in intensity from zero (the lowest level) to one hundred (the highest level) $(L-1)$. The following transformation function$Q_{tf}(i)$ is created using the probability density function$(L-1)$:

$$X_i = Q_{tf}(i) = \sum_{x=0}^{i} P_{df}(i) = \sum_{x=0}^{i} \frac{t_i}{t} \text{ where } 0 \leq i \leq L-1 \tag{2}$$

Where $t_i$is the percentage of gray-valued pixels i, L is the total number of gray levels and t is the total number of pixels in the image. Adaptive histogram equalization addresses the issue of over-brightness caused by the enhancement of background contrast and loss of information. It achieves this by dividing the image into smaller rectangular areas known as tiles and performing histogram equalization on each tile individually. This local contrast enhancement, however, can lead to the amplification of noise. To mitigate this drawback, contrast limiting is employed[12].

2.3 Feature Extraction

Segmentation plays a crucial role in the identification of oral cancer. The initial step involves dividing the input image into multiple pieces through the process of segmentation. Following pre-processing, various features such as form characteristics, texture, and edge features are extracted. One such feature is the Local Tetra Pattern (LTrP), which captures information based on the distribution of coded edges in two directions (positive or negative), it divides the image into local neighborhoods, typically with a specific radius. For each pixel in the image, it looks at the relationship between this pixel and its neighbors within the defined radius. LTrP encodes the relationships between the pixel and its neighbors into a pattern. These patterns are then used as feature vectors. However, there is room for improvement in terms of splitting edges in more than just two directions[13]

2.4 Feature Selection using hybrid Horse Herd Optimization and Lion optimizer algorithm

To enhance classification performance, a novel optimization algorithm is introduced by combining the Horse Herd Optimization (HHO) and Lion Optimizer Algorithm (LOA). The selection of HHO is motivated by its computational efficiency and adaptability to various metaheuristic approaches. In this approach, features are chosen from the extracted feature set using the Hybrid Horse Herd Lion Optimization (HHHLO) technique. The Horse Herd Optimization algorithm is designed based on the social behaviors of horses at different stages of life, utilizing six key characteristics: grazing, hierarchy, sociability, imitation, defense mechanism, and roaming. It explores optimal solutions by incorporating two distinctive behaviors of lions: territorial defense and territorial takeover[14]. Territorial defense occurs between resident males and wandering males, while territorial takeover happens when a mature resident male replaces an old resident male. However, the HHO algorithm may encounter local optima issues when dealing with complex data, leading to poor search capabilities. To address this challenge, the proposed research methodology incorporates the Lion Optimizer process to update positions. This approach combines the strengths of both algorithms, providing a global search capability with a fast convergence rate. It effectively harnesses the high search efficiency of HHO while benefiting from the dynamic optimization capability of the Lion Optimizer Algorithm. In this method, the selected features from each image are treated as "Horses." The first step of the Horse Optimization Algorithm involves number of features, size, randomly generating a population across the solution space. Each solution is represented by:

$$\delta^{x}_{I,Age} = \vec{V}^{x}_{I,Age} + \delta^{x}_{(I-1),Age} Age = \lambda, \mu, \eta, \omega \tag{3}$$

Where, $\delta^{x}_{I,Age}$is the xth horse's location, $\overrightarrow{V}^{x}_{I,Age}$is its vector of velocity of the xth horse, and Age is the current iteration and *I* refers to the horse's age range

Grazing: Horses, similar to herbivorous creatures, depend on plants for their survival. Horses engage in grazing behavior throughout their lifespan, starting from a young age and continuing indefinitely. It has been observed that horses graze within their local vicinity, *gr* as indicated by the coefficient. Equation (4) specifically characterizes the social aspects of grazing behavior.

$$\overrightarrow{GR}^{x}_{I,Age} = g\vec{r}^{x}_{I,Age}(u+l)\left[\delta^{x}_{(I-1)}\right] \text{ Age } = \lambda, \mu, \eta, \omega \tag{4}$$

$$g\vec{r}^{x}_{I,Age} = g\vec{r}^{x}_{(I-1),Age}\delta_{\alpha_{gr}} \tag{5}$$

Where, $\delta^{x}_{(I-1)}$is the horse's propensity to graze and $\overrightarrow{GR}^{x}_{I,Age}$is the xth horse's grazing motion parameter at the Ith iteration. $\overrightarrow{GR}^{x}_{I,Age}$ is linearly decreased at each iteration loop according to a reduction factor defined byα_{gr}. l and u represent the bottom and upper grazing spaces, respectively.

Hierarchy: The population of horses is split into two groups: leaders and followers. Follower horses are always drawn to leaders.

$$\vec{H}^{x}_{I,Age} = \vec{h}^{x}_{I,Age}\left[\delta^{\text{best}}_{(I-1)} - \delta^{x}_{(I-1)}\right] \text{ Age } = \lambda, \mu, \eta \tag{6}$$

$$\overrightarrow{h}^{x}_{I,Age} = \overrightarrow{h}^{x}_{(I-1),Age}\delta_{\alpha_h} \tag{7}$$

Sociability: The way of life of horses is similar to that of other social animals. They form a group to improve their chances of surviving and to facilitate escape from assailants.

$$\vec{S}^{x}_{I,Age} = \vec{S}^{x}_{I,Age}\left[\left(\frac{1}{T}\sum_{i=1}^{T}\delta^{i}_{(I-1)}\right)\right] \text{ Age } = \mu, \eta \tag{8}$$

$$\overrightarrow{s}^{x}_{I,Age} = \overrightarrow{s}^{x}_{(I-1),Age}\delta_{\alpha_s} \tag{9}$$

Where T represents the size of the population. The x$^{\text{th}}$ horse's social motion vector at the I$^{\text{th}}$ iteration is represented by $\overrightarrow{S}^{x}_{I,Age}$ and is linearly decreased by a reduction factor represented by α_s. The $\overrightarrow{S}^{x}_{I,Age}$shows the horse's fascination in the herd in iteration I. For horses of age $\mu$and η, it is set to 0.2 and 0.1, respectively. The value of $\overrightarrow{s}^{x}_{I,Age}$ is updated using equation (9).

Imitation: A horse is capable of mimicking the actions of other horses. The mathematical formula for the social behavior of imitation is as follows:

$$I\vec{M}^{x}_{I,Age} = i\vec{m}^{x}_{I,Age}\left[\left(\frac{1}{P^{T}_{\text{best}}}\sum_{i=1}^{P^{T}_{\text{best}}}\hat{\delta}^{i}_{(I-1)}\right) - \delta^{x}_{(I-1)}\right] \text{ Age } = \eta \tag{10}$$

$$i\vec{m}^{x}_{I,Age} = i\vec{m}^{x}_{(I-1),Age}\delta_{\alpha_{im}} \tag{11}$$

The P^{T}_{best} represents the count of horses within the current population that occupy the most advantageous positions in terms of fitness.

Defense: The defense mechanism of horses is designed to enable them to evade horses in unfavorable positions, which are significantly distant from optimal ones. Equation (12) demonstrates that the coefficient d is assigned a negative value to prevent the current horse from occupying undesirable locations.

$$\vec{D}^{x}_{I,Age} = \vec{d}^{x}_{I,\text{ Age}}\left[\left(\frac{1}{P^{T}_{\text{worst}}}\sum_{i=1}^{P^{T}_{\text{worst}}}\hat{\delta}^{i}_{(I-1)}\right) - \delta^{x}_{(I-1)}\right] \text{ Age } = \lambda, \mu, \eta \tag{12}$$

$$\overrightarrow{d}^{x}_{I,Age} = \overrightarrow{d}^{x}_{(I-1),Age}\delta_{\alpha_d} \tag{13}$$

Roam: The migration of horses from one grazing area to another in search of nourishment can be simulated by modeling their movement as random transitions between locations.

$$\overrightarrow{R}^{x}_{I,Age} = \overrightarrow{r}^{x}_{I,Age} * P\delta^{x}_{(I-1)} Age = \eta, \omega \tag{14}$$

$$\overrightarrow{r}^{x}_{I,Age} = \overrightarrow{r}^{x}_{(I-1),Age}\delta_{\alpha_r} \tag{15}$$

In equation (15), the value of $\overrightarrow{r}^{x}_{I,Age}$ is updated by incorporating a reduction factor represented byα_r .

Upon the initialization of optimization variables for the HHO population, the fitness function is employed to evaluate the effectiveness of the solutions generated by the optimization algorithm. The mathematical representation of this fitness function, denoted as (FF), is presented below:

$$FF = \alpha_1 \text{ Error } + \alpha_2 \frac{|s_f|}{|N_f|} \tag{16}$$

Here, Error refers to the number of features, S_f indicates the number of chosen features, and N_f reflects the classification error. The range [0, 1] includes $\alpha_1 = 1 - \alpha_2$ the factors $\alpha_1$and α_2. Additionally, each variable set taken into account for the HHO population was computed. The HHO population is updated and the variables set is determined once again, this time randomly, employing the Lion Optimization Algorithm. The effectiveness of each solution is assessed by evaluating their respective fitness functions, enabling the identification of the best solution. Subsequently, the current populations are updated using either the HHO or Lion Optimization Algorithm, depending on the reliability of the fitness function. If the probability of fitness is greater than 0.05, HHO is utilized; otherwise, LOA is employed to update the current solutions. After updating the population, the best result is determined based on the individual solution's fitness function. The final step involves verifying whether the stopping conditions are met. If they are satisfied, the algorithm returns the best possible solution. Otherwise, the preceding steps, starting from the computation of the probability, are repeated until the stopping conditions are met. The motion vector of horses at different ages during each cycle of the algorithm can be represented by equation (17), considering the aforementioned behavior patterns.

$$\begin{aligned}
\vec{V}^x_{I,\lambda} &= \overrightarrow{GR}^x_{I,\lambda} + \vec{D}^x_{I,\lambda} \\
\vec{V}^x_{I,\mu} &= \overrightarrow{GR}^x_{I,\mu} + \vec{H}^x_{I,\mu} + \vec{S}^x_{I,\mu} + \vec{D}^x_{I,\mu} \\
\vec{V}^x_{I,\eta} &= \overrightarrow{GR}^x_{I,\eta} + \vec{H}^x_{I,\eta} + \vec{S}^x_{I,\eta} + I\vec{M}^x_{I,\eta} + \vec{D}^x_{I,\eta} + \vec{R}^x_{I,\eta} \\
\vec{V}^x_{I,\omega} &= \overrightarrow{GR}^x_{I,\omega} + I\vec{M}^x_{I,\omega} + \vec{R}^x_{I,\omega}
\end{aligned} \tag{17}$$

Equation (17) performs the updating process in the horse herd optimization method if$f\left(P^T_{new}\right) < f\left(P^T_{best}\right)$, and Equation (20) is used by the female lion if not. A fertility assessment is carried out after the pride generation process. The lion is the laggard in this situation, and the laggardness rate is described as being improved by oneR_{Lag}. If $FF\ (L_{Male})$is becoming better, the phrase "reference fitness"FF^* is used. The territorial defense $R^{Max}_{Lad}$is initiated when the laggardness rate exceeds a certain threshold. The lioness's fertility is also affected by the sterility rateR_{Str}, which is increased by 1 when R_{Str} it surpasses R^{Max}_{Str}. Equations (18), (19), and (20) are used to update the lioness in this situation.

$$L^{\text{Female}+}_p = \begin{cases} L^{\text{Female ifp}}_q = q \\ L^{\text{Female}}_p \text{ otherwise} \end{cases} \tag{18}$$

$$L^{\text{Female}}_q = \min\left[L^{\text{Max}}_q, \max\left(L^{\text{Max}}_q, \Delta_q\right)\right] \tag{19}$$

$$\Delta_q = \left[L^{\text{Female}}_q + (0.1\,\text{rand}_2 - 0.05)\left(L^{\text{Male}}_p - \text{rand}_1\, L^{\text{Female}}_p\right)\right] \tag{20}$$

Where Δ the term denotes the female updating function, the random numbers between 0 and 1 are represented by rand_1 and rand_2 . and the random number between 1 and Length, which is written as q. When the changed lioness $L^{\text{Female}+}$ is considered L^{Female} mating is carried out for improvement. Once the updating procedure is complete, if a replacement $L^{\text{Female}+}$ is not available L^{Female} it is confirmed that the location is fertile in order to produce the best cubs.

2.5 Oral cancer prediction using DCNN-LSTM

In this research methodology, a Deep Convolutional Neural Network (DCNN) is applied for the prediction of oral cancer. The DCNN model is specifically designed for processing and analyzing visual data, such as images and videos. DCNNs have revolutionized computer vision tasks and are widely used in various applications, including image classification, object detection, facial recognition, and medical image analysis. In this research the DCNN is utilized to extract pertinent information from the oral cancer dataset. However, the convolution neural network algorithm encountered difficulties in aligning the input data, resulting in increased computational complexity. To address this alignment issues in the input data, we introduce an aligned layer processed with the assistance of Long Short-Term Memory (LSTM) which is a type of recurrent neural network

(RNN) architecture designed to handle sequences of data effectively reducing computational complexity and delivering superior performance for survival prediction.

In the research methodology, the determination of the number of pooling layers is based on the utilization of average pooling. The DCNN typically consists of convolution layers, pooling layers, and possibly fully connected layers. The architecture includes the number of layers, filter sizes, and the number of channels in each layer. We use two convolution layers: for the DCNN architecture the first with 64 channels of size 5×5 and the second with 128 channels of size 5×5. We employ a channel pitch of 2 pixels and a maximum pool channel size of 2×2. After max pooling, the weights are normalized and managed at a fully connected layer. Following this, we introduce a SoftMax layer after two fully connected layers for performance evaluation [15]. In order to train the LSTM, a memory compiler technique is employed, which iteratively updates the network weights based on the training data. This approach is chosen as a ReLU layer is lighter in weight compared to LSTM while still retaining the beneficial qualities of LSTM. Subsequently, a dense layer, serving as a fully connected layer, is introduced. The dense layer calculates all potential dot products, and to resolve the compatibility issues with the preceding layers, we introduce a flattened layer. This enables the output from the ReLU layer to be seamlessly fed into the LSTM. Finally, for class score computation, we employ a dense output layer with a sigmoid activation function. The rectified linear activation function, which is a simple calculation returning the input value, is used within the neural network to enhance its ability to understand intricate data patterns. As illustrated in Figure 2 , each of the hidden layers makes advantage of the ReLU implementation work. It is depicted as,

$$F(\Psi) = Max(0, \Psi)$$

Where $\Psi$are the neurons' contributions. The ReLU improvement work no longer includes unlimited initiation work. The connecting layer links the numerous features that the additional component provides. The inclusion of neighborhood response standardization following each convolutional layer in the activation function helps mitigate the issue of overfitting within the model. The solution is in equation (22).

$$\Psi_i = \frac{\Psi_i}{\left(\sigma + \left(\rho \sum_j \Psi_j^2\right)\right)^{\nu}}$$

The variables in equation (22), Ψ_i stand for input pixel esteem and hyperparameters are σ, $\rho$and $\nu \in Q$, respectively.

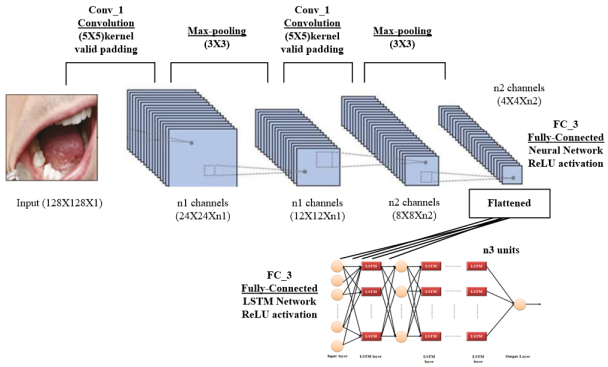

Fig 2. Structure for DCNN-LSTM network

The forget gate erases information that is no longer necessary to understand the state of the cell. In this layer, two inputs are multiplied by weight matrices after which a bias is added. The forgetting process is described in more detail in Equation (23) below:

$$F_l = \varpi\left(\varepsilon_f\left(o_{h-1}, \omega_i\right] + b_f\right) \tag{23}$$

Where, $F_l$denotes the forget gate output, $\varepsilon_f$and b_f indicates the weight value and bias values for input value, ϖ specifies the sigmoid function, and o_{h-1} represents the output of the last LSTM unit at time $h-1$. The block's input and memory are used to determine the output. The expression for the output layer is derived in equation (24),

$$A_l = \varpi\left(\varepsilon_a\left(o_{h-1}, \omega_i\right] + b_a\right) \tag{24}$$

Where, $A_l$denotes the output gate, ε_a and b_a indicates the weight and bias values of the output layer in equation (24). Thus, the DCNN-LSTM predicts the oral cancer as malicious and normal.

The DCNN-LSTM architecture is chosen because of its unique ability to harness spatial information from medical images through the DCNN while simultaneously modelling temporal dependencies within sequences of images using LSTM. To optimize its performance for the specific task of oral cancer prediction, the DCNN undergoes a meticulous process of pretraining and fine-tuning, ensuring it focuses on the most relevant image features. Simultaneously, the LSTM component is meticulously adapted to manage the temporal aspect of the data, effectively capturing how patterns and features evolve over time in the context of oral cancer. This combination of DCNN and LSTM within the DCNN-LSTM architecture creates a synergistic model that takes full advantage of both spatial and temporal information. Such an approach makes it exceptionally well-suited for the intricate task of oral cancer prediction, with the choice of parameters and adaptation strategies carefully tailored to the dataset and the specific research objectives at hand.

3 Results and Discussion

The effectiveness of the suggested DCNN-LSTM technique is assessed in this section. An Intel Core i5-2450M CPU laptop with 6GB RAM is used to evaluate the implementation of the proposed MRI picture semantic segmentation. The MATLAB R2022b program is used to carry out this procedure. A variety of evaluation criteria are used to evaluate the success of the proposed attempt for oral cancer categorization to existing techniques. There is discussion of the model estimates and results.

3.1 Dataset description

A benchmark dataset from the Kaggle repository is used to validate the DCNN-LSTM model's performance in classifying OCs (available at https://www.kaggle.com/shivam17299/oral-cancer-lips-and-tongue-images) [16]. The data set contains images of lips and tongue which are classified in cancerous and non-cancerous groups which of size 29 Mega Bytes in total the images in various ENT hospitals of Ahmedabad and classified the images with the help of ENT doctors. Figure 3 presents images of both malignant and non-cancerous mouth and tongue samples from our collection, totalling 131 photos. Among these, 87 are categorized as malignant, while 44 are non-cancerous. We applied data augmentation techniques to expand this dataset, resulting in a total of 1310 photos. Subsequently, we divided this image dataset into two subsets: a training set, which constitutes 90% of the data (with 10% reserved for validation), and a testing set, comprising the remaining 10%.

Fig 3. Depicts a sample set of tongue images

3.2 Performance analysis for Oral cancer classification

The oral cancer to accurately predict overall survival (OS) rate by using proposed DCNN-LSTM is analysed with the existing Deep Neural Network (DNN), Artificial Neural Network (ANN) and K-Nearest Neighbors (K-NN) according to the precision, recall, accuracy, specificity, sensitivity, and F-Measure metrics. In the context of image classification, we compared the predicted class with the expected class, determined through composite annotation. Binary image classification yielded results in the form of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). To evaluate the effectiveness of our

proposed system, we employed the following metrics:

Accuracy:

The gathering of general data on TP and TN includes the correctness of the suggested study.

$$UC = \frac{\mathrm{TN} + \mathrm{TP}}{(\mathrm{TN} + \mathrm{TP} + \mathrm{FN} + \mathrm{FP})} * 100\% \tag{25}$$

Sensitivity:

Sensitivity is a metric used to determine the proportion of accurately detected instances to real positives.

$$\mathrm{SEN} = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FN})} * 100\% \tag{26}$$

TP stands for a genuine positive, and FN for a false negative.

Specificity:

The ratio between an occurrence that was properly detected and an actual negative may be used to gauge specificity.

$$SPE = \frac{\mathrm{TN}}{(\mathrm{TN} + \mathrm{FP})} * 100\% \tag{27}$$

F-score is used to determine the developed model's correctness. We discover that accuracy and recall must be calculated in order to get the F-score.

$$\mathrm{PRE} = \frac{1}{\Omega} * [(TP) \cdot (TP + FP)] \tag{28}$$

$$\mathrm{REC} = \frac{1}{\Omega} * [(TP) \cdot (TP + FN)] \tag{29}$$

$$\mathrm{F1} = \frac{2 * PRE * REC}{PRE + REC} \tag{30}$$

Where, Ω defines the class count.

In Figure 4 shows that the comparison analysis of performance metrics has been illustrated. In Figure 4 (a) shows that the accuracy analysis of the proposed DCNN-LSTM based oral cancer classification is compared with the existing DNN, ANN and KNN approaches with respect to the accuracy metrics. Here, the accuracy of the proposed methodology is above 0.97 but other three methods have 0.95 for DNN, 0.92 for ANN and 0.9 for the KNN approach. The sensitivity and specificity analysis of the proposed and existing DNN, ANN and KNN is showed in Figure 4 (b & c). Here the sensitivity and specificity of proposed DCNN-LSTM method is 0.94 and 0.97. And also, the sensitivity and specificity value are worst for the existing KNN when compared with the existing DNN and ANN approach. But the DNN and ANN methods also has the lowest sensitivity and specificity value. Thus, the graphical representation shows that the proposed approach achieves better performance than the existing research approaches.

In Figure 5 shows that comparison analysis of precision, recall and F-measure has been illustrated. The precision and recall value of the proposed method is above 0.9 but the DNN value is below 0.8, ANN value is below 0.7 and KNN value is below 0.6 attains less precision, recall and F-measure values when compared to the proposed research approach. Because the suggested study approach employs a distribution function-based output layer and a superior activation function in the oral cancer survival rate prediction. The F-Measure is the combination of precision and recall metric. The F-measure of the proposed method is above 0.945 higher for the proposed method than the existing methodologies such as DNN method is 0.83, ANN approach is 0.68 and KNN technique is 0.64 respectively. As a result, the graphical depiction demonstrates that the proposed technique outperforms the already used existing approaches.

Figure 6 (a) shows a comparative study of the error values of many research and existing approaches, including DNN, ANN and KNN approach. As shown in Table 1 , to illustrate the comparison analysis of previous work. The DCNN-LDTM model's precision-recall curves under different TR/TS data size conditions are shown in Figure 6 (b). The results showed that the DCNN-LDTM model has classified data with a maximum of OC under two categories, namely cancer and non-cancer. Under two classes, it has been noted that the DCNN-LDTM model has shown superior precision-recall values. The proposed method's

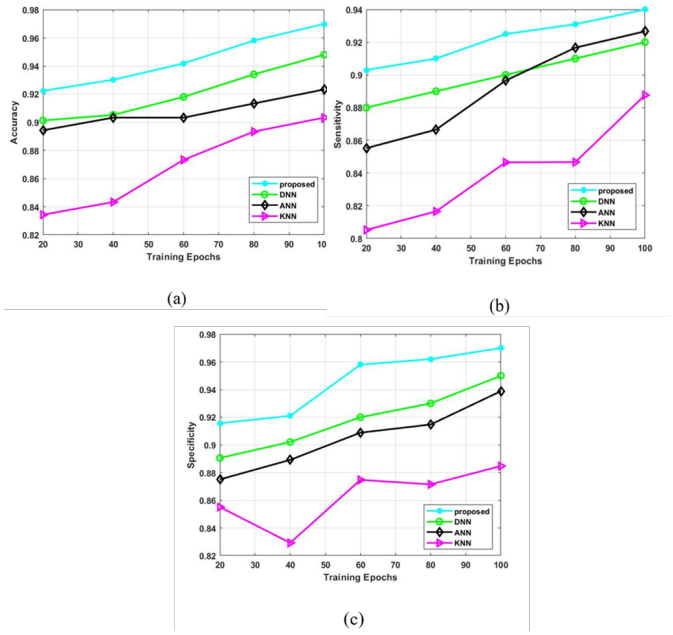

Fig 4. Comparison analysis of (a) Accuracy (b) Sensitivityand (c) Specificity

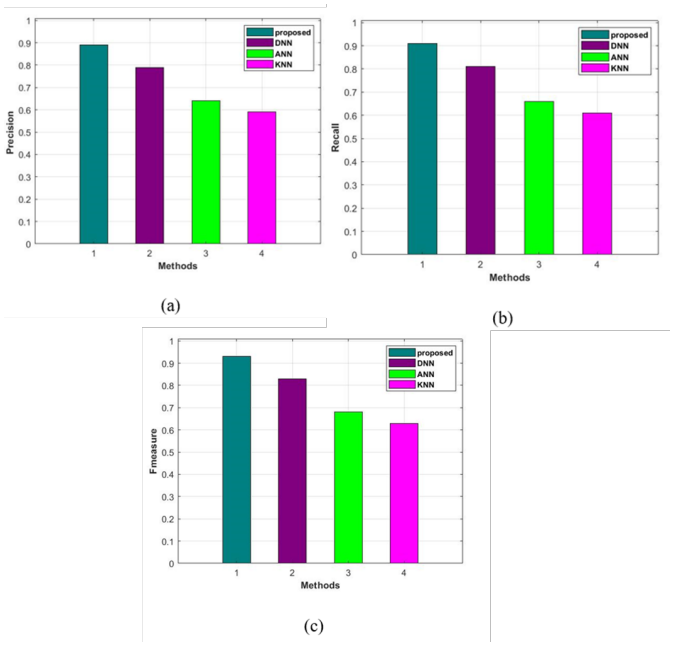

Fig 5. Comparative analysis of the proposed approach with the existing DNN, ANN and KNN based on (a) Precision (b) Recall and (c) F-measure

Table 1. Comparison analysis of previous published work

Methods	Accuracy (%)	Sensitivity (%)	Precision (%)
LD [9]	0.87	0.81	0.87
DT [9]	0.86	0.75	0.82
ResNet50 [7]	0.92	-	-
DNN [10]	0.96	0. 93	-
Proposed	**0.97**	**0.93**	**0.89**

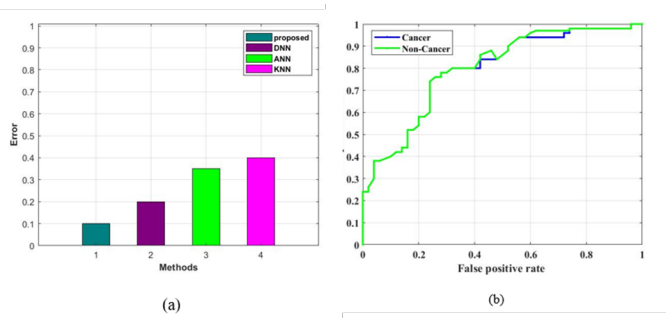

Fig 6. Pictorial plot for (a) Error analysis and (b) ROC of DCNN-LDTM model under distinct TR/TS data

error value in this case less than 0.1, while the high error values of the existing techniques DNN, ANN, and KNN are 0.2, 0.35, and 0.4, respectively. From the above figures, we can understand that compared to the existing approach, the proposed method is better in oral survival rate analysis using oral cancer images. Thus, it demonstrates how well the research architecture supports OS rate that is predicting the survival time of Oral cancer predicting process system. The results consistently showcase the superiority of the proposed DCNN-LDTM model, with error values consistently below 0.1, in contrast to other existing techniques including DNN, ANN, and KNN, which have higher error values of 0.2, 0.35, and 0.4. Furthermore, precision-recall analysis demonstrates the DCNN-LDTM model's exceptional ability to classify oral cancer data into cancer and non-cancer categories. The proposed method stands out in terms of innovation, as with the unique combination of deep learning techniques by the utilization of Local Tetra Pattern (LTrP) for feature extraction, and the application of the HHHLO algorithm for feature selection. This novel approach has not only surpassed previous reports in performance but also advanced the field of oral cancer detection. Our study underscores the significant contributions and innovations that our research brings to the domain of oral cancer diagnosis and prognosis.

3.3 Limitations of proposed methodology

The proposed model is come with its own challenges as well some of them are listed as follows

- **Computational Complexity:** The introduction of LSTM layers in conjunction with DCNN can significantly increase computational complexity. Training and deploying such a model may require substantial computational resources, which could be a limitation for institutions or researchers with limited access to high-performance computing infrastructure.
- **Data Requirements:** Deep learning models, including DCNN-LSTM, often require large amounts of data for effective training. If the dataset available for oral cancer is small or lacks diversity, the model may struggle to generalize to a broader population. Data limitations could impact the model's performance and utility.
- **Hyper parameter Tuning:** The methodology involves numerous hyper parameters, such as the number of layers, filter sizes, and LSTM parameters. Tuning these hyper parameters effectively can be a time-consuming and complex process. Inadequate hyper parameter tuning may result in suboptimal model performance.
- **Interpretability:** Deep learning models, including DCNNs, are often considered difficult to interpret. While the methodology combines DCNN with LSTM, which can model temporal dependencies, understanding how the model makes predictions may still be a challenge. Lack of interpretability can limit the acceptance of the model in clinical practice.
- **Availability of Expertise:** Implementing and fine-tuning a complex model like DCNN-LSTM requires expertise in deep learning and medical imaging. Limitations may arise if the required expertise is not readily available. And for the HHHLO algorithm.
- **Algorithm Complexity:** Complex algorithms like HHHLO can be computationally intensive, requiring significant time and resources to execute. This complexity may limit their practicality for real-time or large-scale applications.
- **Sensitivity to Parameters:** Many optimization algorithms, including hybrid ones, often involve multiple parameters that need to be tuned for optimal performance. Finding the right parameter settings can be a challenging and time-consuming task, and the algorithm's effectiveness may be sensitive to these settings.
- **Scalability Issues:** The algorithm's performance may deteriorate as the problem size or dimensionality increases. Scalability limitations can be a significant challenge when dealing with large datasets or complex optimization problems.
- **Need for Expertise:** Implementing and fine-tuning complex optimization algorithms often requires expertise in the specific algorithm and the problem domain. This can be a limitation for users who lack this expertise.

Taking on these obstacles offers a possible path for improving ongoing research in the sector. By tackling these concerns, researchers can build the framework for optimization algorithms that are more reliable and efficient, hence expanding their applicability to a larger range of issues and enhancing performance. Pushing the limits of optimization algorithms and their useful applications depends on this continual endeavour to get over these restrictions.

4 Conclusion

This study proposes an effective oral cancer prediction method using DCNN-LSTM technique and oral squamous cell carcinoma images. Here, the input image is pre-processed by being cropped and the image contrast level is adjusted using CLAHE. The feature selection of oral squamous cell carcinoma images is done by the HHHLO and the predict the oral cancer using DCNN- LSTM. Due to its enhanced performance, the proposed method can be utilised as a pre-processing tool for feature extraction, feature selection and classification. In an image of the oral squamous cell carcinoma, it may also be utilized to oral cancer prediction. The novelty of the study lies in its holistic approach, which integrates cutting-edge deep learning and sequential processing techniques for oral cancer prediction. In experimental analysis the performance of the proposed approach is analyzed with the existing research approaches. The proposed selection and prediction the proposed DCNN-LSTM is compared with the existing DNN, ANN and KNN approaches. The attained accuracy, sensitivity, specificity, precision and F-measure of our proposed method are 0.97%, 0.94%, 0.97%, 0.89% and 0.94% correspondingly. Thus, the proposed method has demonstrated its superiority to outperform existing techniques based on various performance metrics. Similar to this, the proposed technique achieves superior results based on other performance e analyses of various matrices. This study envisions the potential application of the proposed approach in real-time oral cancer diagnosis using photographs, which could substantially contribute to early detection and enhance patient outcomes.

Acknowledgement

There is no funding from any Research or Funding Agency. The authors would also like to thank the Dept. of Computer Science, Hindustan Institute of Science and Technology, Chennai for the continuous support. The authors would like to thank the reviewers and the editor of the journal for their helpful comments which have greatly improved this paper.

References

1) Das N, Hussain E, Mahanta LB. Automated classification of cells into multiple classes in epithelial tissue of oral squamous cell carcinoma using transfer learning and convolutional neural network. *Neural Networks*. 2020;128:47–60. Available from: https://doi.org/10.1016/j.neunet.2020.05.003.
2) Lakshmanaprabu SK, Mohanty SN, Shankar K, Arunkumar N, Ramirez G. Optimal deep learning model for classification of lung cancer on CT images. *Future Generation Computer Systems*. 2019;92:374–382. Available from: https://doi.org/10.1016/j.future.2018.10.009.
3) Ilhan B, Guneri P, Wilder-Smith P. The contribution of artificial intelligence to reducing the diagnostic delay in oral cancer. *Oral Oncology*. 2021;116:105254. Available from: https://doi.org/10.1016/j.oraloncology.2021.105254.
4) Dharani R. Adaptive Coati Deep Convolutional Neural Network-based Oral Cancer Diagnosis in Histopathological Images for Clinical Applications. 2023. Available from: https://doi.org/10.21203/rs.3.rs-3043490/v1.
5) Dixit S, Kumar A, Srinivasan K. A Current Review of Machine Learning and Deep Learning Models in Oral Cancer Diagnosis: Recent Technologies, Open Challenges, and Future Research Directions. *Diagnostics*. 2023;13(7):1353. Available from: https://doi.org/10.3390/diagnostics13071353.
6) Prabhakaran R, Mohana J. Stochastic gradient descent-based convolutional neural network to detect and classify oral cavity cancer. *Soft Computing*. 2023;27(13):9169–9178. Available from: https://doi.org/10.1007/s00500-023-08283-w.
7) Bansal K, Bathla RK, Kumar Y. Deep transfer learning techniques with hybrid optimization in early prediction and diagnosis of different types of oral cancer. *Soft Computing*. 2022;26(21):11153–11184. Available from: https://doi.org/10.1007/s00500-022-07246-x.
8) Huang Q, Ding H, Razmjooy N. Optimal deep learning neural network using ISSA for diagnosing the oral cancer. *Biomedical Signal Processing and Control*. 2023;84:104749. Available from: https://doi.org/10.1016/j.bspc.2023.104749.
9) Myriam H, Abdelhamid AA, El-Kenawy ESMSM, Ibrahim A, Eid MM, Jamjoom MM, et al. Advanced Meta-Heuristic Algorithm Based on Particle Swarm and Al-Biruni Earth Radius Optimization Methods for Oral Cancer Detection. *IEEE Access*. 2023;11:23681–23700. Available from: https://doi.org/10.1109/ACCESS.2023.3253430.
10) Ding H, Huang Q, Rodriguez D. Modified Locust Swarm optimizer for oral cancer diagnosis. *Biomedical Signal Processing and Control*. 2023;83:104645. Available from: https://doi.org/10.1016/j.bspc.2023.104645.
11) Gao J, Lyu T, Xiong F, Wang J, Ke W, Li Z. MGNN: A Multimodal Graph Neural Network for Predicting the Survival of Cancer Patients. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2020.
12) Soltani R, Jaouadi A. Institute of Electrical and Electronics Engineers (IEEE). . Available from: https://doi.org/10.36227/techrxiv.20408919.v1.
13) Miarnaeimi F, Azizyan G, Rashki M. Horse herd optimization algorithm: A nature-inspired algorithm for high-dimensional optimization problems. *Knowledge-Based Systems*. 2021;213:106711–106711.
14) Yazdani R, Alipour-Vaezi M, Kabirifar K, Kojour AS, Soleimani F. A lion optimization algorithm for an integrating maintenance planning and production scheduling problem with a total absolute deviation of completion times objective. *Soft Computing*. 2022;26(24):13953–13968.
15) Figueroa KC, Song B, Sunny S, Li S, Gurushanth K, Mendonca P, et al. Interpretable deep learning approach for oral cancer classification using guided attention inference network. *Journal of Biomedical Optics*. 2022;27(01):15001. Available from: https://doi.org/10.1117/1.JBO.27.1.015001.

16) Saber A, Sakr M, Abo-Seida OM, Keshk A, Chen H. A Novel Deep-Learning Model for Automatic Detection and Classification of Breast Cancer Using the Transfer-Learning Technique. *IEEE Access*. 2021;9:71194–71209. Available from: https://doi.org/10.1109/ACCESS.2021.3079204.