Automatic Detection and Classification of Oral Cancer from Photographic Images Using Attention Maps and Deep Learning

Sayyada Hajera Begum*1, P. Vidyullatha2

Abstract*:* Deep learning – Convolutional Neural Networks (DL-CNN) has shown a lot of potential in identifying cancerous and non-cancerous oral lesions from oral photographic images. Moreover, the accuracy of CNNs can be improved by guiding the model to concentrate on the cancerous areas rather than not so important surrounding areas. The paper proposes to develop a DL-CNN model that directly focusses on cancerous areas in the lip, tongue and cheek images. The proposed CNN model applies transfer learning using DenseNet201 as base model for the detection. The model works by identifying region of interests (RoIs) and generating attention maps for the images that helps the model in understanding the area of focus by highlighting it and for classifying the oral lesions correctly. The results demonstrated effectiveness of the approach by giving an accuracy of 84.7% for the classification of oral cancer lesions.

Keywords: O*ral photographic images, Transfer learning, Deep Learning, Convolutional Neural networks (CNN), Attention maps*

1. Introduction

Oral cancer is one of the most commonly found cancers in developing countries. Lack of awareness and knowledge among the public regarding oral cancer has become a vital reason for late prognosis of cancerous lesions [1]. Identifying which Oral lesions are prone to transform into cancerous lesions requires experienced clinicians and specialists. Recent successes of computational intelligence and computer vision for automated diagnosis in healthcare have successfully transformed medical image analysis using various Deep Learning techniques [2]. Automated classification of cancerous and non-cancerous oral lesions can support physicians by easing the manual effort and cancerous and non-cancerous oral lesions can support physicians by easing the manual effort and at the same time giving inexpensive access to life-saving diagnosis [3].

As per WHO over six lakh oral cancer cases every year are diagnosed ss potentially malignant disorders (OPMD) which are identified as visible reddish, whitish, or whitish-reddish lesions in the oral region. OPMDs are more prevalent in Asian countries as much as 11%. The OPMDs have a 1% chance of malignant transformation which later leads to painful invasive

treatment and high mortality. Early diagnosis of these OPMDs becomes very essential to decrease the treatment cost and to increase the survival rate. Due to the resemblance of OPMDs' appearance with that of inflammatory lesions appearance, it is sometimes misdiagnosed or late diagnosis might lead to OPMDs transforming into cancer [4].

Chewing betel quid, tobacco, and smoking are major causes of oral cancer. OPMDs can be detected during a routine checkup by an experienced dentist. On identification of a suspicious oral skin lesion, the patients are referred to consult a specialist for further analysis. This detection of suspicious oral lesions at an early stage can reduce the treatment cost significantly. If oral lesions are undetected, they can easily spread to the neck and lungs. Thus, on-time identification of OPMD will reduce the chances of transforming into cancer. Due to the high cost of cancer screening and its treatment, it has become important to introduce a cost-effective and efficient screening approach for OPMDs [5].

The manual analysis of biopsy slides for detecting cancerous cells consumes more time and is prone to human errors. Recent studies have discovered that DL models can help clinicians by reducing the diagnosis time and improving the overall efficacy of the process. Thus DL models can support clinicians in better decision-making. Among the various DL models, the CNN model became more famous due to its specialty in texture classification in images. Recently oral cancer detection has been applied to photographic

1Research Scholar, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.
2Associate Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.
*Corresponding Author Email:
sayyada.hajera07@gmail.com

images but due to the unavailability of a proper dataset, the research has been limited. Moreover, the transfer learning approach has been widely applied in medical imaging to get better results from the CNN models [6].

2. Related Work

The growing success of Deep Learning techniques (DLTs) has successfully identified visual patterns from photographic images leading to various models building for the diagnosis of oral cancer.

Roshan Alex [7] used two DL approaches namely ResNet-101 and Faster-RCNN for image classification and object detection respectively for better results. Their proposed model showed effective progress with an F1 score of 87%. Mohammed Shamim [8] implemented a DCNN model using Resnet-50 and successfully classified cancerous tongue lesions. The proposed DCNN model was able to perform multiclass detection of tongue lesions into geographic, strawberry, hairy, fissured, and leucoplakia with an accuracy of 97%.

Figueroa et al. [9] proposed a CNN model with a two-stage training process using a Guided Network (GAIN) on oral photographic images. The proposed network achieved an accuracy of 86.38% after stage 2 training. Huiping Lin et al. [10] proposed using HRNet as a base model and training a CNN model on the images taken by smart cellphones. The model achieved an F1-score value of 83.6% on HRNet.

Baichen Ding [11] implemented a CNN model for the detection of dental carries that can transform into cancer by oral photographs taken by smartphones using the YOLOv3 algorithm. The trained model achieved a precision of 76.92%. Bofan Song [12] proposed an image classification model using dual-mode autofluorescence and white light images based on CNN. The VGG-CNN model achieved an accuracy of 86.9% with 4-fold cross-validation.

Gizem [13] classified oral skin lesions into non-cancerous, cancerous and potentially cancerous using a detector network that works with YOLOv5 and a classifier model. The model achieved an average precision of 95% using YOLO.

Fahed Jubair et al. [14] proposed a hybrid CNN model using EfficientNet-B0 as base model. The proposed CNN model successfully detected cancerous images with an accuracy of 85.0%. Nanditha [15] combined Resnet-50 and VGG-16 models and applied them to photographic images of oral lesions to get 96.2% accuracy in results. Begum et al. [16] proposed a hybrid model based on DenseNet to classify histopathological images as cancerous and non-cancerous and successfully achieved 91% accuracy.

Chan et al. [17] proposed a DL-CNN model comprising two layers, the lower layer does segments and marks ROIs and the second layer detects oral cancer. The model reported a specificity of 71.29% and a sensitivity of 96.87%. Very recently, DLTs have been applied for oral cancer detection on an oral photographic image

dataset, but the dataset contains fewer images [18], still, authors exhibited the efficacy of DLTs with improved accuracy, specificity, and sensitivity to detect the oral cancer lesions. Small and imbalanced datasets can create hindrance in training DL models properly and to avoid this, datasets can be balanced in the pre-processing stage by applying SMOTE algorithm [19].

Most of the studies mentioned above are based on microscopic images taken using innovative imaging systems. However, oral lesion images can be easily clicked using Smartphone cameras and can be provided as input to various DL models. Hence a DL model capable of working with photographic images is the need of the hour.

Thus, the research explores the advantages of DL-CNN models for the diagnosis of Oral cancer lesions. The proposed model applies a pre-trained model with transfer learning and modifies it with extra added layers for efficient detection and classification of specific patterns of the suspicious oral lesions in the photographic images. In particular, we classified photographic image datasets as cancerous and non-cancerous classes using a DL-CNN-based classification. We also present the performance of our model using various performance evaluators by applying transfer learning and data pre-processing.

Section 3 of the paper explains the dataset used, and algorithms applied to the dataset for pre-processing such as region of interest (RoI) detection, attention map generation, and model formulation. Section 4 describes the experimental results and also provides a comparison with other notable works utilizing oral photographic images for cancer detection. Later, the paper is concluded in section 5.

3. Methodology

This section describes the dataset being used; algorithm adapted for RoI detection and attention map generation, and DL-CNN model formulation.

3.1. Dataset

The efficiency of any DL-CNN model also depends on the dataset chosen for training. The oral photographic image dataset used in the research work is accumulated using the Google search engine. We also used a few oral photographic images from the source [20]. The images are of different sizes, lightening conditions, and different regions of the mouth making them naturally diversified as per the physician's perception. The dataset images are verified and annotated manually by a certified medical practitioner before applying to the model. All the images are stored in JPEG format. The images are categorized into two directories with cancer and non-cancerous labels. A total of 1000 mouth, lip, and tongue images are available, out of which 700 images cater to cancer and 300 cater to non-cancer. For maintaining uniformity, all the images have been resized to 224×224 dimensions and have been

normalized in the 0-1 range. The dataset is randomly split for training and testing in the ratio of 80:20. Fig.1 shows different samples of noncancerous images and cancerous images. Fig. 2 shows different images with the variation and position of the oral cancer region.


Fig. 1. Sample Non-Cancerous (Left) and Cancerous (Right) Images
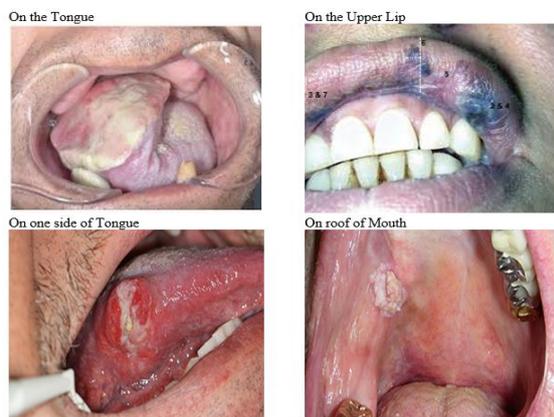
Fig. 2. Sample Images of Oral Cancer for Different Parts of the Mouth

### 3.2. Region of Interest (RoI) Detection in the Images

Regions of Interest (ROIs) are an essential component of object detection tasks, as they indicate the location of an object within an image. RoIs are usually bounded by rectangular shapes that surround objects in an image. They are defined by four parameters: X and Y coordinates at the upper-left corner of the box, in addition to the width and height of the box. RoIs are used to mark the objects by specifying the region of the image containing the object. The proposed work uses a region proposal network (RPN) to mark objects in the image with boxes. Region Proposal is the first step in the object detection framework. Regions of Interest are highlighted as rectangular boxes referred to as bounding boxes. It detects regions with a high likelihood of an image as foreground regions and the rest as background regions. The RPN generates multiple candidates bounding boxes at different scales and aspect ratios. The boxes are then provided to Intersection Over Union (IoU) measure which is a match between the predicted and true box. The greater the overlap, the better the score. Fig. 3 gives a demonstration of IoU on images. The intersection is given by expression (1).
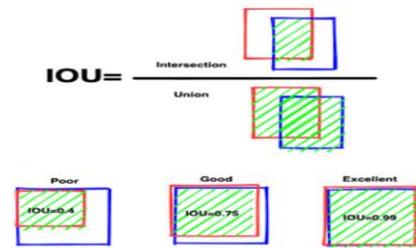
**Fig. 3.** Intersection Over Union, True Image (Blue Box) And Predicted Image (Red Box)

$$IoU = \frac{(Bt \cap Bp)}{(Bt \cup Bp)} \qquad (1)$$

where Bt is the true box and Bp is the predicted box. These candidates' bounding boxes are then passed through a CNN model that extracts features from each bounding box. The features are then fed into a classifier, which predicts the probability of object availability in the bounding box. Once the classifier has predicted the probability, a non-maximum suppression (NMS) function is applied which identifies the most likely bounding box for each object and discards overlapping bounding boxes that are less likely to contain the object. The remaining bounding boxes are then used to localize the object within the image. The final image with Identified RoI is then cropped to get a focused cancerous image for further analysis. Fig. 4 gives a flow chart for identifying ROI in the images.
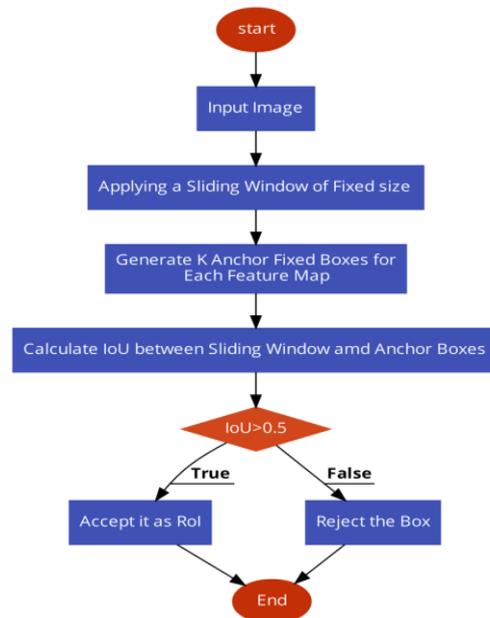
**Fig. 4.** Flow Chart Representing ROI Identification

Fig. 5 shows some sample cancerous images with identified ROI and their cropped counterpart for better results.
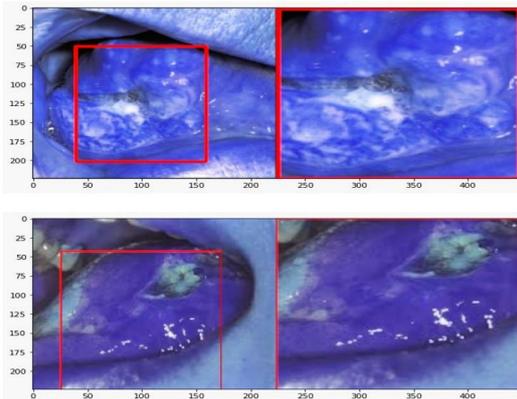
Fig. 5. Detection of Oral Cancer Location with ROI Cropping

3.3. Generating Attention Maps for RoI

Attention maps are created from the cropped bounding boxes to focus on the region of importance (ROI) within the image. The attention map highlights the pixels within the bounding box that are most informative for the given task. We have used COLORMAP_JET color map available in the OpenCV library to create visualizations for heatmaps. This color map ranges from blue to green to yellow to red, with higher values being represented by warmer colors. Specifically, the colors in the COLORMAP_JET are mapped as follows: Blue (low pixel values), Green (medium-low pixel values), Yellow (medium-high pixel values), and Red (high pixel values). Fig. 6 shows the process of creating attention maps from bounding boxes. Fig. 7 represents cancerous images and their respective attention maps.
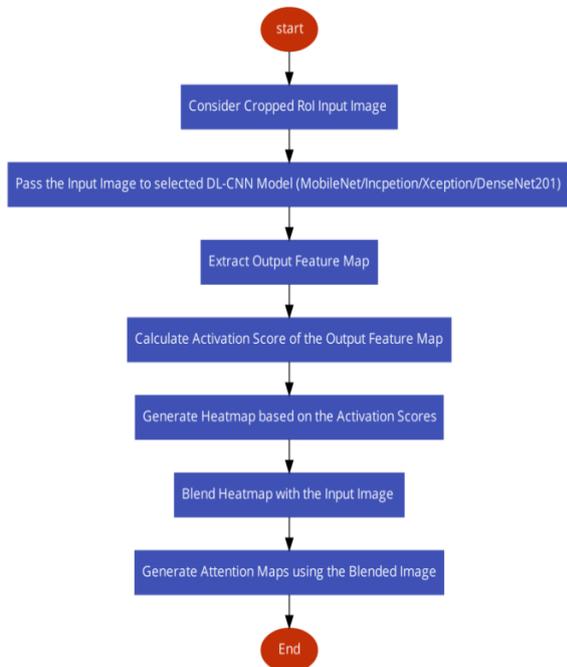

Fig. 6. Flow Chart for Generating Attention Maps
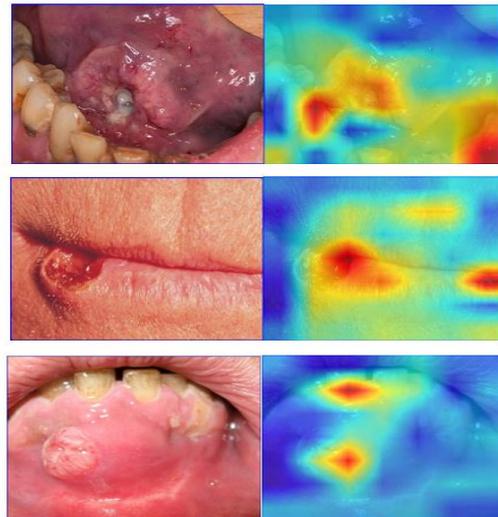
Fig. 7. Original Cancerous Image (Left), Respective Attention Maps (Right)

3.4. Model Formulation

CNN is a Deep Learning model containing specific regions of lower layers transferred to the higher layers to process raw pixel data for image classification. CNN has four key layers: A convolutional layer that extracts features and generates feature maps from the input images, A rectified linear unit (ReLU) function layer that provides acceleration for implementing complex functions, a max pool layer that reduces computational complexity, and a fully connected end layer for the classification. A CNN model is created by using a combination of one or more of the above layers, and its hyperparameters are tuned to achieve better detection and classification results. A general CNN structure is shown in Fig. 8.
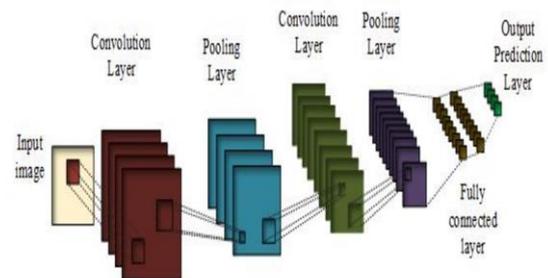
Fig. 8. General CNN Model

Several pre-trained DL-CNN models are available in the Keras library [21] which can be used for classifying images. In our proposed work, we have specifically considered MobileNet [23], InceptionV3 [24], Xception [25], and DenseNet201 [26] models by utilizing the transfer learning approach. The above models are initially trained on ImageNet [22] dataset and are later fine-tuned over the oral photographic dataset. Table 1 gives a general specification of the selected pre-trained models.

Table 1. Parameter Specifications of selected pre-trained models

DL-CNN Models	No. of Layers	Parameter size (millions)
MobileNet [23]	55	4.30
Inception [24]	189	24.0
Xception [25]	81	22.91
DenseNet [26]	121	8.06

To perform the CNN modeling, we started with one of the CNN pre-trained models and then additional layers of Convolution, MaxPooling, Dropout, Flatten, and Dense layers are added to build a hybrid proposed model.

The convolution layers apply a 3x3 filter of stride [1 1] to the image to generate feature maps. Max-Pooling layer applies a 2x2 pixel window of stride [2 2]. The model is trained with 80% of the images and 20% for validation. The proposed model is then evaluated for performance using Precision, Recall, F-measure, and Accuracy parameters. The following steps illustrate the formulation of our proposed DL-CNN.

1. Initially, the oral photographic images are pre-processed and then provided as input to RPN to find RoIs in the images.
2. Images are then cropped to get the potential RoIs.
3. The potential RoIs are generated and then converted to attention maps by blending heatmaps with the original images.
4. Generated attention maps are then used for training the CNN model for classification.
5. Feature maps are generated from individual images by applying convolution filters in the Convolutional layer.
6. To accomplish non-linear transformation on the input images ReLU function is applied.
7. Images from step 6 are then given to the MaxPooling layer to generate feature maps.
8. A DropOut layer is then used to avoid over fitting the model.
9. Steps 5-8 are repeated 2 more times.
10. The pooled features generated from the above steps are first flattened before applying to the Dense layer which enhances the speeding time.

4. Experimental Results

To assess the proposed model, we formed two directories with cancer and noncancer labels. We collected a total of 700 photographic images with cancer lesions and 300 noncancer photographic images, for a total of 1000 images. The dataset is then divided into 80:20 ratios containing an equal distribution of both directories. The proposed work has been implemented using Keras and TensorFlow. We used Google Colab and used python coding to implement the proposed work.

4.1. Performance Evaluators

To asses the efficacy of the proposed work we calculate Precision, Recall, F-measure, and Accuracy metrics for the model. The precision value is calculated in (2),

$$\text{Prec} = \frac{tp}{tp+fp} \times 100 \tag{2}$$

True and false positive and negative values are recorded by tp, fp, tn and fn respectively. The recall evaluator is calculated in (3) as,

$$\text{Recall} = \frac{tp}{tp+fn} \times 100 \tag{3}$$

The correctness of the classifier can also be evaluated using F-Measure score. F-Measure is calculated as in (4)

$$\text{F} - \text{Measure} = 2 \times \frac{[\text{Precision} \times \text{Recall}]}{[\text{Precision} + \text{Recall}]} \tag{4}$$

The accuracy is calculated as in (5),

$$\text{Acc} = \frac{tp+tn}{tp+fn+fp+tn} \times 100 \tag{5}$$

The hyper-parameters for the experiment were selected heuristically and are as follows: Input Image Size is 224×224; each batch is of 15 inputs; optimizer used is Adam optimizer, 30 epochs are considered with a learning rate of 1e-3. Table 2 specifies the performance evaluators for modified hybrid models using different base models.

Table 2. Classification Results for Modified Models using Pre-Trained Models

DL-CNN Model	Prec	Recall	F-Measure	Accuracy
MobileNet	81.65	86.12	83.80	76.47
InceptionNet	80.60	89.10	85.42	78.55
Xception	78.65	90.32	84.23	76.47
DenseNet201	83.65	93.80	89.12	84.70

Fig. 9 to Fig.12 represents loss and accuracy graphs for the modified InceptionNet-V3, MobileNetV3, Xception, and DenseNet201 DL-CNN models. The four selected models used for diagnosis have achieved an accuracy of 76.47%, 78.55%, 76.47%, and 84.70% respectively. The denseNet201 model has achieved the highest accuracy of 84.70% and is selected as the base model.
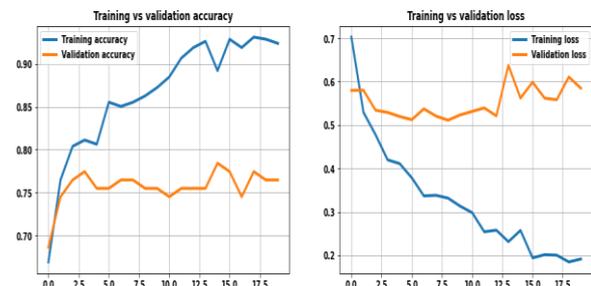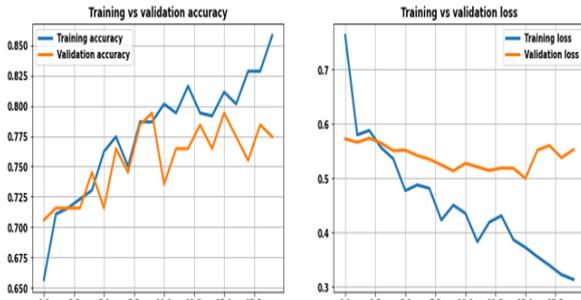

Fig. 9. Accuracy and Loss graphs of MobileNet

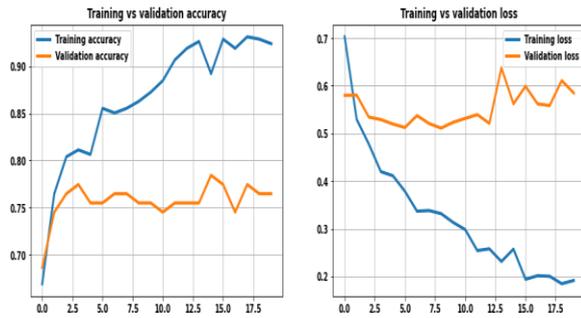**Fig. 10.** Accuracy and Loss graphs of InceptionNetV3
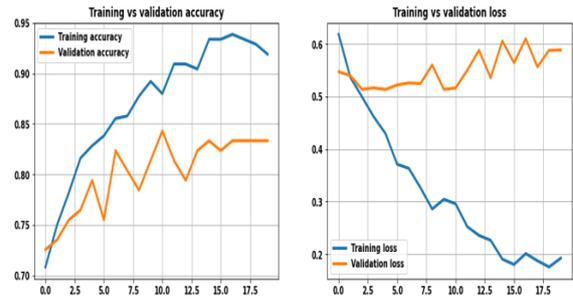

Fig. 12. Accuracy and Loss graphs of DenseNet201

Fig. 11. Accuracy and Loss graphs for Xception

It can also be noted that other performance evaluators are surprisingly good with 83.65%, 93.80%, and 89.12% of Precision, Recall, and F-measure with DenseNet201 model. The hybrid model is then built

using DenseNet201 as the base model. The addition of extra layers in the model improves the overall efficiency of the proposed model.

Fig. 13 represents the confusion matrices of various base models. The matrix depicts the mapping of true values to the predicted values. A true-true label shows a cancerous tissue predicted as cancerous and a non-cancerous tissue predicted as non-cancerous. The confusion matrix provides a visual analysis of the proposed model for successful classification of the images.
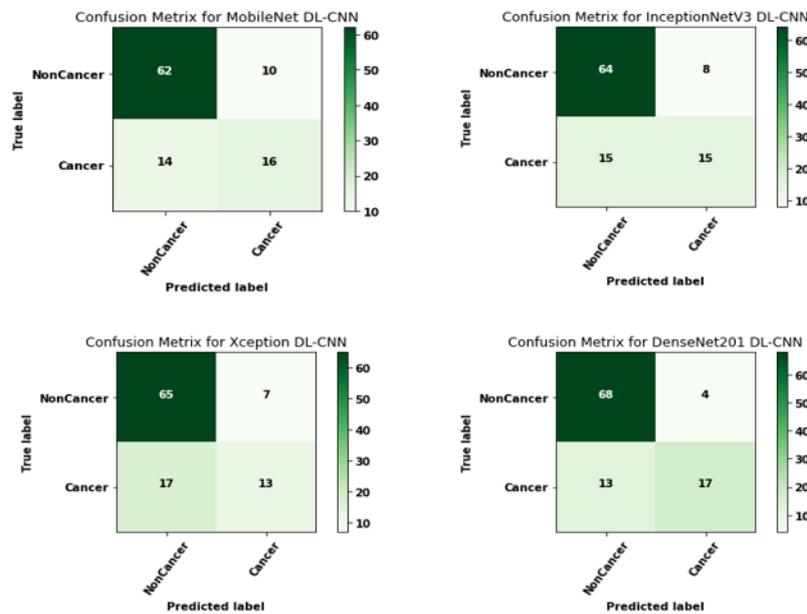

Fig. 13. Confusion Matrix for Different Models

Fig. 14 depicts the detection results of a few oral photographic images which have both cancerous and noncancerous patients' images. Value 0 represents a cancerous image and value 1 represents a non-cancerous. Thus Detection=0 and True=0 represent an image detected as cancerous and a cancerous image. Similarly, Detection=1 and True=1 represent images detected as non-cancerous and non-cancerous images.
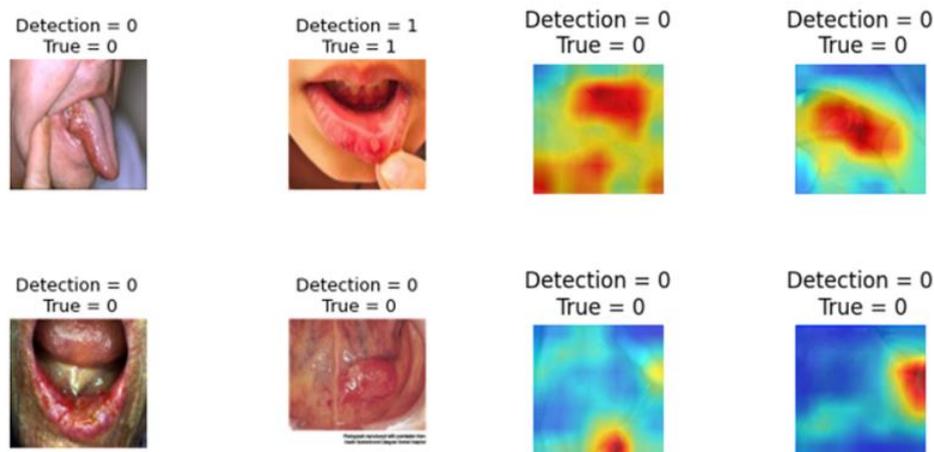
Fig. 14. Detection Results from photographic images (left) and attention maps (right)

4.2. Proposed Work in comparison to Other remarkable Works

Table 3 shows other similar works with the proposed work. Compared to the above-mentioned works, we have considered a relatively large dataset of 1000 oral photographic images for Oral Cancer detection and achieved better performance. However, research works using different datasets come with varied challenges and cannot be compared directly.

Welikala et al. [7] showcased a hybrid model for oral cancer detection. They utilized ResNet model, and achieved 87.07% F1-score for 2155 image dataset whereas our F1 score is 89.12% for the 1000 image dataset. Shamim et al. [8] conducted a study focusing on photographic images of tongue tissues and achieved an accuracy of 98% using the VGG19 model. However, the sensitivity was relatively lower at 89% for 200 images which is less compared to our proposed work. Huiping Lin et al. [10] applied HRNet on 455 images and achieved 83.60% accuracy. Baichen Ding [11] utilized YOLOV3 on 570 images and reported 76.90% accuracy.

Table 3. Comparison with Notable Work on Oral Photographic Dataset

Author(s)	DL-CCN Model(s) Applied	Total No. of Images	Performance Parameters
Roshan et al. [7]	ResNet101	2155	87.0% (F1-score)
Shamim et al. [8]	VGG19	200	97.0% (Accuracy)
Huiping Lin et al. [10]	HRNet	455	83.6% (Accuracy)
Baichen Ding [11]	YOLOV3	570	76.9% (Accuracy)
Song et al. [12]	VGG16	130	86.9% (Accuracy)
Jubair et al. [14]	EfficientNet	716	85.0% (Accuracy)
Nanditha et al. [15]	ResNet50 & VGG16	332	95.0% (Accuracy)
Proposed Work	DenseNet201	1000	84.7% (Accuracy)

Song et al. [12], a smartphone-based intra-oral dual-modality imaging platform was employed. With a dataset consisting of 66 normal images and 64 suspicious images, they achieved 86.90% accuracy using VGG16 model. It is worth noting that the achieved accuracy was higher due to the smaller number of images used. Jubair et al. [14] utilized EfficientNet-B0 as base model on 716 images. The authors reported 85.00% accuracy. Nanditha et al. [15] employed a combination of ResNet and VGG16 models and achieved 95.10% accuracy using 332 images. The images considered are fewer compared to our proposed work.

5. Conclusion

The proposed research work aims at developing a DL-based approach for detecting and classifying oral cancer (OC) from oral photographic images. To achieve this, the work employed pre-trained deep learning models such as MobileNet, InceptionV3, Xception, and DenseNet201. These models were customized by adding extra layers to improve the accuracy of OC detection. The experiment aimed to identify the best pre-trained DL-CNN model, modified with our proposed CNN layers, to achieve the highest classification accuracy for OC in oral photographic images. Furthermore, our proposed work explored the potential of using raw oral photographic images for OC classification through the DL-CNN technique. Among the various models tested, the modified DenseNet201 DL-CNN model demonstrated superior performance metrics. We successfully classified OC from photographic images, indicating the potential for implementing DL approaches for oral cancer diagnosis.

References

[1] Abati, S.; Bramati, C.; Bondi, S.; Lissoni, A.; Trimarchi, M., "Oral Cancer and Precancer: A Narrative Review on the Relevance of Early Diagnosis," Int. J. Environ. Res. Public Health 2020, 17, 9160. https://doi.org/10.3390/ijerph17249160

[2] Jeyaraj, Pandia Rajan, and Edward Rajan Samuel

Nadar. "Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm," Journal of cancer research and clinical oncology 145 (2019): 829-837.

[3] Fu, Qiuyun, et al. "A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study," EClinical Medicine 27 (2020): 100558.

[4] Mortazavi, Hamed, Maryam Baharvand, and Masoumeh Mehdipour. "Oral potentially malignant disorders: an overview of more than 20 entities," Journal of dental research, dental clinics, dental prospects 8.1 (2014): 6.

[5] Warnakulasuriya, Saman. "Oral potentially malignant disorders: A comprehensive review on clinical aspects and management.," Oral oncology 102 (2020): 104550.

[6] Palaskar, Rutwik, et al. "Transfer learning for oral cancer detection using microscopic images," arXiv preprint arXiv:2011.11610 (2020).

[7] Roshan Alex Welikala, Paolo Remagnino, Jian Han Lim, Chee Seng Chan, Senthilmani Rajendran, Thomas George Kallarakkal, Rosnah Binti Zain, Ruwan Duminda Jayasinghe, "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer," IEEE Access, vol. 8, pp. 132677-132693, 2020. https://doi.org/10.1109/access.2020.3010180.

[8] Mohammed Zubair M Shamim, Sadatullah Syed, Mohammad Shiblee, Mohammed Usman, Syed Jaffar Ali, Hany S Hussein, Mohammed Farrag, "Automated Detection of Oral Pre-Cancerous Tongue Lesions Using Deep Learning for Early Diagnosis of Oral Cavity Cancer," The Computer Journal, Volume 65, Issue 1, January 2022, Pages 91–104, https://doi.org/10.1093/comjnl/bxaa136.

[9] Figueroa, Kevin Chew, et al. "Interpretable deep learning approach for oral cancer classification using guided attention inference network," Journal of biomedical optics 27.1 (2022): 015001-015001.

[10] 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.

[11] Ding, Baichen, et al. "Detection of dental caries in oral photographs taken by mobile phones based on the YOLOv3 algorithm," Annals of Translational Medicine 9.21 (2021).

[12] Song et al., "Automatic classification of dual-modality, smartphone-based oral dysplasia and malignancy images using deep learning," Biomed. Opt. Express, vol. 9, no. 11, pp. 5318–5329, 2018.

[13] Tanriver, G.; Soluk Tekkesin, M.; Ergen, O. "Automated Detection and Classification of Oral Lesions Using Deep Learning to Detect Oral Potentially Malignant Disorders," Cancers 2021, 13, 2766. https://doi.org/10.3390/cancers13112766

[14] 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. (2021) doi: 10.1111/odi.13825.

[15] Nanditha, B. R., et al. "An ensemble deep neural network approach for oral cancer screening," (2021): 121-134.

[16] Begum, Sayyada Hajera, and P. Vidyullatha. "Deep Learning Model for Automatic Detection of Oral squamous cell carcinoma (OSCC) using Histopathological Images," International Journal of Computing and Digital Systems, vol.13, no.1, (Apr-23)

[17] Chan, Chih-Hung, et al. "Texture-map-based branch-collaborative network for oral cancer detection," IEEE transactions on biomedical circuits and systems 13.4 (2019): 766-780.

[18] M. A. Muqeet, A. B. Mohammad, P. G. Krishna, S. H. Begum, S. Qadeer and N. Begum, "Automated Oral Cancer Detection using Deep Learning-based Technique," 2022 8th International Conference on Signal Processing and Communication (ICSC), Noida, India, 2022, pp. 294-297, doi: 10.1109/ICSC56524.2022.10009448.

[19] Begum, Sayyada Hajera, et al. "A Lightweight Deep Learning Model for Automatic Diagnosis of Lung Cancer," 2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC). IEEE, 2022.

[20] https://www.kaggle.com/shivam17299/oral-cancer-lips-and-tongue-images

[21] Martín Abadi, et al., TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[22] Krizhevsky, I. Sutskever, G.E. Hinton, "ImageNet classification with deep convolutional neural networks," Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.

[23] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510–4520.

[24] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, "Rethinking the inception architecture for computer vision," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, 2818–2826.

[25] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510–4520.

[26] Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q., "Densely Connected Convolutional Networks," Proceedings of the IEEE Conference on

Computer Vision and Pattern Recognition, Honolulu, 21-26, July 2017, 4700-4708.

[27] Kumar, B. P. ., Reddy, N. H. ., Das, R. R. ., Dey, D. ., Sona, D. R. ., & M, A. . (2023). Encryption and Decryption of Images with Pixel Data Modification Using Hand Gesture Passcodes. International Journal on Recent and Innovation Trends in Computing and Communication, 11(4), 118–125. https://doi.org/10.17762/ijritcc.v11i4.6394

[28] Mwangi, J., Cohen, D., Costa, R., Min-ji, K., & Suzuki, H. Optimizing Neural Network Architecture for Time Series Forecasting. Kuwait Journal of Machine Learning, 1(3). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/132

[29] Sharma, R., & Dhabliya, D. (2019). A review of automatic irrigation system through IoT. International Journal of Control and Automation, 12(6 Special Issue), 24-29. Retrieved from www.scopus.com