IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125167

ORALSCREEN-ORAL CANCER DETECTION USING DEEP LEARNING

Prof Ashwini D S[1],

Amithashree H R[2], Charan M[3], Venkatesh R S[4]

Assistant Professor, Electronics and Communication, NIE Institute of Technology, Mysore, India[1]

Student, Electronics and Communication, NIE Institute of Technology, Mysore, India[2-4]

Abstract: Oral cancer is a severe public health problem with a high fatality rate, and early identification is critical for increasing survival rates. We propose a deep learning-based technique for detecting oral cancer utilizing two types of oral cavity pictures: normal oral images and histopathologic images in this work. We trained our models on a huge dataset of oral cavity pictures using two cutting-edge convolutional neural network (CNN) models, ResNet50 and VGG16, using transfer learning. According to our findings, ResNet50 had an accuracy of 95% and VGG16 had an accuracy of 94% in identifying oral cavity pictures as cancerous or non-cancerous.

We have integrated our model into OralScreen, an online tool that can be utilized by both patients and medical experts such as physicians and histopathologists. Our findings show that deep learning-based techniques have the potential to increase the accuracy and efficiency of oral cancer diagnosis dramatically.

Keywords: Deep Learning, CNN, ResNet50, VGG16.

I. INTRODUCTION

With over 350,000 new cases identified each year, global cancer is a major public health problem. Early identification is critical for better patient outcomes and for lowering morbidity and death rates linked with this condition. Traditional techniques of diagnosing oral cancer, such as visual inspection, biopsy, and histopathology analysis, are time-consuming, invasive, and need the use of experienced specialists. Deep learning algorithms, on the other hand, such as convolutional neural networks (CNNs), can learn subtle patterns and characteristics from large-scale datasets, making them useful for medical picture analysis. Deep learning-based techniques for oral cancer diagnosis have recently been developed, using CNNs to analyse pictures of the oral cavity and identify probable malignant lesions.

VGG16 and ResNet50 are popular deep-learning architectures that have been used among these methodologies.VGG16 is a 16-layer CNN architecture that has produced state-of-the-art performance on image classification tasks, whereas ResNet50 is a 50-layer CNN architecture that was created to solve the issue of disappearing gradients in deep neural networks. In this study, we used a dataset of normal oral cavity images and histopathologic images to train a deep learning model for detecting oral cancer. We trained our model using the ResNet50 and VGG16 methods and compared their performance.

The ResNet50 algorithm obtained a better accuracy of 95% than the VGG16 method, which achieved a lower accuracy of 94%. As a result, we chose the ResNet50 algorithm as the superior model for our online application, Oral screen. It is worth noting that our model was trained in a single model utilizing both types of data, normal oral images, and histopathologic images. Our findings show that deep learning-based techniques have a high potential for enhancing the accuracy and efficiency of oral cancer diagnosis, which can lead to improved outcomes.

II. METHODOLOGY

Image acquisition plays a crucial role in the development of deep learning models for oral cancer detection. The quality and type of images used for training, validation, and testing of the models can significantly impact the accuracy of the model. Accurate medical imaging is essential for correctly diagnosing medical conditions as incorrect or incomplete information from medical images can lead to misdiagnosis, inappropriate treatment, and potential harm to patients. The dataset considered involves four classes of images namely normal oral image, cancer oral image, normal histopathologic image, and OSCC histopathologic image having image counts of 44,87,2435 and 2511 respectively.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ※ Impact Factor 8.102 ※ Peer-reviewed / Refereed journal ※ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125167

In the next stage, the Image augmentation technique is used to artificially increase the size of a dataset by creating new versions of the original images with modifications. This helps improve the performance of machine learning models by reducing overfitting and improving generalization. In the context of oral cancer detection, image augmentation techniques can be applied to both cancer and non-cancer oral images. In our study, we used two types of image augmentations: vertical flip and noise.

Vertical flip augmentation involves flipping the images vertically, which can help the model learn to recognize cancerous lesions from different angles. This is particularly useful when training the model on a limited dataset of images taken from a single angle. The noise augmentation involves adding random noise to the images, simulating the variations that may occur in real-world images. This helps the model learn to distinguish between true cancerous lesions and other variations that may look similar to cancerous lesions but are not. The total image count after augmentation is 5522.

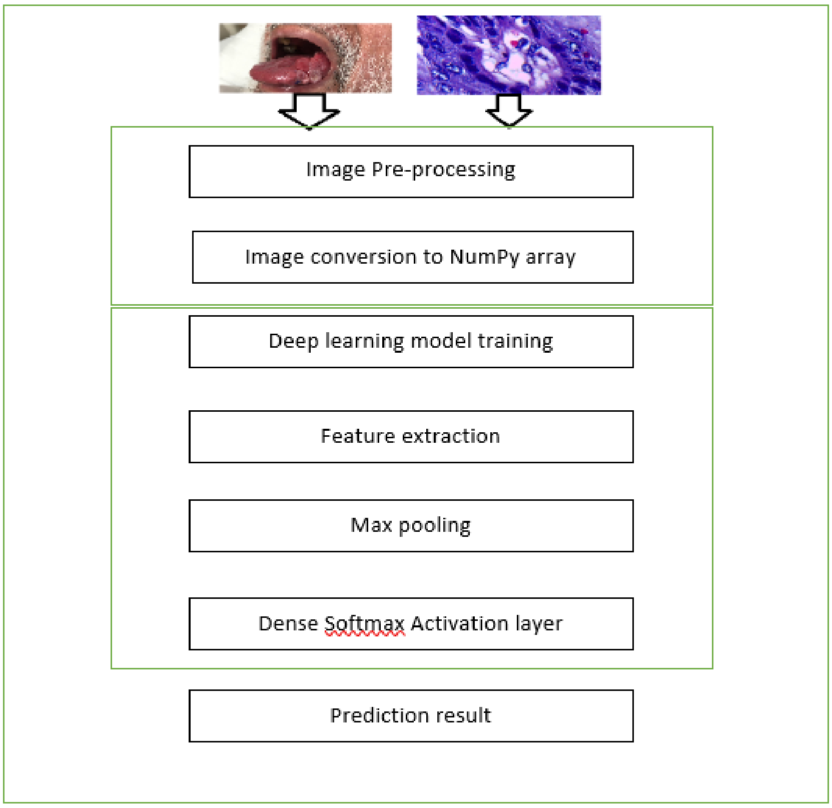


Fig.1 block diagram

Converting an image into a NumPy array involves reading the image file and then using the NumPy library to represent the image as an array. The first step is to load the image data from a folder using the from_folder function in the ImageDataLoaders class of the fastai library. Next, the dataset is split into a training set and a validation set using the training parameter, and a validation percentage of 20% is specified with valid_pct=0.2. A random seed of 66 is also set to ensure reproducibility.

After the data is split, various image transformation techniques are applied to the images. The item_tfms parameter specifies the image transformations to be applied to each image before it is loaded into the model. In this case, a RandomResizedCrop transformation is used to randomly crop and resize the images to a size of 224 pixels. The min_scale parameter specifies the minimum scaling factor for the crop, which is set to 0.5 in this case.
Finally, the Normalize.from_stats(*imagenet_stats) function is used to normalize the pixel values of the images based on the mean and standard deviation of the pixels in the ImageNet dataset. Overall, these transformations help to improve the accuracy and performance of the deep learning model in image classification tasks.

A convolutional neural network (CNN) learner is initialized using the fastai library. The cnn_learner() function takes three arguments: the data, the pre-trained model to use (in this case, VGG16 with batch normalization and ResNet50), and a list of metrics to use during training. The data argument is the output of the ImageDataLoaders function that loads the data from the specified directory and applies data augmentation and normalization. The models module from the

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁙ Impact Factor 8.102 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125167

PyTorch library provides pre-trained CNN models, and VGG16 is one such model that is commonly used for image classification tasks. By passing models.vgg16_bn and resnet50 as the second argument, the code initializes a pre-trained VGG16 model with batch normalization layers and the ResNet50 model. Finally, the metrics argument specifies the evaluation metrics to use during training, such as accuracy and error rate. Overall, a CNN learner is initialized with a pre-trained VGG16 and ResNet50 models and prepares for training on the given data.

Confusion matrix is a popular technique used to evaluate the performance of a deep learning model. It is a table that summarizes the predictions made by a model and compares them to the actual values. The matrix is organized into four sections, representing the true positive, true negative, false positive, and false negative values.

In the context of medical image analysis, a confusion matrix can be used to evaluate the performance of a deep learning model for oral cancer detection. After training the model, a set of test images is used to validate the model's performance. The model's predictions are compared to the actual cancer status of the images, and a confusion matrix is created.

In deep learning, PKL files are commonly used to store trained models and their associated weights, as well as other information such as the model architecture, optimizer parameters, and training history. These files can be loaded into a Python environment and used to make predictions on new data and fine-tune with additional training on new data.

ORALSCREEN, a web-based application was developed using the Flask framework for the deep learning model trained on oral cavity images. Flask is a lightweight web application framework written in Python. It is designed to make it easy to build web applications with Python by providing useful abstractions and simplifying the process of handling HTTP requests and responses.

Flask was used to create a web interface for users to upload an image of their oral cavity and receive a prediction from the model on whether it contains cancerous cells or not. The Flask application would receive the image from the user, process it, and then pass it to the deep learning model for prediction. The prediction result would then be returned to the user via the web interface.

III. RESULTS ANALYSIS

An epoch refers to one complete pass through the entire training dataset during the training process of a machine learning model. During one epoch, the model uses each training example once and updates the weights of the model's parameters to optimize the performance of the model.

The number of epochs required for training depends on the complexity of the problem, the size of the dataset, and the model's architecture. Typically, multiple epochs are required to achieve the desired level of accuracy in the trained model. For each epoch, the values for the training loss, validation loss, accuracy, error rate, and time taken for the epoch are provided.

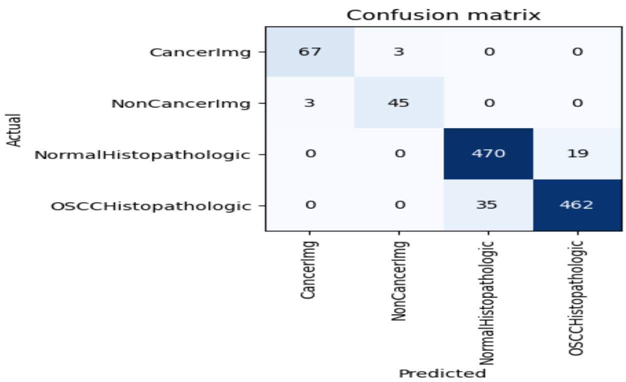
epoch	train_loss	valid_loss	accuracy	error_rate	time
0	1.172860	0.432098	0.846014	0.153986	08:10
1	0.623861	0.275905	0.903986	0.096014	02:33
2	0.414110	0.208271	0.923007	0.076993	02:27
3	0.315262	0.162475	0.933877	0.066123	02:31
4	0.248565	0.157909	0.941123	0.058877	02:30
5	0.198151	0.139684	0.945652	0.054348	02:34

Fig.2 Performance analysis of VGG14.

epoch	train_loss	valid_loss	accuracy	error_rate	time
0	0.981382	0.447504	0.857790	0.142210	02:21
1	0.503516	0.219068	0.923007	0.076993	02:19
2	0.338690	0.159139	0.940217	0.059783	02:20
3	0.213638	0.148576	0.948370	0.051630	02:16
4	0.156651	0.112683	0.960145	0.039855	02:20
5	0.135313	0.107459	0.956522	0.043478	02:20

Fig.3 Performance analysis of ResNet50.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125167



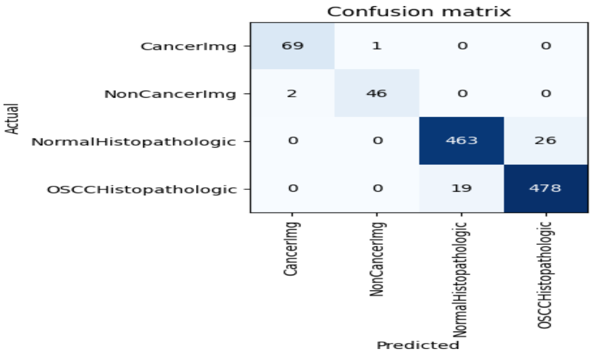


Fig.4 Confusion matrix of VGG16

Fig.5 Confusion matrix of ResNet50

The training loss and validation loss indicate how well the model is fitting the training data and the validation data respectively. A lower loss value suggests better performance. The accuracy indicates how well the model is classifying the images into their respective classes. The error rate indicates the proportion of misclassified images. Finally, the time taken for each epoch is also provided. This is useful for tracking the efficiency of the training process.

The final accuracy for VGG16 AND ResNet50 are obtained as 94.56 and 95.65 with error rates being 0.05 and 0.04 respectively. The confusion matrix analysis of VGG16 and ResNet50 shows that the prediction is better with the latter having less number of false positive and false negative results which regards Resnet50 as the better classifier model.

IV. CONCLUSION

In this paper, the project ORALSCREEN aimed to develop an automated system for oral cancer screening using deep learning models. We evaluated the performance of two popular models, VGG16 and ResNet50, and found that ResNet50 outperformed VGG16 in terms of accuracy and error rate. The confusion matrix also showed that ResNet50 had a higher true positive rate and lower false positive rate for both cancer and non-cancer classes, indicating better classification performance. Therefore, ResNet50 was selected as the model of choice for web-based application development.

The application developed using the Flask framework allows users to upload oral images and get real-time predictions for cancer and non-cancer cases. In the future, we plan to further improve the accuracy of our model by incorporating more advanced deep learning techniques such as transfer learning and ensembling. We also hope to expand our dataset to include more diverse and representative samples to ensure the robustness of our model across different populations.

Overall, our study demonstrates the potential of deep learning models in the field of oral cancer screening and provides a foundation for future research in this area. a deep-learning oral cancer detection model has been discussed. That collects oral cavity images anddemonstrates results for automating the early detection of oral cancer.

REFERENCES

[1]. ‘‘Recent advances in convolutional neural networks,’’ J. Gu, W. Zhenhua, K. Jason, L. Ma, A.Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, and T. Chen, 2019
[2]. ‘‘Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,’’ H.-C. Shin, H. R. Roth, M. Gao, L.Lu, Z.Xu, I.Nogues, J.Yao, R.M. Summers, and D. Mollura,2016
[3]. ‘‘Translated learning: Transfer learning across different feature spaces,’’ W. Dai, Y.Chen, G. Xue,Q. Yang, and Y. Yu,2018
[4]. “Oral Cancer Detection using Machine Learning and Deep Learning Techniques” Nanditha B R,Geetha Kiran A Sanathkumar M P,2022
[5]. ‘‘Improving computer-aided detection using convolutional neural networks and random view aggregation,’’ H. R. Roth, L. Lu, J. Liu, J.Yao, A. Seff, K.Cherry, L.Kim, and R. M. Summers,2016
[6]. “Optical coherence tomography (oct) and chest X-ray images for classification” Kermany D, Goldbaum M.Labeled.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁑ Impact Factor 8.102 ⁑ Peer-reviewed / Refereed journal ⁑ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125167

In: Mendeley Data. 2018
[7]. “Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer”. Welikala RA. IEEE Access, vol. 8, 132677-132693, 2020
[8]. “Detection of Oral Cancer Using Machine Learning Classification Methods” Prabhakaran R, Mohana J. Int J of Elect Eng and Tech 2020.
[9]. “Detection of Oral Cancer Using Deep Neural Based Adaptive Fuzzy System in Data Mining Techniques” K Lalithamani, A Punitha .International Journal of Recent Technology and Engineering (IJRTE), issue 7, p. 2277–3878 , 2019
[10]." Oral Epithelial Dysplasia ComputerAided Diagnostic Approach", David Adel, John Mounir, Mahmoud El Shafey, IEEE, 2018.
[11]." Cureth Alert System to Detect Oral Cancer", D.Padmini Pragna, Sahithi Dandu, Meenakzshi M, C. Jyotsna, Amudha J, International Conference on Inventive Communication and Computational Technologies, 2017.