Proceedings of the Fourth International Conference on Electronics and Sustainable Communication Systems (ICESC-2023)
IEEE Xplore Part Number: CFP23V66-ART; ISBN: 979-8-3503-0009-3

2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC) | 979-8-3503-0009-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICESC57686.2023.10193479

Analysis of Deep Learning based Optimization Techniques for Oral Cancer Detection

K J Subha[1]
Department of ECE
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
Chennai, India
subhakj@veltech.edu.in

M.Anto Bennet[2]
Department of ECE
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
Chennai, India
drmantobenet@veltech.edu.in

Gaddam Pranay[3]
Department of ECE
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
Chennai, India
gaddampranay508@gmail.com

Ketham Bharadwaj[4]
Department of ECE
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
Chennai, India
kethambharadwaj@gmail.com

Polu Vikram Reddy[5]
Department of ECE
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
Chennai, India
poluvikramkumareddy@gmail.com

***Abstract*— The main aim is to examine the application of deep learning technique in detecting oral cancer at an early stage. The focus is on evaluating the performance of various optimization techniques to deep learning technique through analyzing images of the mouth and throat for oral cancer detection. The goal is to determine the most effective method for identifying oral cancer and gain insights into potential improvements. The importance of early detection of oral cancer cannot be overstated, as it plays a crucial role in improving patient outcomes through earlier treatment and higher survival rates. The findings have the potential to contribute to the development of more accurate and efficient methods for oral cancer diagnosis and make a positive impact in the field.**

Keywords*— *Deep Learning, Oral Cancer Detection, Optimization Techniques, Convolutional Neural Network, Transform Techniques.

I. INTRODUCTION

Oral cancer named Oral Squamous Cell Carcinoma (OSCC) occurrence poses a significant concern for public health with millions of cases diagnosed each year around the world. Early detection is key to improving survival, as the disease is highly treatable in its early stages. Deep learning demonstrated significant potential in the field of medical image analysis, with several studies demonstrating its potential for accurate and efficient detection of various diseases, including cancer. Therefore, the main aim is to analyze deep learning-based optimization techniques for oral cancer detection [1]. The project will investigate application of Transform Techniques to deep learning model Convolutional Neural Network (CNN), to analyze their effectiveness in detecting oral cancer.

Additionally, the study compare and contrast the results of different optimization techniques to determine the most efficient way to optimize a deep learning model for oral cancer detection [2]. Oral cancer affects the mouth and throat, with an approximate estimation of 377,000 novel cases and 177,000 deaths around the world in 2020, in accordance with World Health Organization (WHO). Early detection is key to improving survival, as the disease is highly curable in its early stages. Therefore, there is a growing need for accurate and effective methods for early detection and diagnosis of oral cancer. As a part of AI, deep learning has shown significant potential in the field of medical images analysis [3]. Deep learning models can autonomously learn and extract features from medical images, plays a major role in detecting diseases accurately, including cancer. In recent years, there has been increase in interest in detecting Oral cancer using deep learning. Detecting Oral cancer using deep learning involves training a deep neural network on a huge dataset of oral cancer images. A neural network is designed to learn from image patterns and identify features indicative of oral cancer. After training, the model can be used to detect oral cancer in new images with high accuracy.

Oral cancer detection using deep learning has several pros. First, deep learning models can analyze huge amounts of data fast and accurately. This helps reduce the workload of healthcare professionals and the diagnosis of oral cancer will be done quickly. Second, deep learning models can learn from a variety of image types, including X-rays, CT scans, and MRIs, making them versatile tools for detecting oral cancer. Finally, deep learning models can improve diagnostic accuracy, reduce the risk of misdiagnosis, and improve patient outcomes [4].

Notwithstanding the possible advantages of deep learning for oral cancer detection, there are some challenges that need to be considered. One of the biggest challenge is to train high quality data for the models. Another challenge is that expertise in deep learning and oral cancer diagnostics is needed to develop and implement effective models. CNNs play an important role in various tasks such as image processing problems, some of the computer vision assignments are localizing and segmenting images, analyzing videos, detecting obstacles in autonomous vehicles, and recognizing speech in natural language processing. [5]. Since

CNNs play an important role in these emerging and rapidly developing fields, it is very popular in deep learning. In deep learning, convolutional neural networks are a class of deep neural networks most commonly used for image classification. Now, when the term neural networks arise, matrix multiplication will also arise, but that's not the case with ConvNets. It uses special technique called convolution [6]. As per the above-mentioned details the purpose for this study is to analyze the best model for oral cancer detection from the models obtained with the combination of Convolution Neural Network (CNN) and Optimization Techniques.

The remaining of this stduy proceeds in the following order section two covered with Oral cancer dataset required for this work, section three contains Methodology and Block diagram and Section four with Results and Section five concluded with the results obtained.

II. ORAL CANCER HISTOPATHOLOGICAL IMAGE DATASETS

This section provides concise overview of Oral Cancer Histopathological Image Dataset, the term Histopathological Image defines the study of signs of disease by microscopic examination of biopsies or surgical specimens, that are processed and mounted on glass slides. There was a total of 5192 image files in the dataset and the data is been classified into two classes Normal and OSCC. The Normal class contains non-cancer images while the other OSCC class contains the microscopic images of cancer effected tissues. The Data is also been splitted into three parts, they are train set, test set and Validation set. The Train set consists of 4946 images in total and testing set contains 126 images while the validation set consists of 120 images in total. These three parts are used in model according to the requirement in training process.

III. PROPOSED MODEL

This section provides the methodology used to pre-process and train Oral cancer Histopathological images with Transform Techniques, Optimization Techniques, and a CNN model. Transform techniques and Optimization techniques are used for Feature extraction in this. So, for classification of the model the CNN is used. The objective is to classify the images into two classes: Cancer- affected and normal. The proposed method is motivated by the fact that although the basic CNN model has been implemented in previous studies for oral cancer prediction, the inclusion of Transform Techniques and Optimization techniques with the CNN model has been found to significantly improve the performance of the absolute model. Therefore, we aim to leverage these techniques to enhance the accuracy of our classification model [7]. The model is implemented in different stages namely Pre-processing phase, Feature Extraction phase, and the Classification phase.

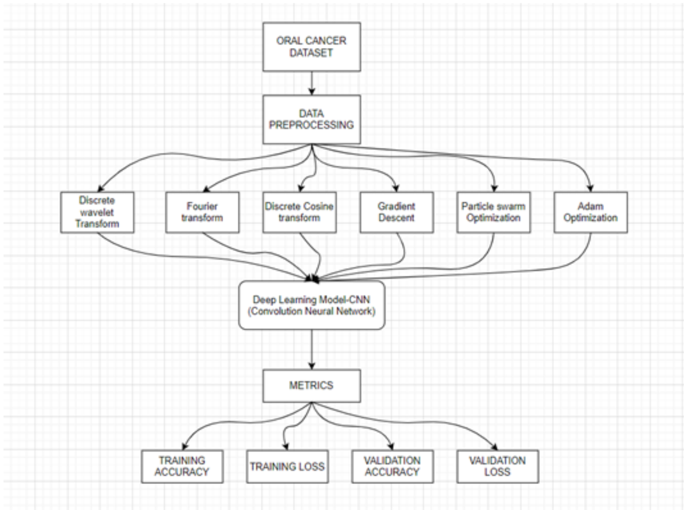

Fig. 1. Block Diagram of the Model

A. Pre-processing phase

In this Initial stage the Data has been undergone re-scaling, the pixel values of every image present in the dataset has been set to same values. The pre-process also involves the elimination of noise and enhance the quality of images. After re-scaling all the images are converted into Gray scale images in case of application of Fourier Transform Technique while the other techniques does not require of Gray scale images. The main features for these images in Oral cancer dataset are Cell Boundary and Cell size. These features will be extracted in the next stage known as Feature Extraction stage.

The first step in pre-processing the oral cancer image dataset is data cleaning. This involves deleting any corrupted or incorrect images, such as missing pixels or mislabelled images. Data cleaning helps ensure the accuracy and completeness of the dataset, which is essential for obtaining accurate machine learning results. After cleaning the data, the next step is to resize the image. Image scaling is the process of reducing or increasing the size of images in a dataset to a specific size, the pixel values of every image present in the dataset has been set to values [224,224]. This step is necessary because the images in the dataset can have different sizes and aspect ratios, which can affect the performance of the machine learning algorithm. Image scaling ensures that all images are the same size and can be easily handled by machine learning algorithms. Image normalization is another important step in the pre-processing of oral cancer image datasets. Image Normalization comprises modifying the pixel values of an image to fit within a particular range, usually between 0 and 1 or -1 and 1. This step is necessary because the pixel values in the image may have different ranges, which may affect the performance of the machine learning algorithm. Normalizing pixel values ensures that they are consistent and can be easily processed by machine learning algorithms. For, the dataset we selected does not contain any noise and the images are of good quality after rescaling, so elimination of noise and enhancement of quality of image procedures are not required

B. Feature extraction phase

Feature extraction plays a crucial role in processing the Oral cancer image dataset for use in machine learning algorithms. The features in this dataset are mainly Cell boundary and cell size, the colour is not included as a feature in this work. The goal of feature extraction is to extract relevant information from the images in the dataset that can be used to train ma- chine learning models. There are several methods for feature extraction, including hand-crafted features and deep learning- based features. The techniques used for Feature extraction here are Transform Techniques and Optimization Techniques.

1) Transform techniques:

Feature extraction using transform techniques is a crucial step in image processing and analysis, especially for medical image datasets such as oral cancer images. This process involves the identification and extraction of significant features from images to facilitate further analysis, classification and diagnosis. The transform Techniques applied for this model are Fourier Transform, Wavelet Transform and Discrete Cosine Transform. Transform techniques can be applied to the image dataset to extract high-frequency components that correspond to edges and boundaries. By setting an appropriate threshold on the wavelet coefficients, edges can be detected and isolated. Then, techniques such as contour extraction or morphological operations can be applied to identify cell boundaries based on the detected edges.

a) Fourier Transform:

One of the most commonly used transform techniques for feature extraction is the Fourier transform. This technique is used to convert an image from the spatial domain to the frequency domain, where the image is represented as a sum of sinusoidal components with different frequencies and amplitudes [8], [9]. The resulting spectrum can reveal important features such as texture, shape, and contrast in the image. The mathematical representation of Fourier transform is

$$f(t) = \int_{\infty}^{-\infty} F(w) e^{-iwt} dt \tag{1}$$

b) Discrete Wavelet Transform:

Another commonly used transform technique is wavelet trans- form. This technique is used to decompose an image into different levels of frequency and scale, and the resulting coefficients represent the image features at different levels of detail [10]. Wavelet transform can be used to spot the features such as boundaries, textures, and patterns in the image. The mathematical representation of Discrete wavelet Transform is

$$T_{mn} = \int_{\infty}^{-\infty} x(t) \psi mn^{t} dt \tag{2}$$

The figure 2 shows the difference between Images from dataset before pre-processing and after applying Transform Techniques.

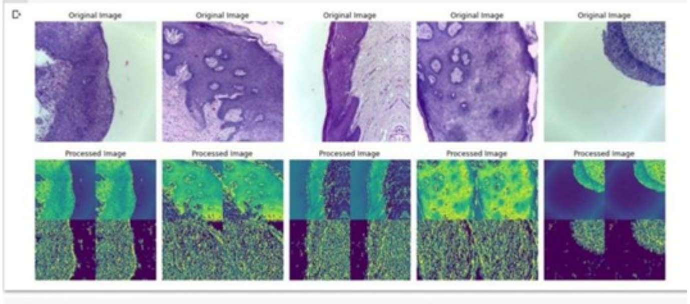

Fig. 2. Images from Dataset after applying Transform Techniques

c) Discrete Wavelet Transform:

The DCT is similar to the Fourier Transform, but it uses a smaller set of basis functions (cosine functions) to represent the image. The resulting coefficients represent the image features at different levels of frequency and scale. The DCT can be applied to the entire image or to smaller regions of interest, such as tumor regions, to extract features that are specific to those regions. the advantages of using the DCT for feature extraction is its computational efficiency, as it requires less computation time than some other transform techniques. Additionally, the DCT is a lossy compression technique, meaning that it can be used to compress the image data while still retaining the important features [11]. This can be useful for reducing storage and transmission requirements in large image datasets.

2) Optimization Techniques:

Feature extraction using optimization techniques is another approach for processing the Oral cancer image dataset for use in machine learning algorithms. This approach involves using optimization algorithms to find a set of features that can best discriminate between the different classes of images in the dataset [12]. During training, SGD is used to optimize the model's parameters, including the weights and biases of the neural network, to minimize the detection loss function. Once the model is trained, it can be used to extract features related to cell size and cell boundary. The Optimization techniques we used here for Feature extraction are as follows

a) Gradient Descent:

This method iteratively adjusts the values of a set of parameters in a model to minimize the error between predicted outputs and actual outputs. For feature extraction, gradient descent can be used to identify the most relevant features that can accurately predict the presence or absence of oral cancer in an image. The features are represented as numerical values and are used as inputs to the model, which then outputs a prediction based on those features.

One of the advantages of using gradient descent for feature extraction is that it can automatically select the most relevant features without requiring manual feature selection. However, it can be computationally expensive and may require a large amount of data to accurately identify the optimal set of features. Additionally, selecting appropriate hyper-parameters such as the learning rate and number of iterations can be challenging [13]. Overall, gradient descent is a powerful tool for feature extraction in oral cancer image datasets and can be used with a range of machine learning models [14]. Proper data preprocessing, quality control, and validation are essential to ensure accurate and reliable results.

b) Adam Optimization:

One of the most popular optimization techniques for feature extraction is Adam, which is short for Adaptive Moment Estimation. Adam optimization technique is a gradient-based optimization algorithm that combines ideas from both Momentum and RMSprop methods. It calculates adaptive learning rates for each parameter, and exponentially decays the average of past gradients to ensure that the optimization process is efficient and accurate. This step involves extracting relevant features from the pre-processed images using Adam optimization technique [15]. The algorithm uses a neural network with multiple layers to learn the features and adjust the weights of the network. The extracted features can be used for further analysis, such as classification or clustering. The algorithm can be further optimized by fine-tuning the hyper-parameters or using other optimization techniques such as stochastic gradient descent (SGD).

c) Stochastic Gradient Descent:

Stochastic Gradient Descent (SGD) is known as an optimization technique that can be used for feature extraction in oral cancer image dataset analysis. SGD works by updating the parameters of a model based on the gradient of the loss function. In oral cancer image analysis, the loss function measures the error between the predicted and actual values of a target variable, such as the existence or non-existence of cancer cells. By randomly selecting a small batch of samples from the dataset, SGD can efficiently extract features that are most relevant for accurate diagnosis and treatment of oral cancer [16]. The model is trained to identify the subset of features that are most important for accurately predicting the presence or absence of cancer cells in the biopsy samples. Using SGD for feature extraction in oral cancer image dataset analysis can improve the accuracy of oral cancer diagnosis and treatment, ultimately leading to better patient outcomes. However, it should be used in conjunction with other diagnostic tools and techniques and under the guidance of a qualified medical professional.

d) RMSprop Optimization:

RMSprop is another optimization technique that can be used for feature extraction in oral cancer image dataset analysis. RMSprop adjusts the learning rate of the model based on the root mean square of the gradients. This methodology can be beneficial to prevent the model from getting stalled in a local minimum during optimization. In oral cancer image analysis, RMSprop can efficiently extract features that are most relevant for accurate diagnosis and treatment of oral cancer. The model is trained to identify the subset of features that are most important for accurately predicting the presence or absence of cancer cells in the biopsy samples. Using RMSprop for feature extraction in oral cancer image dataset analysis can improve the accuracy of oral cancer diagnosis and treatment, ultimately leading to better patient outcomes. However, it should be used in conjunction with other diagnostic tools and techniques and under the guidance of a qualified medical professional. The accurate diagnosis and treatment of oral cancer require a comprehensive approach that includes imaging analysis, biopsy, and other diagnostic tests.

e) Adadelta Optimization:

Adadelta is an optimization technique that can be used for feature extraction in oral cancer image dataset analysis. It is an extension of the Adagrad algorithm that aims to address the problem of diminishing learning rates. Adadelta uses a moving window of the past gradients to dynamically modify the model's learning rate. In oral cancer image analysis, Adadelta can efficiently extract features that are most relevant for accurate diagnosis and treatment of oral cancer. The model is trained to identify the subset of features that are most important for accurately predicting the presence or absence of cancer cells in the biopsy samples. One advantage of using Adadelta for feature extraction in oral cancer image dataset analysis is that it requires very little memory compared to other optimization techniques [17].

This makes it a useful technique for large datasets with limited memory resources. However, it is important to note that Adadelta may require longer training times compared to other optimization techniques, and the choice of optimization technique Ought to be contingent on particular requirements and characteristics of the oral cancer image dataset being analyzed. As with any diagnostic tool, the use of Adadelta for feature extraction should be under the guidance of a qualified medical professional.

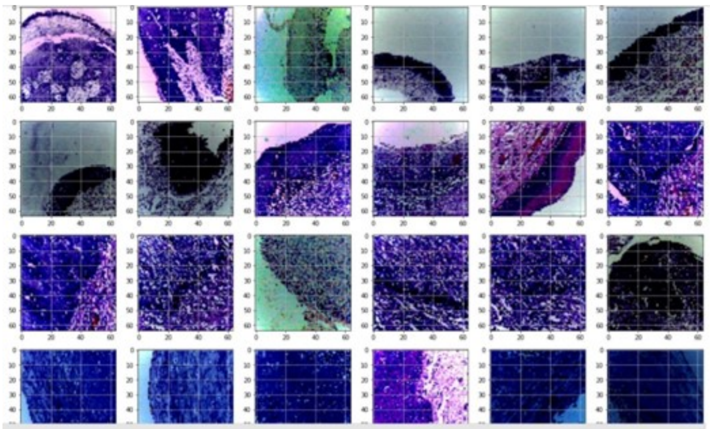

Fig. 3. Images from Dataset after Optimization

The figure 3 shows the images of data after applying optimization in process of training the model. After applying feature extraction using optimization techniques, we need to select the optimal set of features that can best discriminate between different classes of images.

C. Classification phase

CNN (Convolutional Neural Network) is the main algorithm we used here for classification, but some layers has been changed or applied for the CNN network to gain more precise results. The main reason for using CNN in this work is, CNN provides better accuracy for image processing among all the binary classifiers and making changes in CNN network will give us the best results. The CNN architecture typically involves several layers, including convolutional layers, pooling layers, activation layers, and fully connected layers. The convolutional layers extract features from the input images, while the pooling layers down sample the feature maps to reduce the dimensionality of the data, the pooling can be done by various types such as Max Pool (gives the maximum values from the obtained features as input to next layer), Average Pool (gives the Average mean value of obtained features as the input to next layer) [18].

Activation layers introduce non-linearity into the network, allowing it to model complex relationships, there are also number of activation functions can be used in Convolutional Neural Network such as soft-max function, Rectified Linear unit (ReLU) function. In this CNN model ReLU is used as activation function in Neural network. Finally, fully connected layers are used to classify the input images. During the training process, the CNN learns to identify the patterns and features that are important for oral cancer detection. This is done by adjusting the weights of the network using back-propagation algorithm [19]. Once the CNN is trained, it can be evaluated using a validation set or a testing set of images. Performance of the trained model is typically measured using performance metrics such as accuracy, precision and recall.

IV. RESULTS AND ANALYSIS

This section presents the results of the models and will be analysed, the models have been evaluated based on various performance metrics such as training accuracy, loss, validation accuracy, validation loss, precision, and recall.

TABLE I. PERFORMNACE METRICS

MODELS TRAINED	METRICS					
	$T_{ACCURACY}$	*T_{LOSS}*	*$V_{ACCURACY}$*	*V_{LOSS}*	*Precision*	*Recall*
CNN	0.5073	0.6934	0.246	0.6962	0.754	1.0
DWT+CNN	0.5077	0.6931	0.754	0.6858	0.753	1.0
Fourier Transform +CNN	0.5077	0.6931	0.754	0.685	0.75	1.0
Adam+CNN	0.8142	0.4276	0.858	0.37	0.8737	0.8411
SGD+CNN	0.6933	0.3077	0.787	0.47	0.8737	0.884
Gradient Descent +CNN	0.6759	0.6173	0.7778	0.512	0.75	0.91
Adadelta+ CNN	0.8229	0.3856	0.7857	0.4038	0.89	0.82
RMSprop+ CNN	0.4932	0.502	0.754	0.4851	0.75	0.91

Fig. 4. Performance Metrics of the model.

Firstly the basic CNN model has a low training accuracy of 50.73% and a high loss of 0.6934. Its validation accuracy is only 24.6%, which indicates poor generalization, and its validation loss is also high at 0.6962. The next model is the Discrete wavelet Transform for CNN (Convolution Neural Network), this model has the same training accuracy and loss as the previous model, but its validation accuracy has improved to 75.4%. However, its validation loss is still high at 0.6858. The precision and recall values are reported as same for previous model. The Next model is Adam Optimization technique for CNN, this model has the highest training accuracy of 81.42% and the lowest loss of 0.4276. It also has a high validation accuracy of 85.8% and a low validation loss of 0.37. Its precision and recall values are also reported and are reasonably good at 0.8737 and 0.8411, respectively. The next model is Stochastic Gradient Descent which is having a bit lower results compared to Adam optimizer other than precision. The next model is Gradient Descent Optimization technique for CNN, this model has a lower training accuracy of 67.59% and a higher loss of 0.6173. Its validation accuracy is 77.78%, and its validation loss is even higher at 0.512. Its precision and recall values are reported and are reasonable at 0.75 and 0.91, respectively. The other model is AdaDelta Optimization technique for CNN, this model has a high training accuracy of 82.29% and a low loss of 0.3856. Its validation accuracy is 78.57%, and its validation loss is even lower than the previous models at 0.4038. Its precision and recall values are reported and are very good at 0.89 and 0.82, respectively, and the final model RMSprop is having lowest training accuracy and high training loss compared to previous models.

CONCLUSION

Oral cancer detection using histopathological image dataset can be good supporting in medical imaging, so, till now the accuracy for oral cancer detection is not more than 76% in previous records, in this work there has been an effort to increase the Accuracy by applying optimization techniques. Based on the metrics data acquired from training, the model that appears to have the best performance metrics is the one trained with the Adam optimizer. This model has a higher training accuracy (0.8142) and a lower training loss (0.4276) than all other models listed. Additionally, it has a high validation accuracy (0.858) and a low validation loss (0.37), indicating that it does not overfit to the train data and is able to generalize well to new data. Each specific dataset has its own feature, so, as for this oral cancer detection using this oral cancer image dataset, the best model is training with CNN using Adam optimization technique, even though Adadelta is having more training accuracy, it is having less validation accuracy and more validation loss compared to Adam optimization applied model. By comparing metrics of all the models trained, model trained with Adam+CNN has given better results with better accuracy and less loss. So, the suggested model for Oral cancer detection is model trained using Adam and CNN classification.

REFERENCES

[1] Lin, Huiping, et al. "Automatic detection of oral cancer in smartphone- based images using deep learning for early diagnosis." Journal of Biomedical Optics 26.8 (2021): 086007-086007.

[2] Alabi, Rasheed Omobolaji, et al. "Deep machine learning for oral cancer: from precise diagnosis to precision medicine." Frontiers in Oral Health 2 (2022): 97.

[3] Welikala, R.A., Remagnino, P., Lim, J.H., Chan, C.S., Rajendran, S., Kallarakkal, T.G., Zain, R.B., Jayasinghe, R.D., Rimal, J., Kerr, A.R. and Amtha, R., 2020. "Automated detection and classification of oral lesions using deep learning for early detection of oral cancer". IEEE Access, 8, pp.132677-132693.

[4] Nanditha, B. R., A. Geetha, H. S. Chandrashekar, M. S. Dinesh, and S. Murali. "An ensemble deep neural network approach for oral cancer screening." (2021): 121-134.

[5] Ariji, Y., Kise, Y., Fukuda, M., Kuwada, C., Ariji, E. (2022). Segmentation of metastatic cervical lymph nodes from CT images of oral cancers using deep-learning technology. Dentomaxillofacial Radiology, 51(4), 20210515.

[6] Warin, K., Limprasert, W., Suebnukarn, S., Jinaporntham, S., Jantana,

P. (2021). Automatic classification and detection of oral cancer in photographic images using deep learning algorithms. Journal of Oral Pathology Medicine, 50(9), 911-918.

[7] Bansal, Khushboo, R. K. Bathla, and Yogesh Kumar. ”Deep transfer learning techniques with hybrid optimization in early prediction and diagnosis of different types of oral cancer.” Soft Computing 26.21 (2022): 11153-11184.

[8] Mironovova, Martina, and Jir´ı B´ıla. ”Fast fourier transform for feature extraction and neural network for classification of electrocardiogram signals.” 2015 Fourth International Conference on Future Generation Communication Technology (FGCT). IEEE, 2015.

[9] Wu, Hao, Xue Ma, and Chenglin Wen. ”Multilevel fine fault diagnosis method for motors based on feature extraction of fractional fourier transform.” Sensors 22.4 (2022): 1310.

[10] Saxena, S. C., Vinod Kumar, and S. T. Hamde. ”Feature extraction from ECG signals using wavelet transforms for disease diagnostics.” International Journal of Systems Science 33.13 (2002): 1073-1085.

[11] Dabbaghchian, Saeed, Masoumeh P. Ghaemmaghami, and Ali Aghagolzadeh. ”Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology.” Pattern recognition 43.4 (2010): 1431-1440.

[12] Myriam, Hadjouni, et al. ”Advanced Meta-Heuristic Algorithm Based on Particle Swarm and Al-Biruni Earth Radius Optimization Methods for Oral Cancer Detection.” IEEE Access (2023).

[13] Aydog˘du, O¨ zge, and Murat Ekinci. ”An approach for streaming data feature extraction based on discrete cosine transform and particle swarm optimization.” Symmetry 12.2 (2020): 299.

[14] Haji, Saad Hikmat, and Adnan Mohsin Abdulazeez. ”Comparison of optimization techniques based on gradient descent algorithm: A review.” PalArch’s Journal of Archaeology of Egypt/Egyptology 18.4 (2021): 2715-2743.

[15] Suresha, Halebeedu Subbaraya, and Srirangapatna Sampathkumaran Parthasarathy. ”Alzheimer disease detection based on deep neural network with rectified Adam optimization technique using MRI analysis.” 2020 Third International Conference on Advances in Electronics, Com- puters and Communications (ICAECC). IEEE, 2020.

[16] Zheng, Qinghe, et al. ”Layer-wise learning based stochastic gradient descent method for the optimization of deep convolutional neural network.” Journal of Intelligent Fuzzy Systems 37.4 (2019): 5641-5654.

[17] Kokilambal, S. ”Intelligent content based image retrieval model using adadelta optimized residual network.” 2021 International Conference on System, Computation, Automation and Networking (ICSCAN). IEEE, 2021.

[18] Bazi, Yakoub, et al. ”Simple yet effective fine-tuning of deep CNNs using an auxiliary classification loss for remote sensing scene classification.” Remote Sensing 11.24 (2019): 2908.

[19] Das, N.; Hussain, E.; Mahanta, L.B. ”Automated classification of cells into multiple classes in epithelial tissue of oral squamous cell carcinoma using transfer learning and convolutional neural network”. Neural Netw. 2020, 128, 47–60.