Research Article GU J Sci 38(2): 1006-1019 (2025) DOI: 10.35378/gujs.1480477

Gazi University

Journal of Science

http://dergipark.gov.tr/gujs

Examination of ResNet and DenseNet Architectures in Early Diagnosis of Oral Cancer: An Evaluation

Betul SUREN , Mutlu AKAR*

Yildiz Technical University, College of Arts and Sciences, Department of Mathematics, Davutpasa Campus, 34220, Esenler, Istanbul, Türkiye

Highlights

- This paper focuses on the classification of lesion images for early diagnosis of oral cancer.
- A classification was made using convolutional neural networks in the study.
- A highly precise and efficient classification accuracy were obtained.

Article Info

Received: 08 May 2024
Accepted: 04 Jan 2025

Keywords

Oral cancer
Deep learning
ResNet
DenseNet

Abstract

Oral cancer holds a significant position among head and neck cancers and is encountered quite frequently. Oral cancer, which is the eleventh most common type of cancer worldwide, causes approximately 177,000 deaths and 350,000 new cases every year [1, 2]. The most commonly observed type of oral cancer is Oral Squamous Cell Carcinoma (OSCC) [6] which comprises about 90% of the cases [7]. The survival rate for OSCC is low due to the frequent late diagnosis [19]. This also underscores the importance of early diagnosis. Convolutional neural networks (CNN) are highly preferred for their high performance in early diagnosis. In this study, the early diagnosis of oral cancer has been investigated through the utilization of CNN. Additionally, two models are selected for each of the two different CNN architectures. Classification is carried out with varying hyperparameters on these four models, and the resulting classification accuracies were examined. Furthermore, the two architectures are compared in terms of their performance, highlighting the differences in accuracy and efficiency. The accuracy values for the DenseNet and ResNet architectures in this classification problem are investigated. Models were selected with varying layer depths within each architecture to understand how the number of layers affected classification accuracy. Furthermore, these processes are carried out with different optimizers and epoch numbers, aiming to explore the influence of optimizer choices and varying epoch numbers on classification accuracy. As a result of the study, the highest accuracy rate was measured as 97.01%, achieved using the DenseNet201 architecture with the SGD optimizer.

1. INTRODUCTION

Oral cancer held a significant position among head and neck cancers, with a notably high incidence. Oral cancer caused approximately 177,000 deaths and 350,000 new cases, and it was ranked as the eleventh most common type of cancer worldwide every year [1, 2]. The significance of oral cancer was presented. Two-thirds of the global incidence of oral cancer occurred in low or middle-income countries [3]. Furthermore, it ranked sixth among high risk types of cancer incidence globally and in middle-income countries [4]. The oral cavity and all sub-regions, lip, and oropharynx cancers were involved by this cancer type [5].

Oral Squamous Cell Carcinoma (OSCC), which was commonly observed, was identified as the prevalent type of oral cancer [6], composing about 90% of the disease [7].

Mucosal carcinomas, including oral squamous cell carcinoma, were evaluated separately due to differences in position, anamnesis, etiology, and methodology. Oral cancer was usually evaluated differently from carcinomas that occurred in the oropharynx because oropharyngeal cancers were mostly resulted from the human papillomavirus (HPV), while oral cancer was typically caused by tobacco and alcohol use [8].

*Corresponding author, e-mail: makar@yildiz.edu.tr

But the incidence of oral and oropharyngeal cancer increased due to the HPV origin in the last ten years [9-11]. Additionally, it can be stated that oral cancer risk factors were altered based on the geographical region and the lifestyle habits of the local inhabitants [12]. For example, among the foremost risk factors in Western countries, smoking and alcohol use were considered [12], while in South Asia and Pacific countries, chewing betel nuts and smoking were identified as important risk factors [13]. So, to sum up, the major risk factors for this disease included the use of all tobacco products, alcohol consumption, chronic inflammation, betel nut chewing, and HPV [14-18].

The regions where oral cancer was typically observed included the lips, floor of the mouth, gums, and tongue, the oral cavity, and various other regions [7].

The survival rate for OSCC is typically kept low because the diagnosis is often delayed [19]. The five-year survival rate for OSCC generally ranged from 15% to 50%, with most diagnoses being made at an advanced stage, typically in the third or fourth stage [20]. However, if the diagnosis was made in the early stages, particularly in the first and second stages, the survival rate was believed to have reached around 80%. [21, 22].

The disease could be diagnosed at an advanced stage of approximately 50%. The reasons for this are believed to be the lack of significant symptoms in patients during the early stages, or the patient not seeking medical support until experiencing obvious symptoms such as pain, bleeding, and a mass in the oral cavity or neck, and the belief that lymphatic spread occurs [22].

When there was a delay in diagnosis for more than a month, the probability of the disease being in an advanced stage increased [23]. Also, when the disease advanced and access to the location where the tumor was located became difficult, the prognosis worsened [9]. Biomarkers could be used in the diagnosis of OSCC [24].

For this, a biopsy specimen is taken from the patient by the pathologist and examined under the microscope. However, this evaluation was conducted solely through the visual examination of cell structure, formality, tissue distribution, and cancer level [19].

Therefore, the results are considered qualitative, and specificity with sensitivity is considered uncertain. For both this reason and due to increased awareness about the dangers of oral cancer, the demand for true and early diagnosis techniques increased [25].

Many patients avoided the biopsy because of reasons such as cost or various other factors. Therefore, in the early stages, patients often delayed or avoided biopsy, hindering early diagnosis. For all these reasons, researchers recommended more accessible diagnostic methods [26].

One of the mentioned diagnostic approaches entailed diagnosing oral cancer by examining risk factors, conducting laboratory tests, and analyzing lesion images through the utilization of artificial neural networks (ANN) [27]. In this study, the goal was to achieve successful results using the powerful architectures of artificial neural networks for a disease where early diagnosis is of vital importance. In this context, a comprehensive analysis was carried out using models and hyperparameters that were believed to provide the highest accuracy. The different models and hyperparameters used in the study also serve as a guiding resource for future research.

The studies conducted on the subject were mentioned in the continuation of this section.

A smartphone-based neural network model was suggested by Uthoff et al. (2018) [28], and sensitivity, specificity, positive estimator, and negative estimator values ranging from 80% to 90% were obtained. In another study, it was demonstrated by Xu et al. (2019) [29]. that three-dimensional CNN are superior to two-dimensional CNN in the early diagnosis of oral cancer.

High performance was demonstrated by CNN, making them highly preferred. A deep learning-based multilayer neural network model was proposed by Gupta et al. (2019) [30], achieving an accuracy of 91.65% for the training dataset and 89.3% for the test dataset. A divided deep CNN model has been proposed by Jeyaraj et al. (2019) [31], and a classification accuracy of 91.4%, sensitivity of 0.94, and specificity of 0.91 were achieved for a 100-image dataset. Additionally, an accuracy of 94.5% was achieved for a 500-image dataset.

A study was presented by Fu et al. (2020) [32] in which they achieved an AUC (Area Under the Curve) of 98.3%, a sensitivity of 94.9%, and a specificity of 88.7%. The ResNet-101 algorithm was utilized in a study presented by Welikala et al. (2020) [33], and a training accuracy of 78.30% was achieved.

The utilization of AlexNet, VGG-16, VGG-19, and ResNet-50 models was presented in a study by Das et al. (2020) [34], and the highest classification accuracy among these four models was achieved with the ResNet-50 model at 92.15%. Additionally, Das et al. introduced their self-developed CNN model, and better results were obtained with 97.5% accuracy compared to the transfer learning performed with the four models that had been previously used. The SVM model was utilized in a study by Chu et al. (2020) [35], and a training accuracy of 70.59% was achieved. A study was presented by Lin et al. (2021) [36], and data preprocessing was applied to the smartphone-based images. The study utilized the CNN-based HRnet algorithm, and sensitivities of 83.0%, specificity of 96.6%, precision of 84.3%, and an F1-score of 83.6% were achieved.

The artificial neural network-based prediction models were developed by Alhazmi et al. (2021) [37], and twenty-nine variables were addressed for each case in the data set. These twenty-nine variables associated with patients were used to train the model. Sensitivity of 85.71%, specificity of 60%, and accuracy of 78.95% were obtained.

A study utilizing the FCM (Ex vivo fluorescent confocal microscopy) method was presented by Shavlokhova et al. (2021) [38], and the dataset was subjected to data preprocessing to prevent overfitting. MobileNet was employed as a deep learning model, and achievements included a sensitivity of 47% and a specificity of 96%. DenseNet-121 and R-CNN models were used by Warin et al. (2021) [39], and an accuracy of 99%, a sensitivity of 100%, and an F1-score of 99% were obtained for DenseNet-121. For R-CNN, a sensitivity of 76.67%, a recall of 82.12%, and an F1-score of 79.31% were obtained.

EfficientNet-B0 was utilized by Jubair et al. (2021) [40], and an accuracy of 85%, a specificity of 84.5%, a sensitivity of 86.7%, and an AUC of 92.8% were achieved. The advantages of the ResNet-50 and VGG-16 models were combined in a deep learning model developed by Naditha B R et al. (2021) [41], and an accuracy of 96.2%, a sensitivity of 98.14%, and a specificity of 94.23% was achieved.

In the study presented by Gizem et al. (2021) [27], EfficientNet-b4, Inception-v4, DenseNet161, Ensemble, and ResNet152 models were utilized. For EfficientNet-b4, a precision of 0.869, recall of 0.855, and F1-score of 0.858 were achieved. For Inception-v4, a precision of 0.877, recall of 0.855, and F1-score of 0.858 were obtained. For DenseNet161, a precision of 0.879, recall of 0.841, and F1-score of 0.844 were achieved. For Ensemble, a precision of 0.849, recall of 0.841, and F1-score of 0.843 were achieved. Finally, for ResNet152, a precision of 0.826, recall of 0.812, and F1-score of 0.811 were achieved.

A study was presented by Alkahadar et al. (2021) [42], and the decision tree model was employed. An accuracy of 76% was achieved. A transfer learning-supported model was proposed by Rahman et al. (2022) [43], and an algorithm based on AlexNet was utilized. As a result of the study, a training accuracy of 90.06% was achieved. A study was presented by Deif et al. (2022) [44]. In the study, VGG-16, AlexNet, ResNet50, and Inception-V3 were utilized for feature extraction from the dataset, and binary particle swarm optimization (BPSO) was employed to select the best features. Classification was performed using XGBoost. The highest classification accuracy of 96.3% was achieved when Inception-V3 and BPSO were utilized.

The MobileNet-V3 and EfficientNet-V2 models were used by Liyanage et al. (2023) [45]. For EfficientNet-V2, a recall of 64%, precision of 61%, and an F1-score of 62% were achieved. For MobileNet-V3, a recall, precision, and F1-score of 64% were obtained. In the study presented by Zhou et al. (2023) [46], the ResNet50 model was utilized for the classification task, resulting in a precision of 92.86%, recall of 91.84%, F1-score of 92.24%, specificity of 96.41%, sensitivity of 91.86%, and an AUC value of 98.95%. For the detection task, the YOLOV5 model was employed, resulting in a sensitivity of 98.70%, a precision of 98.70%, recall of 79.51%, F1-score of 88.07%, and an AUC value of 90.89%.

A CNN model (CLAHE + GLCM + ICNN) was proposed by Manikandan et al. (2023) [47], and an accuracy of 97.32% was achieved.

Furthermore, different machine learning methods were preferred for the detection of various diseases and types of cancer where early diagnosis is crucial, and successful results were achieved [48-50].

The primary aim of this study is to achieve rapid and reliable results in oral cancer diagnosis using transfer learning. Additionally, efforts are being made to identify the most suitable hyperparameters that align with the objectives of the study, while emphasizing the impact of hyperparameters on the utilized architectures. The contributions of this study are as follows.

- We train CNN models on our dataset, which is believed to yield effective results for early diagnosis of oral cancer.
- We investigate the results by applying data preprocessing to the dataset, with the expectation that it will enhance the outcome.
- We aim to achieve the best results by experimenting with different hyperparameters on the four models we have selected.
- We conduct training on four models using two different optimizers and five different epochs. Thus, we aim to determine the best optimizer and number of epochs to achieve optimal results for this classification problem.

2. MATERIAL AND METHODOLOGY

The rapid development and proliferation of artificial intelligence are evident across many domains of life. Among these domains, perhaps the most important is the healthcare sector. Sub-disciplines of artificial intelligence such as machine learning and deep learning offer highly valuable solutions in healthcare. Early diagnosis, especially, holds critical importance for many life-threatening diseases. This is where machine learning and deep learning come into play. These methods are preferred for their ability to provide fast and reliable results. There are many deep learning techniques available for early diagnosis. In this study, classification supported by transfer learning will be performed using several architectures of CNN. The obtained results will be evaluated on various hyperparameters. All these processes are carried out using Google Colab and the Python programming language.

A publicly available dataset obtained from Kaggle is used [51]. The dataset comprises a total of 5,192 images. Additionally, the images in the dataset are divided into two classes: normal and oscc. CLAHE (Contrast Limited Adaptive Histogram Equalization) is applied on images to increase the contrast in the images and make the details more obvious. Clahe is an image processing technique used to enhance the contrast in images. Images are divided into subregions and separate histogram equalization is applied for each subregion. Additionally, excessive contrast increase is also prevented. This leads to the preference for this technique. After this process, the dataset is divided into 80% for training and 20% for testing purposes. Consequently, the training dataset comprises a total of 4,154 images, while the test dataset contains 1,038 images.

2.1. Pre-processing

A contrast enhancement process is applied on the data set to increase the contrast in the images and make the details more obvious. CLAHE is preferred for this procedure to be applied. This work is being conducted in Python, and the createCLAHE function of the OpenCV library is being used to apply this algorithm to

images. Additionally, the dataset is being subjected to normalization and size adjustment operations after these processes. The size of each image in the dataset is 224x224. The dataset has been divided into 80% for training and 20% for testing. As a result, the training dataset contains a total of 4,154 images, while the test dataset contains 1,038 images.

Different optimization algorithms and 5 different numbers of epochs are used for training. In each model, CrossEntropyLoss loss function and a learning rate of 0.001 are employed. CrossEntropyLoss is preferred as successful results are obtained when comparing predicted labels with the actual labels in classification problems. For classification, 4 different CNN models are selected: ResNet101, ResNet152, DenseNet169, and DenseNet201. These models are trained first with the Adam (Adaptive Moment Estimation) optimization algorithm, CrossEntropyLoss loss function, and a learning rate of 0.001. The models used have been pre-trained with ImageNet weights. Evaluations were conducted for 5 different epoch values. Subsequently, similar procedures are repeated by selecting the SGD (Stochastic Gradient Descent) optimization algorithm to investigate the effect of the optimizer.

As an example, Figures 1 and 2 show two samples from the open access dataset obtained from the Kaggle platform, before and after the application of the CLAHE method [51].

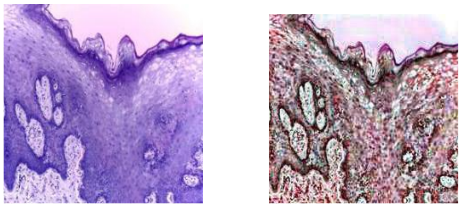

Figure 1. *Original image (Normal) & Clahe image (Normal) [51]*

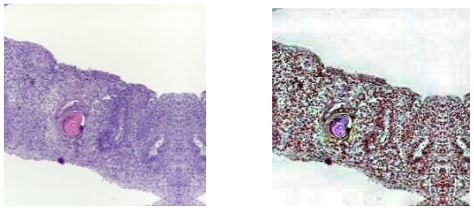

Figure 2. *Original image (OSCC) & Clahe image (OSCC) [51]*

2.2. Classification Method

Classification is carried out through transfer learning with the ResNet101, ResNet152, DenseNet169, and DenseNet201 algorithms. These methods are preferred for being trained faster, for achieving better performance, and for having computation costs reduced. Additionally, each model is pre-trained with ImageNet weights.

One of the models used, ResNet101, is a variant of the ResNet architecture, which was developed by Microsoft. It was first introduced in 2015 by Kaiming He et al. [52] in the article titled ’Deep Residual Learning for Image Recognition’. This architecture is composed of residual blocks instead of conventionally used layers. Skip connections are included in these blocks, which makes it easier to learn by directly adding the input of the block to the output of the block. ResNet101 is composed of 101 layers, indicating that this architecture is deep and can handle complex tasks on large datasets. On the other hand, ResNet152 is equipped with 152 layers, making it the deepest and most complex version of the ResNet architecture.

Another model used, DenseNet169, is a variant of the DenseNet architecture, which was first introduced in 2017 by Huang et al. [53] in the article titled ’Densely Connected Convolutional Networks.’ This architecture, which is much more advanced than residual learning, features a dense connectivity structure where the outputs of all preceding layers are connected to the outputs of each layer. In other words, all outputs coming from previous layers are passed as input to the next layer, and each layer receives the

outputs from all preceding layers to form its own output. DenseNet169 is a variant of the DenseNet architecture and consists of 169 layers. On the other hand, DenseNet201 has more layers and consists of 201 layers.

3. CONCLUSION

In this work, the accuracy values of these four models, each equipped with the same hyperparameters and operating on the same dataset, are compared. Additionally, while evaluating the models, the impact of the number of layers on accuracy is being investigated. These four models have been evaluated separately for two different optimizers. These four models are separately evaluated for two different optimizers. Throughout the study, CrossEntropyLoss is being used as the loss function, and a learning rate of 0.001 is being employed.

Below, the results for epochs are being provided for four models with the use of the Adam (Adaptive Moment Estimation) optimizer.

Table 1. *For accuracy value*

Epoch	ResNet101	ResNet152	DenseNet169	DenseNet201
10	0.9229	0.8854	0.9470	0.8940
20	0.8988	0.8969	0.9393	0.9037
30	0.9037	0.8921	0.9374	0.9470
40	0.9066	0.8882	0.9200	0.9432
50	0.8950	0.9075	0.9528	0.9268

The results, obtained by starting with 10 epochs and increasing the epoch count by 10 for each of the four different models, using Adam as the optimizer and CrossEntropyLoss as the loss function, have been presented in Table 1 based on accuracy values. Accuracy provides the ratio of correct predictions to the total data points. This metric is important in measuring overall performance. A high accuracy value indicates a high probability of the model making correct predictions. While there are many factors that influence this metric, changes in accuracy based on the number of epochs can be discussed in Table 1.

Table 2. *For precision value*

Epoch	ResNet101	ResNet152	DenseNet169	DenseNet201
10	0.9314	0.9041	0.9544	0.9024
20	0.8871	0.8895	0.9500	0.8961
30	0.9144	0.8808	0.9441	0.9592
40	0.9033	0.8957	0.9305	0.9530
50	0.9079	0.9118	0.9644	0.9288

The precision values, obtained by increasing the epoch count for four different models using Adam optimizer and CrossEntropyLoss as the loss function, are presented in Table 2. Precision expresses how many of the examples the model identifies as positive are actually positive. It is one of the preferred metrics for measuring the performance of models in classification problems.

Table 3. *For recall value*

Epoch	ResNet101	ResNet152	DenseNet169	DenseNet201
10	0.9247	0.8761	0.9458	0.8958
20	0.9189	0.9082	0.9344	0.9165
30	0.8935	0.9016	0.9354	0.9400
40	0.9153	0.8891	0.9150	0.9330
50	0.8858	0.9118	0.9402	0.9322

In Table 3, recall values for four different models have been presented based on the epoch count, with Adam optimizer having been used and CrossEntropyLoss having been employed as the loss function. Recall

provides the rate at which true positives are correctly predicted. Therefore, it is considered a crucial metric and high results are desired.

Table 4. *For F1-score*

Epoch	ResNet101	ResNet152	DenseNet169	DenseNet201
10	0.9280	0.8898	0.9500	0.8990
20	0.9027	0.8987	0.9421	0.9061
30	0.9038	0.8910	0.9397	0.9495
40	0.9092	0.8923	0.9226	0.9428
50	0.8945	0.9118	0.9521	0.9304

Table 4 presents F1-score values based on the epoch count for four different models, using Adam optimizer, CrossEntropyLoss as the loss function, and a learning rate of 0.001. When an evaluation is considered for the four models, Table 4 can be utilized. F1-score is the harmonic mean of precision and recall values, representing the evaluation of these two metrics together. It determines the accuracy and success of the model.

Table 5. *For specificity value*

Epoch	ResNet101	ResNet152	DenseNet169	DenseNet201
10	0.9208	0.8957	0.9484	0.8920
20	0.8779	0.8854	0.9447	0.8904
30	0.9140	0.8830	0.9395	0.9549
40	0.8974	0.8873	0.9255	0.9534
50	0.9047	0.9028	0.9653	0.9207

The results for the specificity value are presented in Table 5. Specificity provides the percentage of true negatives correctly identified. Having a high value of this measure is crucial in avoiding misdiagnosis for individuals who are not diseased.

Below, the results for epochs are being provided for four models with the use of the SGD (Stochastic Gradient Descent) optimizer.

Table 6. *For accuracy value*

Epoch	ResNet101	ResNet152	DenseNet169	DenseNet201
10	0.9557	0.9624	0.9566	0.9644
20	0.9461	0.9412	0.9634	0.9615
30	0.9644	0.9422	0.9547	0.9557
40	0.9441	0.9634	0.9499	0.9701
50	0.9461	0.9663	0.9461	0.9653

The results, obtained by starting with 10 epochs and increasing the epoch count by 10 for each of the four different models, using SGD as the optimizer and CrossEntropyLoss as the loss function, have been presented in Table 6 based on accuracy values. The conclusion that the increase in epoch count has different effects on each model can be derived from the table.

Table 7. *For precision value*

Epoch	ResNet101	ResNet152	DenseNet169	DenseNet201
10	0.9601	0.9642	0.9610	0.9754
20	0.9461	0.9404	0.9712	0.9766
30	0.9660	0.9496	0.9705	0.9544
40	0.9555	0.9652	0.9562	0.9737
50	0.9527	0.9594	0.9507	0.9744

The precision values, obtained by increasing the epoch count for four different models using SGD optimizer and CrossEntropyLoss as the loss function, are presented in Table 7. It has been concluded from the table that the epoch count affects the precision value differently for each model.

Table 8. *For precision value*

Epoch	ResNet101	ResNet152	DenseNet169	DenseNet201
10	0.9487	0.9659	0.9556	0.9556
20	0.9561	0.9457	0.9564	0.9470
30	0.9642	0.9391	0.9390	0.9614
40	0.9338	0.9615	0.9454	0.9682
50	0.9379	0.9725	0.9435	0.9555

In Table 8, recall values for four different models have been presented based on the epoch count, with SGD optimizer having been used and CrossEntropyLoss having been employed as the loss function.

Table 9. *For precision value*

Epoch	ResNet101	ResNet152	DenseNet169	DenseNet201
10	0.9543	0.9650	0.9582	0.9653
20	0.9510	0.9430	0.9637	0.9615
30	0.9650	0.9443	0.9544	0.9578
40	0.9445	0.9633	0.9507	0.9709
50	0.9452	0.9659	0.9470	0.9648

Table 9 presents F1-score values based on the epoch count for four different models, using SGD optimizer, CrossEntropyLoss as the loss function, and a learning rate of 0.001. When an evaluation is considered for the four models, Table 9 can be utilized.

Table 10. *For specificity value*

Epoch	ResNet101	ResNet152	DenseNet169	DenseNet201
10	0.9623	0.9583	0.9577	0.9738
20	0.9339	0.9365	0.9705	0.9764
30	0.9645	0.9455	0.9707	0.9493
40	0.9548	0.9653	0.9546	0.9721
50	0.9541	0.9602	0.9487	0.9750

The results for the specificity value are provided in Table 10. Based on this table, an analysis can be conducted to determine which model performs better in correctly identifying true negatives. Additionally, the effects of epoch count on this metric can also be discussed. The F1-Score Graphic for Adam Optimizer & F1-Score Graphic for SGD Optimizer is given in Figure 3.

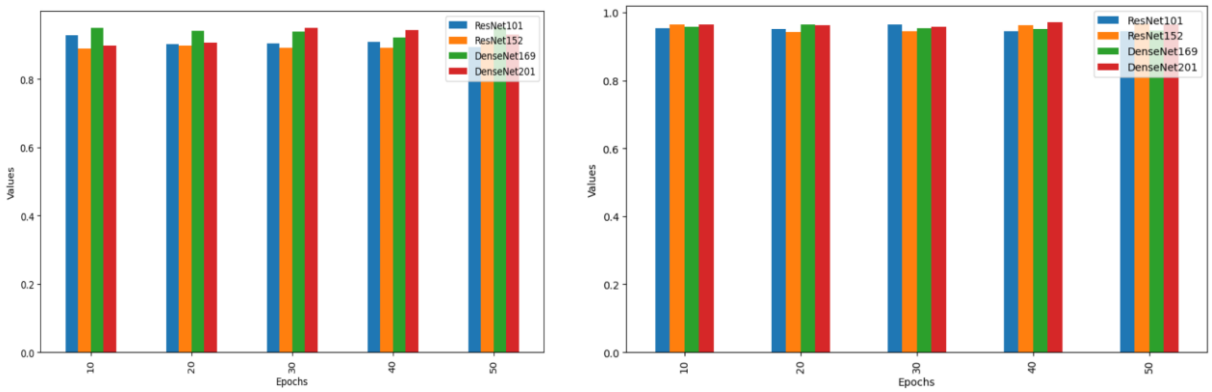

Figure 3. *F1-score Graphic for Adam Optimizer & F1-score Graphic for SGD Optimizer*

The observation of the impact of the optimizer on the models is enabled by this study. Similarly, within the same architecture, the observation of how the classification accuracy of the model is influenced by the number of layers can be made. Additionally, an evaluation is being conducted between architectures to investigate the results provided by two different architectures for the same problem. Furthermore, an examination is being conducted on the extent to which the model accuracy is affected by the number of epochs.

As can be seen from the results of the study, success has been achieved in the classification problem by these four models, each with different architectures. However, an increase in the number of layers results in an increase in depth and a significant impact on the classification accuracy. Similarly, it is observed that the accuracy is influenced by the number of epochs and the optimizer, which are also considered significant factors. If the number of epochs is insufficient, the risk of underfitting is encountered, and an excessive number of epochs can lead to the problem of overfitting.

4. DISCUSSION

In this research, a classification is being performed on Oral Squamous Cell Carcinoma biopsy images using different CNN models with various hyperparameters. To analyze the models used, accuracy (AC), precision (PR), specificity (SP), recall (RE), and F1-score (F) parameters are employed. The highest results obtained are highlighted in Tables 11 and 12. Based on the results obtained from the performance parameters, it can be said that the models used and the hyperparameters applied on the models have yielded highly successful results in this classification problem. The conducted study can be preferred for early diagnosis in the healthcare field. Due to its fast results and low cost, this method can be highly preferred for diagnosis in low-income, underdeveloped countries where oral cancer is prevalent, helping to gain sufficient time for treatment. Thus, the mortality and morbidity rates of the disease can be reduced, thereby increasing the survival rate. Additionally, the impact of hyperparameters on models and architectures is also being explained by this study. This also sheds light on future studies to be conducted.

The comparison of the findings with previous studies is given in Tables 11 and 12.

Table 11. *Comparing findings with prior studies. (Proposed models for Adam optimizer)*

Author	Models	AC (%)	F (%)	RE (%)	SP (%)	PR (%)
Welikala et al. [33]	ResNet101	N/A	78.30	93.88	N/A	67.15
Chu et al. [35]	SVM, KNN	70.59	N/A	N/A	84.12	N/A
Alhazmi et al. [37]	ANN	78.95	N/A	N/A	60.00	N/A
Rahman et al. [43]	AlexNet	90.06	90.15	N/A	87.38	N/A
Proposed Model1	ResNet101	92.29	92.80	92.47	92.08	93.14
Proposed Model2	ResNet152	90.75	91.18	91.18	90.28	91.18
Proposed Model3	DenseNet169	**95.28**	**95.21**	**94.58**	**96.53**	**96.44**
Proposed Model4	DenseNet201	94.70	94.95	94.00	95.49	95.92

Table 12. *Comparing findings with prior studies. (Proposed models for SGD optimizer)*

Author	Models	AC (%)	F (%)	RE (%)	SP (%)	PR (%)
Welikala et al. [33]	ResNet101	N/A	78.30	93.88	N/A	67.15
Chu et al. [35]	SVM, KNN	70.59	N/A	N/A	84.12	N/A
Alhazmi et al. [37]	ANN	78.95	N/A	N/A	60.00	N/A
Rahman et al. [43]	AlexNet	90.06	90.15	N/A	87.38	N/A
Proposed Model1	ResNet101	96.44	96.50	96.42	96.45	96.60
Proposed Model2	ResNet152	96.63	96.59	**97.25**	96.53	96.52
Proposed Model3	DenseNet169	96.34	96.37	95.56	97.07	97.12
Proposed Model4	DenseNet201	**97.01**	**97.09**	96.82	**97.64**	**97.66**

ACKNOWLEDGEMENTS

This work was supported by the Research Fund of the Yildiz Technical University, Project Number: FYL-2024-6104.

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A., "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries", CA: A Cancer Journal for Clinicians, 68(6): 394-424, (2018). DOI: 10.3322/caac.21492

[2] Kalavrezos, N., Scully, C., "Mouth cancer for clinicians part 2: epidemiology", Dental Update, 42(4): 354-359, (2015). DOI: 10.12968/denu.2015.42.4.354

[3] Sankaranarayanan, R., Ramadas, K., Amarasinghe, H., Subramanian, S., Johnson, N. "Oral cancer: prevention, early detection, and treatment", Cancer: Disease Control Priorities, Third Edition, 3: (2015). DOI: 10.1596/978-1-4648-0349-9_ch5

[4] Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries", CA: A Cancer Journal for Clinicians, 71(3): 209-249, (2021). DOI: 10.3322/caac.21660

[5] World Health Organization. Oral Health. (2018). Available online: https://gco.iarc.fr/today/data/factsheets/cancers/1-Lip-oral-cavity-fact-sheet.pdf accessed on 11 March (2020). Access: https://gco.iarc.fr/today/en

[6] Torres-Rosas, R., Torres-Gómez, N., Hernández-Juárez, J., Pérez-Cervera, Y., Hernández-Antonio, A., Argueta-Figueroa, L., "Reported epidemiology of cancer of the lip, oral cavity and oropharynx in Mexico", Revista Medica del Instituto Mexicano del Seguro Social, 58(4): 494-507, (2020). DOI: 10.24875/rmimss.m20000075

[7] Warnakulasuriya, S., "Global epidemiology of oral and oropharyngeal cancer", Oral Oncology, 45(4-5): 309-316, (2009). DOI: 10.1016/j.oraloncology.2008.06.002

[8] Speight, P.M., Farthing, P.M., "The pathology of oral cancer", British Dental Journal, 225(9): 841-847, (2018). DOI: 10.1038/sj.bdj.2018.926

[9] Chow, L.Q.M., "Head and neck cancer. Reply", The New England Journal of Medicine, 382(20): e57, (2020). DOI: 10.1056/NEJMc2001370

[10] Wong, T.S.C., Wiesenfeld, D., "Oral cancer", Australian Dental Journal, 63: S91-S99, (2018). DOI: 10.1111/adj.12594

[11] Joseph, B.K., "Oral cancer: prevention and detection", Medical Principles and Practice, 11(Suppl. 1): 32-35, (2002). DOI: 10.1159/000057776

[12] Su, Y.F., Chen, Y.J., Tsai, F.T., Li, W.C., Hsu, M.L., Wang, D.H., Yang, C.C., "Current insights into oral cancer diagnostics", Diagnostics, 11(7): 1287, (2021). DOI: 10.3390/diagnostics11071287

[13] Chattopadhyay, I., Verma, M., Panda, M., "Role of oral microbiome signatures in diagnosis and prognosis of oral cancer", Technology in Cancer Research & Treatment, 18: 1533033819867354, (2019). DOI: 10.1177/1533033819867354

[14] Trimarchi, M., Bertazzoni, G., Bussi, M., "Cocaine induced midline destructive lesions", Rhinology, 52(2): 104-111, (2014). DOI: 10.4193/Rhino13.112

[15] Trimarchi, M., Bellini, C., Fabiano, B., Gerevini, S., Bussi, M., "Multiple mucosal involvement in cicatricial pemphigoid", Acta Otorhinolaryngologica Italica, 29(4): 222, (2009).

[16] Biafora, M., Bertazzoni, G., Trimarchi, M., "Maxillary sinusitis caused by dental implants extending into the maxillary sinus and the nasal cavities", Journal of Prosthodontics, 23(3): 227-231, (2014). DOI: 10.1111/jopr.12123

[17] Trimarchi, M., Bondi, S., Della Torre, E., Terreni, M.R., Bussi, M., "Palate perforation differentiates cocaine-induced midline destructive lesions from granulomatosis with polyangiitis", Acta Otorhinolaryngologica Italica, 37(4): 281, (2017). DOI: 10.14639/0392-100X-1586

[18] Lanzillotta, M., Campochiaro, C., Trimarchi, M., Arrigoni, G., Gerevini, S., Milani, R., Bozzolo, E., Biafora, M., Venturini, E., Cicalese, M.P., Shone, J.H., Sabbadini M.G., Della-Torre, E., "Deconstructing IgG4-related disease involvement of midline structures: Comparison to common mimickers", Modern Rheumatology, 27(4): 638-645, (2017). DOI: 10.1080/14397595.2016.1227026

[19] Sosiawan, A., Chusnita, R., Laksanti, P.A.M., Anwar, A.A., Ramadhani, N.F., Nugraha, A.P., "Utilization of artificial intelligence-assisted histopathological detection in surveillance of oral squamous cell carcinoma staging: A narrative review", World Journal of Advanced Research and Review, 16(3): 54-59, (2022). Access: https://repository.unair.ac.id/119228/

[20] Chinn, S.B., Myers, J.N., "Oral cavity carcinoma: current management, controversies, and future directions", Journal of Clinical Oncology, 33(29): 3269-3276, (2015). DOI: 10.1200/JCO.2015.61.2929

[21] Silverman, S., Kerr, A.R., Epstein, J.B., "Oral and pharyngeal cancer control and early detection", Journal of Cancer Education, 25: 279-281, (2010). DOI: 10.1007/s13187-010-0045-6

[22] McCullough, M.J., Prasad, G., Farah, C.S. "Oral mucosal malignancy and potentially malignant lesions: an update on the epidemiology, risk factors, diagnosis and management", Australian Dental Journal, 55: 61-65, (2010). DOI: 10.1111/j.1834-7819.2010.01200.x

[23] Gómez, I., Seoane, J., Varela-Centelles, P., Diz, P., Takkouche, B., "Is diagnostic delay related to advanced-stage oral cancer? A meta-analysis", European Journal of Oral Sciences, 117(5): 541-546, (2009). DOI: 10.1111/j.1600-0722.2009.00672.x

[24] Hema Shree, K., Ramani, P., Sherlin, H., Sukumaran, G., Jeyaraj, G., Don, K.R., Santhanam, A., Ramasubramanian, A., Sundar, R., "Saliva as a diagnostic tool in oral squamous cell carcinoma-a systematic review with meta analysis", Pathology & Oncology Research, 25: 447-453, (2019). DOI: 10.1007/s12253-019-00588-2

[25] Ulaganathan, G., Niazi, K.T.M., Srinivasan, S., Balaji, V.R., Manikandan, D., Hameed, K.A.S., Banumathi, A., "A clinicopathological study of various oral cancer diagnostic techniques", Journal of Pharmacy & Bioallied Sciences, 9(Suppl 1): S4, (2017). DOI: 10.4103/jpbs.JPBS_110_17

[26] Beristain-Colorado, M.P., Castro-Gutiérrez, M.E.M., Torres-Rosas, R., Vargas-Treviño, M., Moreno-Rodríguez, A., Fuentes-Mascorro, G., Argueta-Figueroa, L., "Application of neural networks for the detection of oral cancer: A systematic review", Dental and Medical Problems, 61(1): 121-128, (2024). DOI: 10.17219/dmp/159871

[27] Tanriver, G., Soluk Tekkesin, M., Ergen, O., "Automated detection and classification of oral lesions using deep learning to detect oral potentially malignant disorders", Cancers, 13(11): 2766, (2021). DOI: 10.3390/cancers13112766

[28] Uthoff, R.D. Song, B., Sunny, S., Patrick, S., Suresh, A., Kolur, T., Keerthi, G., Spires, O., Anbarani, A., Wilder-Smith, P., Kuriakose, M.A., Birur P., Liang R., "Point-of-care, smartphone-based, dual-modality, dual-view, oral cancer screening device with neural network classification for low-resource communities", PloS One, 13(12): e0207493, (2018). DOI: 10.1371/journal.pone.0207493

[29] Xu, S., Liu, C., Zong, Y., Chen, S., Lu, Y., Yang, L., Ng, E.Y.K., Wang, Y., Wang, Y., Liu, Y., Hu, W., Zhang, C., "An early diagnosis of oral cancer based on three-dimensional convolutional neural networks", IEEE Access, 7: 158603-158611, (2019). DOI: 10.1109/ACCESS.2019.2950286

[30] Gupta, R.K., Kaur, M., Manhas, J., "Tissue level based deep learning framework for early detection of dysplasia in oral squamous epithelium", Journal of Multimedia Information System, 6(2): 81-86, (2019). DOI: 10.33851/JMIS.2019.6.2.81

[31] Jeyaraj, P.R., Samuel Nadar, E.R., "Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm", Journal of Cancer Research and Clinical Oncology, 145: 829-837, (2019). DOI: 10.1007/s00432-018-02834-7

[32] Fu, Q., Chen, Y., Li, Z., Jing, Q., Hu, C., Liu, H., Bao, J., Hong, Y., Shi, T., Li, K., Zou, H., Song, Y., Wang, H., Wang, X., Wang, Y., Liu, J., Liu, H., Chen, S., Chen, R., Zhang, M., Zhao, J., Xiang, J., Liu, B., Jia, J., Wu, H., Zhao, Y., Wan, L., Xiong, X., "A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study", E Clinical Medicine, 27: (2020). DOI: 10.1016/j.eclinm.2020.100558

[33] 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., Amtha, R., Patil, K., Tilakaratne, W.M., Gibson, J., Cheong, S.C., Ann Barman, S., "Automated detection and classification of oral lesions using deep learning for early detection of oral cancer", IEEE Access, 8: 132677-132693, (2020). DOI: 10.1109/ACCESS.2020.3010180

[34] 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 Networks, 128: 47-60, (2020). DOI: 10.1016/j.neunet.2020.05.003

[35] Chu, C.S., Lee, N.P., Adeoye, J., Thomson, P., Choi, S.W., "Machine learning and treatment outcome prediction for oral cancer", Journal of Oral Pathology & Medicine, 49(10): 977-985, (2020). DOI: 10.1111/jop.13089

[36] Lin, H., Chen, H., Weng, L., Shao, J., Lin, J., "Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis", Journal of Biomedical Optics, 26(8): 086007-086007, (2021). DOI: 10.1117/1.JBO.26.8.086007

[37] Alhazmi, A., Alhazmi, Y., Makrami, A., Masmali, A., Salawi, N., Masmali, K., Patil, S., "Application of artificial intelligence and machine learning for prediction of oral cancer risk", Journal of Oral Pathology & Medicine, 50(5): 444-450, (2021). DOI: 10.1111/jop.13157

[38] Shavlokhova, V., Sandhu, S., Flechtenmacher, C., Koveshazi, I., Neumeier, F., Padrón-Laso, V., Jonke, Ž., Saravi, B., Vollmer, M., Vollmer, A., Hoffman, J., Engel, M., Ristov, O., Freudlsperger, C., "Deep learning on oral squamous cell carcinoma ex vivo fluorescent confocal microscopy data: a feasibility study", Journal of Clinical Medicine, 10(22): 5326, (2021). DOI: 10.3390/jcm10225326

[39] Warin, K., Limprasert, W., Suebnukarn, S., Jinaporntham, S., Jantana, P. "Automatic classification and detection of oral cancer in photographic images using deep learning algorithms", Journal of Oral Pathology & Medicine, 50(9): 911-918, (2021). DOI: 10.1111/jop.13227

[40] Jubair, F., Al-karadsheh, O., Malamos, D., Al Mahdi, S., Saad, Y., Hassona, Y., "A novel lightweight deep convolutional neural network for early detection of oral cancer", Oral Diseases, 28(4): 1123-1130, (2022). DOI: 10.1111/odi.13825

[41] Nanditha, B.R., Geetha, A., Chandrashekar, H.S., Dinesh, M.S., Murali, S., "An ensemble deep neural network approach for oral cancer screening", International Journal of Online and Biomedical Engineering, (2021). DOI: 10.3991/ijoe.v17i02.19207

[42] Alkhadar, H., Macluskey, M., White, S., Ellis, I., Gardner, A., "Comparison of machine learning algorithms for the prediction of five-year survival in oral squamous cell carcinoma", Journal of Oral Pathology & Medicine, 50(4): 378-384, (2021). DOI: 10.1111/jop.13135

[43] Rahman, A., Alqahtani, A., Aldhafferi, N., Nasir, M.U., Khan, M.F., Khan, M.A., Mosavi, A., "Histopathologic oral cancer prediction using oral squamous cell carcinoma biopsy empowered with transfer learning", Sensors, 22(10): 3833, (2022). DOI: 10.3390/s22103833

[44] Deif, M.A., Attar, H., Amer, A., Elhaty, I.A., Khosravi, M.R., Solyman, A.A.A., "Diagnosis of oral squamous cell carcinoma using deep neural networks and binary Particle Swarm optimization on histopathological images: an AIoMT approach", Computational Intelligence and Neuroscience, 2022(1): 6364102, (2022). DOI: 10.1155/2022/6364102

[45] Liyanage, V., Tao, M., Park, J.S., Wang, K.N., Azimi, S., "Malignant and non-malignant oral lesions classification and diagnosis with deep neural networks", Journal of Dentistry, 137: 104657, (2023). DOI: 10.1016/j.jdent.2023.104657

[46] Zhou, M., Jie, W., Tang, F., Zhang, S., Mao, Q., Liu, C., Hao, Y., "Deep learning algorithms for classification and detection of recurrent aphthous ulcerations using oral clinical photographic images", Journal of Dental Sciences, 19(1): 254-260, (2024). DOI: 10.1016/j.jds.2023.04.022

[47] Manikandan, J., Krishna, B.V., Varun, N., Vishal, V., Yugant, S., "Automated Framework for Effective Identification of Oral Cancer Using Improved Convolutional Neural Network", IEEE Xplore, 1-7, (2023). DOI: 10.1109/ICONSTEM56934.2023.10142794

[48] Badik, Ş.T., Akar, M., "Machine learning classification models for the patients who have heart failure", Sigma Journal of Engineering and Natural Sciences, 42(1): 235-244, (2024). DOI: 10.14744/sigma.2023.00095

[49] Akar, M., Sirakov, N.M., Mete, M., "Clifford algebra multivectors and kernels for melanoma classification", Mathematical Methods in the Applied Sciences, 45(7): 4056-4068, (2022). DOI: 10.1002/mma.8034

[50] Akar, M., Sirakov, N.M., "Support vector machine skin lesion classification in Clifford algebra subspaces", Applications of Mathematics, 64(5): 581-598, (2019). DOI: 10.21136/AM.2019.0292-18

[51] Available online: https://www.kaggle.com/ashenafifasilkebede/dataset?select=val Access date: 06.01.2022

[52] He, K., Zhang, X., Ren, S., Sun, J., "Deep residual learning for image recognition", IEEE Xplore, 770-778, (2016). DOI: 10.1109/CVPR.2016.90

[53] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., "Densely connected convolutional networks", IEEE Xplore, 4700-4708, (2017). DOI: 0.1109/CVPR.2017.243