구강암 조기발견을 위한 영상인식 시스템

에드워드 카야디, 송미화 [1]
[1] 세명대학교 스마트 IT 학부
e-mail : edw.chydi@gmail.com, mhsong@semyung.ac.kr

Image Recognition System for Early Detection of Oral Cancer

Edward Dwijayanto Cahyadi, Mi-Hwa Song[1]
[1]School of Smart IT , Semyung University

Summary

Oral cancer is a type of cancer that has a high possibility to be cured if it is threatened earlier. The convolutional neural network is very popular for being a good algorithm for image recognition. In this research, we try to compare 4 different architectures of the CNN algorithm: Convnet, VGG16, Inception V3, and Resnet. As we compared those 4 architectures we found that VGG16 and Resnet model has better performance with an 85.35% accuracy rate compared to the other 3 architectures. In the future, we are sure that image recognition can be more developed to identify oral cancer earlier.

1. Introduction

Oral cancer forms in tissues of the oral cavity (the mouth) or the oropharynx (the part of the throat at the back of the mouth)(National institute of cancer by usa.gov). This type of cancer has a more significant life expectancy than other cancer if it is threatened earlier. Image recognition is a subcategory of Computer Vision and Artificial Intelligence, representing a set of methods for detecting and analyzing images to enable the automation of a specific task. Using this technology, we believe that it can detect any early symptoms of oral cancer in the future.

2. Related Research

According to the "Oral cancer prognosis based on clinicopathologic and genomic markers using a hybrid of feature selection and machine learning methods." (Chang et al.2013) paper that was published in 2013, the writer compares the result with several methods such as artificial neural network, adaptive neuro-fuzzy inference system, support vector machine, and logistic regression also they were making the hybrid model of *ReliefF-GA-ANFIS* with 3-input features of drink, invasion and p63. That model achieved the best accuracy (accuracy = 93.81%; AUC = 0.90) for the oral cancer prognosis. In the paper, it has been explained that some data is divided into two; the first one was clinicopathologic variables, and the second one is variable of protein types of p53 and p63. After collection, the data feature selection methods were done to make less data noise; after that, the data was classified with the k-fold cross-validation to compare the machine learning methods in order to get the best result. In the study called "Improving Oral Cancer Outcomes with Imaging and Artificial Intelligence" (B.Ilhan.2020) it has been described that to improve the performance of the AI, three main steps need to be done: preprocessing, image segmentation, and post-processing. Preprocessing needs to be done to remove any unwanted image information contained in the raw images. Image segmentation is executed to improve the accuracy of the AI model. Lastly, the post-processing method is a process to improve the result of the model also to minimize error while avoiding overfitting the data. Convolutional neural networks are a class of artificial neural networks, most commonly applied to analyze visual images. VGG16 is a simple and widely used Convolutional Neural Network (CNN) Architecture used for ImageNet, a large visual database project used in visual object recognition software research. The VGG16 Architecture was developed and introduced by Karen Simonyan and Andrew Zisserman from the University of Oxford, in the year 2014, through their article "Very Deep Convolutional Networks for Large-Scale Image Recognition. Inception v3 is a convolutional neural network to aid in image analysis and object detection and started as a module for Googlenet. ResNet, short for Residual Network, is a specific neural network introduced in 2015 by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun in their paper "Deep Residual Learning for Image Recognition".

3. Research Method

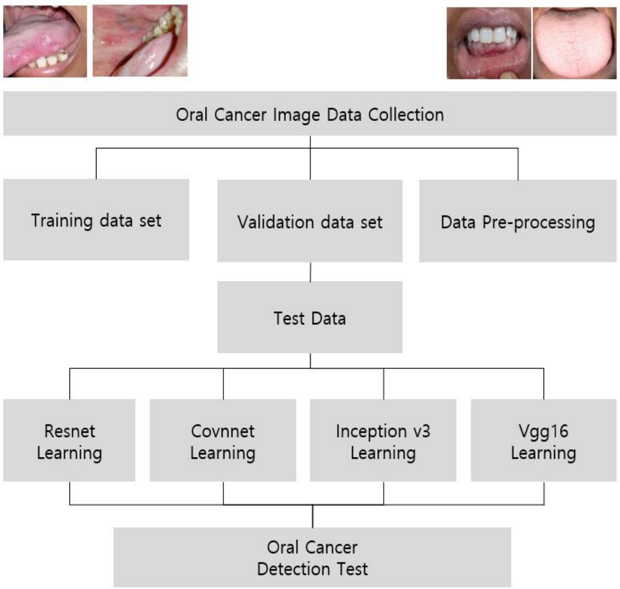

Figure 1. Proposed System Structure

This research uses 488 images divided into two categories, 263 cancer images, and 225 non-cancer images obtained from Kaggle and Mendeley data, as the datasets. To create the machine learning model, we use Python as the programming language and Google Colaboratory as the compiler. According to the previous study, the CNN algorithm performs well in classifying oral cancer images. Based on that, we tried to make a CNN machine learning model with four different CNN architectures: Convnet, VGG16, Inception-v3, and Resnet.

As we can see in the proposed system structure (figure 1), Before building the model, we have done some data pre-processing. First, we use a function to remove duplicate images in the datasets to prevent model overfitting. Then, we split the data into three parts. 70% is for training data sets,15% is for the validation data, and the last 15% is for the testing data. After that, we imported some libraries from Keras for the CNN model training, such as dense, conv2d, flatten, and max pool. Then, we train the data using 4 different architectures of CNN with 32 epochs of each training. In figure 2, we printed out the result of the training the variables of this result include accuracy, f1 score, and loss of the 32 epochs training. As we take a look at the result (figure 2) the VGG16 and the ResNet model has the best performance in maintaining the accuracy compared to other models. On the other hand, Convnet model has the worse performance in maintaining both accuracy and loss rate. And for the f1 score, the Resnet model has the most balance performance compared to other models. On another hand, Convnet model has the worse outcome. Lastly, VGG16 and Inception V3 has similar outcome in term of the f1 score.

	acc_cnn	acc_vgg16	acc_inception v3	acc_resnet	loss_cnn	loss_vgg	loss_incep	loss_resnet	f1_cnn	f1_vgg16	f1_inception v3	f1_resnet
0	0.578125	0.571429	0.476190	0.547619	3.074385	1.593815	1.046796e+02	2.924853	0.572158	0.488388	0.359895	0.305447
1	0.428571	0.484375	0.666667	0.547619	0.907744	0.966802	1.184719e+02	6.709031	0.069231	0.520338	0.537618	0.559685
2	0.595238	0.690476	0.619048	0.500000	0.681336	0.542018	1.166867e+02	1.997127	0.355363	0.552956	0.587287	0.738066
3	0.595238	0.690476	0.857143	0.452381	0.670444	0.519188	3.807380e+01	2.829341	0.379999	0.704627	0.621074	0.721621
4	0.500000	0.833333	0.857143	0.523810	0.702371	0.454966	1.040901e+01	5.468537	0.429296	0.648806	0.565526	0.772729
5	0.571429	0.738095	0.880952	0.578125	0.687270	0.504426	6.105373e+00	2.307843	0.677926	0.676407	0.692520	0.828291
6	0.642857	0.833333	0.952381	0.406250	0.682687	0.375821	2.954462e+00	4.419594	0.285365	0.736209	0.553533	0.778082
7	0.619048	0.880952	0.833333	0.619048	0.683799	0.202720	9.494950e+00	1.067783	0.554915	0.829863	0.608816	0.886453
8	0.619048	0.880952	0.904762	0.666667	0.680299	0.317481	1.789269e+01	2.102859	0.450435	0.767968	0.677913	0.879907
9	0.593750	0.904762	0.952381	0.523810	0.682080	0.233261	1.149012e-01	1.840628	0.584325	0.814874	0.736233	0.950128
10	0.642857	0.928571	0.904762	0.571429	0.652992	0.167574	2.152892e+01	1.265978	0.313867	0.864897	0.713034	0.906282
11	0.571429	0.952381	0.904762	0.562500	0.687073	0.175607	1.056767e+01	1.805073	0.508837	0.845462	0.617360	0.848567
12	0.666667	0.928571	0.952381	0.595238	0.639387	0.194921	8.643863e+00	1.201043	0.470380	0.762037	0.727502	0.892649
13	0.593750	0.906250	0.928571	0.476190	0.662789	0.208336	9.350589e+00	1.683697	0.615030	0.876837	0.661596	0.878354
14	0.625000	0.906250	0.452381	0.531250	0.648437	0.205301	1.362354e+02	1.686324	0.543453	0.872756	0.723422	0.858141
15	0.593750	0.952381	0.921875	0.595238	0.671936	0.131106	9.415502e+00	1.032451	0.599806	0.862062	0.758296	0.881535
16	0.625000	0.976190	1.000000	0.619048	0.637515	0.101578	4.981262e-22	1.130662	0.593258	0.875552	0.575925	0.888809
17	0.547619	0.952381	0.833333	0.547619	0.677979	0.106675	2.032880e+01	1.527926	0.683719	0.782993	0.745175	0.871376
18	0.714286	0.952381	0.976190	0.578125	0.568312	0.154992	2.510534e+00	1.867457	0.681518	0.836597	0.738416	0.919775
19	0.547619	0.952381	0.968750	0.642857	0.679090	0.104557	2.675214e+00	1.174635	0.682922	0.893265	0.825472	0.794808
20	0.578125	0.952381	0.904762	0.619048	0.623613	0.117422	5.722240e+00	1.372234	0.649810	0.863395	0.748559	0.887420
21	0.571429	0.953125	0.928571	0.671875	0.602206	0.163542	1.574038e+00	1.345791	0.566449	0.881244	0.735480	0.936096
22	0.562500	0.952381	1.000000	0.562500	0.606491	0.105717	0.000000e+00	1.224918	0.731345	0.873612	0.763172	0.921128
23	0.796875	0.976190	0.976190	0.593750	0.559169	0.089152	1.394804e+00	1.302573	0.703284	0.907425	0.718490	0.904401
24	0.640625	0.953125	0.952381	0.619048	0.546753	0.130279	4.721954e+00	1.122477	0.689740	0.869056	0.768551	0.841302
25	0.734375	0.952381	0.833333	0.578125	0.573352	0.147781	3.036059e+01	1.605449	0.664075	0.894673	0.684992	0.958040
26	0.738095	0.968750	1.000000	0.619048	0.548760	0.104264	7.882031e-12	1.489281	0.641527	0.904016	0.735031	0.840429
27	0.785714	0.928571	0.976190	0.690476	0.503960	0.127827	6.143348e+00	1.878292	0.679790	0.913460	0.919162	0.944381
28	0.880952	0.968750	1.000000	0.500000	0.422177	0.105839	1.449904e-34	2.803251	0.707861	0.906278	0.744033	0.948123
29	0.809524	1.000000	0.976190	0.500000	0.463091	0.040480	1.565860e+00	3.017413	0.777239	0.891781	0.843455	0.970103
30	0.859375	1.000000	1.000000	0.619048	0.413928	0.060457	1.865987e-08	3.412566	0.731224	0.928290	0.854461	0.893031
31	0.714286	0.984375	1.000000	0.714286	0.505623	0.089137	3.442866e-12	2.691212	0.742450	0.898466	0.815311	0.929060

Figure 2. Accuracy, F1 score, and loss of 32 epoch

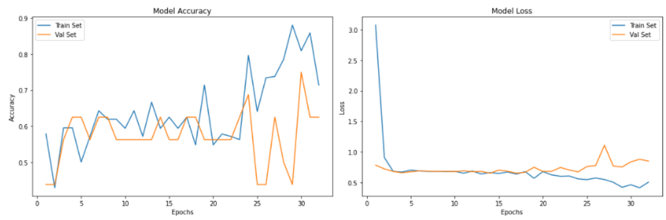

Figure 3. Accuracy and loss graph of convnet

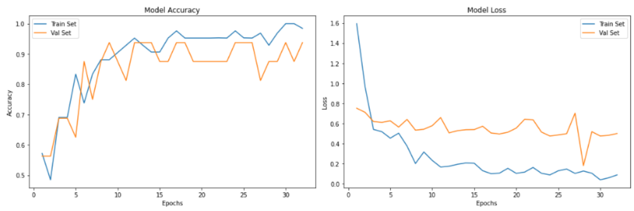

Figure 4. Accuracy and loss graph of VGG16

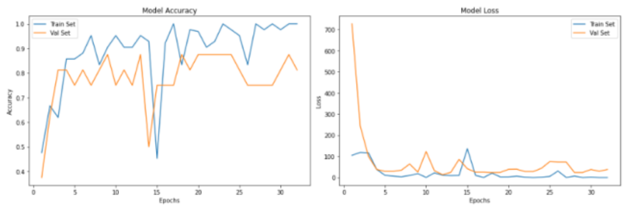

Figure 5. Accuracy and loss graph Inception v3

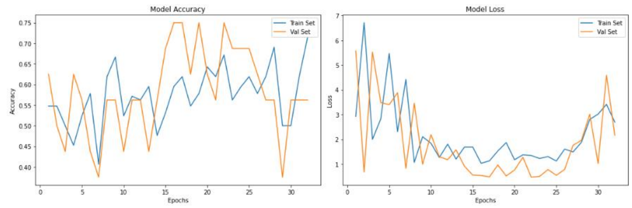

Figure 6. Accuracy and loss graph of Resnet

As we can see in figure 3 the accuracy gap between the training and validation data is quite big compare to the loss rate. For the VGG 16 model (figure 4) the pattern of validation and training data set's accuracy are similar but on the other hand the loss rate are quite different. Inception V3 (figure 5) has a good result in term of maintaining the gap between training and validation data set's accuracy and loss rate. Lastly the Resnet model (figure 6) has a good performance on balancing loss rate but in term of accuracy rate it has a big gap while running the 15 until 20 epoch. Although we get a good result on the training test, there is some problem that makes the model can't do a perfect classification, such as the datasets that we are using are far from enough to teach the model, in the result the model sometimes overfit and do a bad performance while identifying the validation data

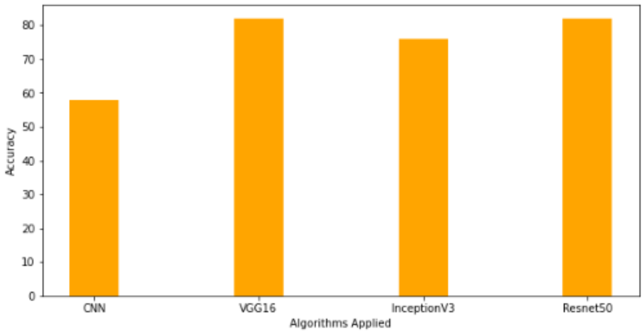

Figure 7. CNN Architecture comparison result

The accuracy of the test data sets by convnet architecture is 58.82%. We did the same for the rest of the model with other architectures. Inception v3 has an 76.47 accuracy rate, followed by the Resnet and VGG16 model, which has a 83.35% % accuracy rate

4. Conclusion

The convolutional neural network is a robust algorithm for image classification. The CNN architectures such as Convnet, VGG16, Inceptionv3, and Resnet have a good performance while doing the classification, but based on our test, vgg16 and Resnet architectures have a significant-good result compared to other architectures. The datasets we are using are not enough for detailed classification, but we can still know how the machine learning works to detect oral cancer symptoms through this model. We hope that in the future this idea can be more developed and it can be used widely to prevent Oral Cancer and save others life.

Reference

[1] Siow-Wee Chang,Sameem Abdul-Kareem,Amir Feisal Merican,Rosnah Binti Zain "Oral cancer prognosis based on clinicopathologic and genomic markers using a hybrid of feature selection and machine learning methods,BMC Bio informatic,vol 170, pp. 2-3, 2013.

[2] B.Ilhan,K.Ling,P.Guneri,P.Wilder-Smith"Improving Oral Cancer Outcomes with Imaging and Artificial Intelligence" Journal of Dental Research, Vol.99, pp. 2-5, 2020.

[3] Jianxin Wu, "Introduction to Convolutional Neural Networks," National key lab for novel software technology Nanjing University, pp. 3-7, 2017.

[4] Shivam Barot, "Oral Cancer (Lips and Tongue) images",Kaggle,kaggle.com/datasets/shivam17299/oral-cancer-lips-and-tongue-images2021.06.25

[5] Mohammed shamim, "Automated detection of oral pre-cancerous tongue lesions using deep learning for early diagnosis of oral cavity cancer.," The Computer Journal,vol 65,pp. 1-8, 2022.

[6] Chandrashekar H S, "oral images dataset", Mendeley Data, https://data.mendeley.com/datasets/mhjyrn35p4/2