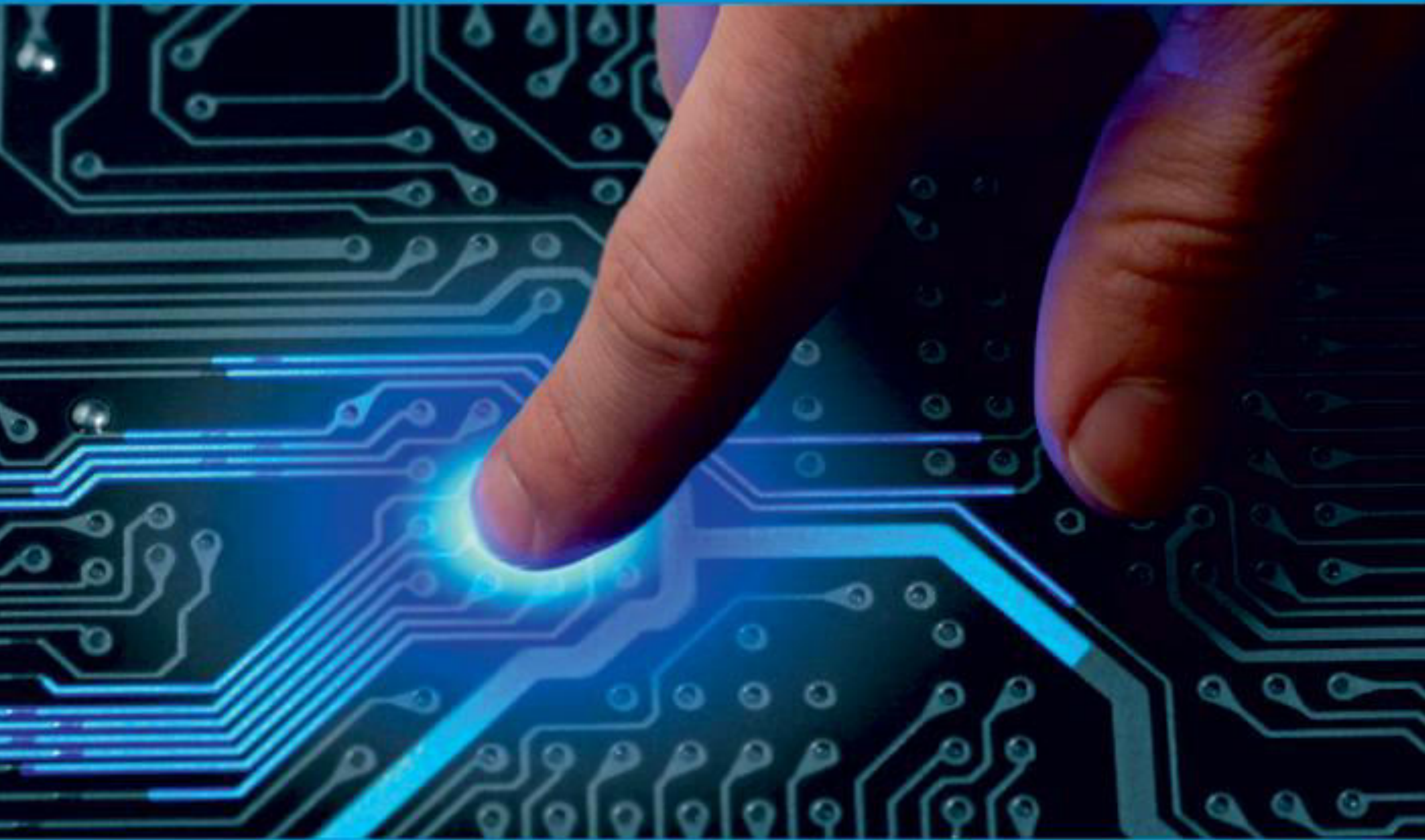
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 6, June 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006152|

Detection of Oral Cancer Using Deep Learning and Microscopic Images

Gowtham SV[1], Rakesh R[2], Aditi AM[3], Bharath GB[4], Prof Shilpa KC[5].

BE Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davanagere , Karnataka, India[1]

BE Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India[2]

BE Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davanagere , Karnataka, India[3]

BE Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India[4]

Assistant Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India[5]

ABSTRACT: Oral cancer is a major global health issue accounting for 177,384 deaths in 2018 and it is most prevalent in low- and middle-income countries. Enabling automation in the identification of potentiallymalignant and malignant lesions in the oral cavity would potentially lead to low-cost and early diagnosis of the disease. Building a large library of well-annotated oral lesions is key. As part of the MeMoSA® (Mobile Mouth Screening Anywhere) project, images are currently in the process of being gathered from clinical experts from across the world, who have been provided with an annotation tool to produce rich labels.A novel strategy to combine bounding box annotations from multiple clinicians is provided in this paper. Further to this, deep neural networks were used to build automated systems, in which complex patterns were derived for tackling this difficult task. Using the initial data gathered in this study, two deep learning based computer vision approaches were assessed for the automated detection and classification of oral lesions for the early detection of oral cancer, these were image classification withResNet-101 and object detection with the Faster R-CNN.Image classification achieved an F1 score of 87.07% for identification of images that contained lesions and 78.30% for the identification of images that required referral. Object detection achieved an F1 score of 41.18% for the detection of lesions that required referral. Further performances are reported with respect to classifying according to the type of referral decision. Our initial results demonstrate deep learning has the potential to tackle this challenging task.

KEYWORDS: Oral cancer, MeMoSA, ResNet, Faster R-CNN.

I. INTRODUCTION

Oral cancer is one of the most common cancers worldwide and is characterized by late diagnosis, high mortality rates and morbidity. GLOBOCAN estimated 354,864 new cases and 177,384 deaths in 2018. Two-thirds of the global incidence of oral cancer occurs in low- and middle- income countries half of those cases are in South Asia. Tobacco use, in any form, and excessive alcohol use are the major risk factors for oral cancer. A factor most prominent in South and Southeast Asia is the chewing of betel quid which generally is comprised of areca nut, slaked lime, betel leaf and may contain tobacco. Nowadays, these quids are available commercially in sachets and are popular in public due to vigorous marketing strategies.

Oral cancer is typically associated with late presentation, particularly in LMICs, where more than two-thirds present at late stages and as a result survival rates are poor. Management of cancers, especially at the late stages, is very costly. The lack of public awareness and the lack of knowledge of health professionals concerning oral cancer is an important reason for late detection.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006152|

Late diagnosis does not need to be a defining attribute as oral cancer is often preceded by visible oral lesions termed as oral potentially malignant disorders (OPMDs) which can be detected during routine screening by a clinical oral examination (COE) performed by a general dentist. If a suspicious lesion is identified the patient is referred to a specialist for confirmation of diagnosis and further management.

Previous studies in India reveal screening has resulted in early diagnosis, down-staging of the disease and reduction in mortality amongst individuals who use tobacco and alcohol. With most of the burden of oral cancer falling on LMICs due to the limited number of specialists and health resources, it is vital that screening programs must offer a low-cost and efficient approach to diagnosis.

II. LITERATURE REVIEW

Texture-based characteristics were the subject of the first articles in the discipline. In contrast to Thomas [18], Krishnan [9] utilised higher order spectra and local binary patterns, as well as rules texturing energy. [10–17], [19], and [20] are examples of recent works that have shifted to deep learning. In order to learn complicated patterns from massive data sets, deep learning employs artificial neural networks with multiple layers of neurons. Deep convolutional neural networks (CNNs) were employed in these articles, and their designs were based on pictures being used as inputs to the CNNs in question. Since AlexNet's victory in the ImageNet image classification competition in 2012, CNNs have grown more popular in the area of computer vision.

Building frameworks around CNNs has made a lot of progress in the field of object detection (predicting bounding boxes and assigning each box to a specific category) for natural image datasets like Pascal VOC (Visual Object Classes), COCO (Common Objects in Context), and other datasets with object classes like cats, dogs, cars, and bicycles. Faster R-CNN [27] and Mask R-CNN [28], which could also produce object instance segmentation, were all part of the R-CNN family of region-based CNN techniques, which employed a two-stage procedure to create the most accurate object detectors to date. One-stage detectors like YOLO (You Only Look Once) and SSD (Single Shot Detector) may be speedier, but they are less accurate than two-stage detectors. object detection frameworks are being researched in the medical imaging industry. The Faster RCNN, for example, has been used to identify colon polyps [31] and to categorise lesions on mammograms [32] Cold sores (herpes labialis) and canker sores, which are innocuous mouth sores, were detected by Anantharaman [20] using a collection of 40 oral pictures and the Mask R-CNN. [20] (aphthous ulcers). Because "instance segmentation" replaced "bounding box detection," the application was given a passing grade.

III. EXISTING SYSTEM

Only a few research on oral cancer have made use of AI technologies. Using chemicals in patients' saliva, researchers have developed a model that could distinguish between oral cancer and periodontitis using machine learning. Using a deep learning algorithm, they were able to determine the tumor's lymphocyte infiltration rate and forecast its prognosis for oral SCC. For the purpose of discovering novel genes associated with oral SCC, Chen et al. (2019) employed one-class learning algorithms, while Chu et al. (2020) reported on the application of machine learning algorithms to predict how oral cancer therapy would proceed (Chen et al., 2019; Chu et al., 2020; Shaban et al., 2019). When investigated by fiber-optic Raman spectroscopy following surgical excision, deep CNNs were shown to be very excellent at recognising the difference between tongue SCC and normal tissues. It was discovered in a recent systematic review that most studies using CNNs for the histological identification of oral precancerous and cancerous lesions had a substantial bias risk.

Problem Statement

In spite of the fact that the mucosa of the mouth is easily observable, oral cancer is frequently discovered at a late stage. Early detection of oral cancer may be made possible at a reasonable cost by using artificial intelligence to detect changes in the oral mucosa that seem worrisome.

Proposed Solution

When developing automated systems to detect oral cancer early, it is critical to gather clinically labelled data that may be utilised to identify the disease. Deep learning can only be effective if it is used on a wide scale. In order to develop this dataset and provide clinical professionals with the tools they need to make detailed annotations, this project is a partnership between individuals from several disciplines of study. Annotations on bounding boxes from multiple

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

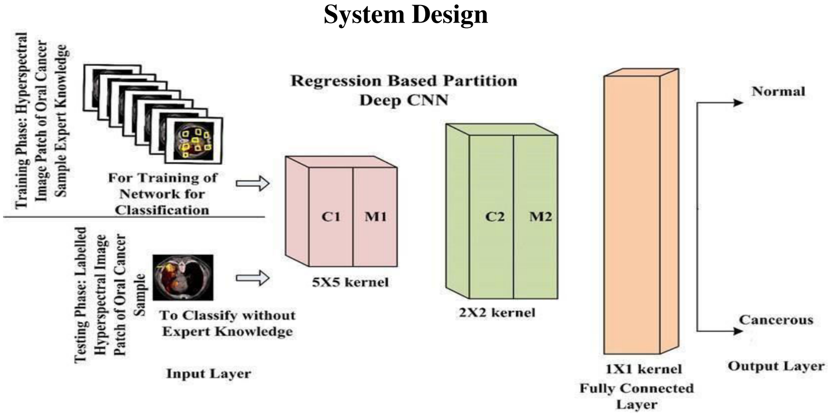
| DOI: 10.15680/IJIRCCE.2022.1006152|

physicians may be combined in a novel manner. This data let us test two distinct methods for automatically finding and classifying oral lesions: a deep learning-based image classification framework and an object identification framework.

Proposed Objectives

- To get the data from different hospitals or a central database of data.
- To use a deep learning neural network to pull out features from pictures of different patients.
- To suggest a deep learning-based method for separating healthy people from people with oral lesions using images.
- To talk about how well different deep learning approaches work.
-

System Design



System Architecture

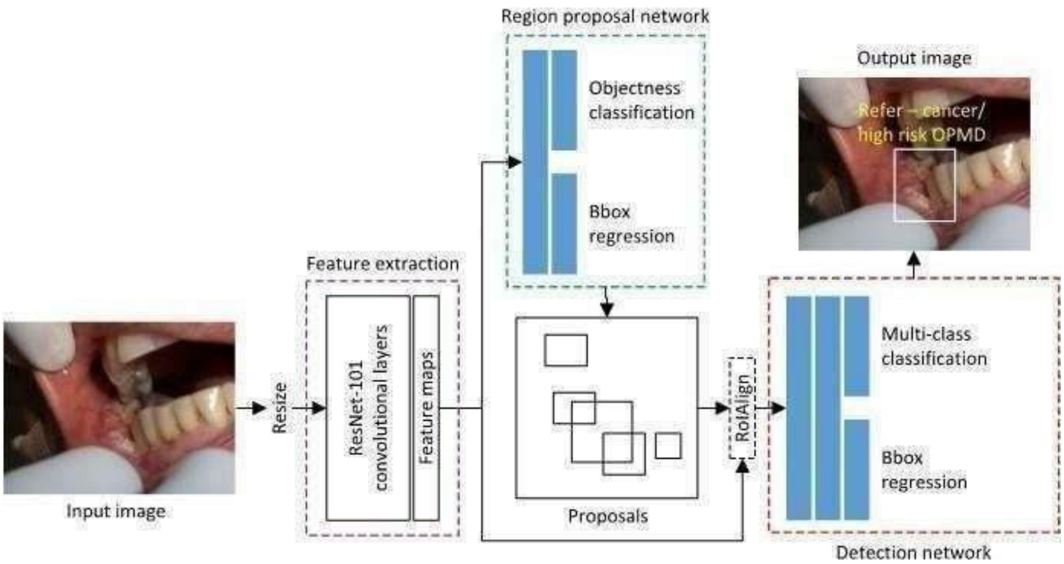


Fig.1: Outline of the faster R-CNN object detection framework applied to four-classoral lesion detection.

Detailed description of the methodology

Computer vision's object detection employs a bounding box to locate items in an image and determine to which class they belong [37]. Faster R-CNN [27], based on deep learning and combining categorization and bounding box regression, was utilised to locate objects. Step one of the Faster R-CNN was accomplished in two stages. As a first phase, the region proposal network (RPN) generated a limited number of object/region suggestions, each of which was given a score based on how similar it was to an actual item. As part of the "detection network," area ideas were broken down into two categories: item types and backgrounds. The convolution layers in both networks were identical. The framework, which was a CNN article, has several common layers (which can also be called the base CNN). Convolutional interpolation added rich hierarchical characteristics to the input picture.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006152|

Consider the RPN as transmitting an image to each point on the feature map by using a CNN as a back-end. A 3 x 3 pixel input is all that is needed for this. There were a number of anchors of various sizes and aspect ratios, all of which were bound by the same procedure (which made reference to the original image). Using a binary classification system, the tiny network determined whether or not an anchor was an item based on its bounding box coordinates as determined by a regression layer. The RPN was able to do all of this in a completely convolutional manner. As a result, the picture was littered with a grid of region ideas. After that, the NMS was switched off, and only the best area suggestions were provided to the detection network thereafter. Each of the remaining region suggestions was converted into a tiny, fixed-size area on the feature map from the basic CNN using a pooling layer called RoIPool [26]. Next, a regression layer and a softmax classification layer were used by the detection network to provide more precise bounding box coordinates for each suggested area as well as an item class and a confidence score. Until recently, the only NMS detections were based on classes. More details may be found in the original article [27].
We used ResNet-101 [33] and the feature pyramid network [38] as the basis CNN for our model due to updates to the original Faster R-CNN publication [27]. It was also found that replacing the RoIPool layer with the RoIAlign layer improved the Mask R-CNN [28]. For the Faster R-CNN model, the COCO dataset [24], which contains 328,000 pictures divided into 80 groups, was employed as a source of transfer learning. [21] Prior to this, we had previously trained ResNet-101 on the ImageNet dataset [21]. Frozing the layers preceding conv5 1 of the basic CNN and utilising our oral lesion dataset to fine-tuning the remainder of the system resulted in the best model being built. It was necessary to develop three distinct object detection models in order to assess the complexity of the job (detailed below). Our models transmit the bounding boxes, class, and confidence score for each detection. To account for each model's number of distinct classes, the number of softmax classification neurons in the detection network was determined.

IV. RESULTS AND DISCUSSION

Fig.2: Above image describes the prediction of oral cancer after uploading the image which represents the oral cancer condition

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006152|

Fig.3: Above image describes the prediction of oral cancer after uploading the image which represents the non-oral cancer condition.

V. CONCLUSION

This project has discussed the collection and annotation of images from the oral cavity and demonstrated results for automating the early detection of oral cancer. The contribution of this project is a novel strategy to combine bounding box annotations from multiple clinicians; followed by the assessment of two different deep learning-based approaches to provide a solution to automation. Our promising initial results demonstrate the effectiveness of deep learning and suggest it has the potential to tackle this challenging task. Performances are set to increase as the dataset grows and this will have a significant impact in lowand middle-income countries where health resources are limited.

REFERENCES

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," CA. Cancer J. Clin., vol. 68, no. 6, pp. 394–424, 2018.
[2] H. Gelband, P. Jha, R. Sankaranarayanan, and S. Horton, Disease Control Priorities, (Volume 3): Cancer. The World Bank, 2015.
[3] J. Rimal, A. Shrestha, I. K. Maharjan, S. Shrestha, and P. Shah, "Risk assessment of smokeless tobacco among oral precancer and cancer patients in eastern developmental region of Nepal," Asian Pacific J. Cancer Prev., vol. 20, no. 2, pp. 411–415, 2019.
[4] J. G. Doss, W. M. Thomson, B. K. Drummond, and R. J. R. Latifah, "Validity of the FACT-H&N (v 4.0) among Malaysian oral cancer patients," Oral Oncol., vol. 47, no. 7, pp. 648–652, 2011.
[5] H. Amarasinghe et al., "Economic burden of managing oral cancer patients in Sri Lanka: a cross-sectional hospital - based costing study," BMJ Open, vol. 9, no. 7, 2019.
[6] R. D. Jayasinghe, L. P. G. Sherminie, H. Amarasinghe, and M. A. Sitheeque, "Level of awareness of oral cancer and oral potentially malignant disorders among medical and dental undergraduates," Ceylon Med. J., vol. 61, no. 2, 2016.
[7] P. Brocklehurst, O. Kujan, L. A. O'Malley, G. Ogden, S. Shepherd, and A.-M. Glenny, "Screening programmes for the early detection and prevention of oral cancer," Cochrane database Syst. Rev., no. 11, 2013.
[8] N. Haron et al., "Mobile phone imaging in low resource settings for early detection of oral cancer and concordance with clinical oral examination," Telemed. e-Health, vol. 23, no. 3, pp. 192–199, 2017.
[9] M. M. R. Krishnan et al., "Automated oral cancer identification using histopathological images: a hybrid feature extraction paradigm," Micron, vol. 43, no. 2–3, pp. 352–364, 2012.
[10] M. Aubreville et al., "Automatic classification of cancerous tissue in laserendomicroscopy images of the oral cavity using deep learning," Sci. Rep., vol. 7, no. 1, p. 11979, 2017.
[11] J. Folmsbee, X. Liu, M. Brandwein-Weber, and S. Doyle, "Active deep learning: Improved training efficiency of convolutional neural networks for tissue classification in oral cavity cancer," in 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 2018, pp. 770–773.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006152|

[12] R. K. Gupta, M. Kaur, J. Manhas, and others, "Tissue Level Based Deep Learning Framework for Early Detection of Dysplasia in Oral Squamous Epithelium," J. Multimed. Inf. Syst., vol. 6, no. 2, pp. 81–86, 2019.
[13] P. R. Jeyaraj and E. R. S. Nadar, "Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm," J. Cancer Res. Clin. Oncol., vol. 145, no. 4, pp. 829–837, 2019.
[14] S. Xu et al., "An Early Diagnosis of Oral Cancer based on Three-Dimensional Convolutional Neural Networks," IEEE Access, vol. 7, pp. 158603–158611, 2019.
[15] B. Song et al., "Automatic classification of dual-modalilty, smartphone-based oral dysplasia and malignancy images using deep learning," Biomed. Opt. Express, vol. 9, no. 11, pp. 5318–5329, 2018.
[16] R. D. Uthoff et al., "Point-of-care, smartphone-based, dual-modality, dual-view, oral cancer screening device with neural network classification for low-resource communities," PLoS One, vol. 13, no. 12, p. e0207493, 2018.
[17] A. Rana, G. Yauney, L. C. Wong, O. Gupta, A. Muftu, and P. Shah, "Automated segmentation of gingival diseases from oral images," in 2017 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT), 2017, pp. 144–147.
[18] B. Thomas, V. Kumar, and S. Saini, "Texture analysis based segmentation and classification of oral cancer lesions in color images using ANN," in 2013 IEEE International Conference on Signal Processing, Computing and Control (ISPCC), 2013, pp. 1–5.
[19] R. Anantharaman, V. Anantharaman, and Y. Lee, "Oro vision: Deep learning for classifying orofacial diseases," in 2017 IEEE International Conference on Healthcare Informatics (ICHI), 2017, pp. 39–45.
[20] R. Anantharaman, M. Velazquez, and Y. Lee, "Utilizing Mask R-CNN for Detection and Segmentation of Oral Diseases," in 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2018, pp. 2197–2204.
[21] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in 2009 IEEE conference on computer vision and pattern recognition, 2009, pp. 248–255.
[22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.
[23] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The Pascal Visual Object Classes Challenge: A Retrospective," Int. J. Comput. Vis., vol. 111, no. 1, pp. 98–136, Jan. 2015.
[24] T.-Y. Lin et al., "Microsoft coco: Common objects in context," in European conference on computer vision, 2014, pp. 740–755.
[25] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587.
[26] R. Girshick, "Fast r-cnn," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1440–1448.
[27] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in Advances in neural information processing systems, 2015, pp. 91–99.
[28] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 2961–2969.
[29] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.
[30] W. Liu et al., "Ssd: Single shot multibox detector," in European conference on computer vision, 2016, pp. 21–37.
[31] Y. Shin, H. A. Qadir, L. Aabakken, J. Bergsland, and I. Balasingham, "Automatic colon polyp detection using region based deep cnn and post learning approaches," IEEE Access, vol. 6, pp. 40950–40962, 2018.
[32] D. Ribli, A. Horváth, Z. Unger, P. Pollner, and I. Csabai, "Detecting and classifying lesions in mammograms with deep learning," Sci. Rep., vol. 8, no. 1, p. 4165, 2018.
[33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
[34] N. Haron et al., "m-Health for Early Detection of Oral Cancer in Low-and Middle-Income Countries," Telemed. e-Health, 2019.
[35] V. C. Raykar et al., "Learning from crowds," J. Mach. Learn. Res., vol. 11, no. Apr, pp. 1297–1322, 2010.
[36] X. Li, B. Aldridge, J. Rees, and R. Fisher, "Estimating the ground truth from multiple individual segmentations with application to skin lesion segmentation," in Proc. Medical Image Understanding and Analysis Conference, UK, 2010, vol. 1, pp. 101–106.
[37] Z.-Q. Zhao, P. Zheng, S. Xu, and X. Wu, "Object detection with deep learning: A review," IEEE Trans. neural networks Learn. Syst., 2019.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006152|

[38] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature pyramid networks for object detection,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2117–2125.
[39] J. Huang et al., “Speed/accuracy trade-offs for modern convolutional object detectors,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 7310–7311.
[40] W. Abdulla, “Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow,” GitHub repository. Github, 2017.
[41] J. K. Chorowski, D. Bahdanau, D. Serdyuk, K. Cho, and Y. Bengio, “Attention-based models for speech recognition,” in Advances in neural information processing systems, 2015, pp. 577–585.
[42] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv Prepr. arXiv1409.1556, 2014.
[43] K. Xu et al., “Show, attend and tell: Neural image caption generation with visual attention,” in International conference on machine learning, 2015, pp. 2048–2057.
[44] F. Wang et al., “Residual attention network for image classification,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 3156–3164.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING