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

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13835

Skin Cancer Classification by Leveraging Segment Anything Model for Semantic Segmentation of Skin Lesion

Mani Abedini

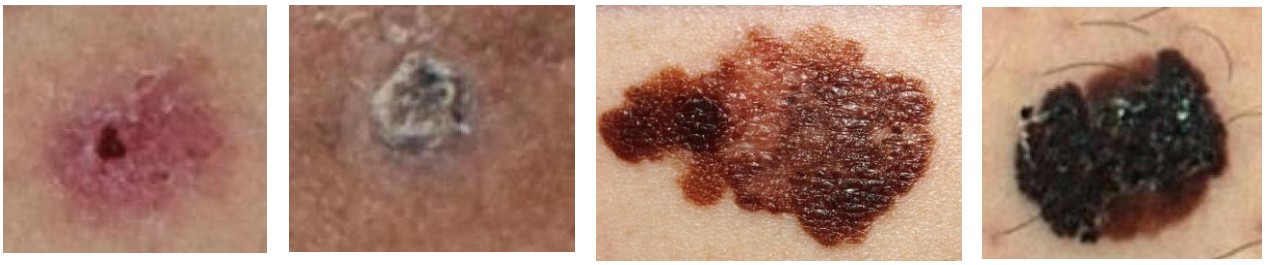
Data Science Manager, AWRostamani, Dubai, UAE[1]

Abstract: Skin cancer is a growing public health concern; while some types of skin cancer are deadly, such as Melanoma, early detection is crucial for effective treatment and improving patient survival rates [1,2,3,4]. In fact, Malignant melanoma accounts for only 2.3% of all skin cancers yet is responsible for more than 75% of skin cancer-related deaths. However, if it is detected at an early stage, it is highly curable; the 10-year survival rate is between 90% and 97% when the tumour thickness is less than 1 mm. Also, the treatment for an early detected cancerous mole is as simple as excision of the lesion, which can prevent metastasis and spread of cancer to other organs. In this research study, we introduce an approach for skin cancer classification using a state-of-the-art deep learning architecture that has demonstrated exceptional performance in diverse image analysis tasks. We have used two publicly available benchmark data sets for training and validating our results: HAM10000 and ISIC2018 datasets. These datasets consist of dermoscopic images captured using Dermatoscopes and carefully annotated by expert dermatologists. Preprocessing techniques, such as normalization and augmentation, are applied to enhance the robustness and generalization of the model. The proposed approach demonstrated the efficacy of extracting relevant features for accurate classification by leveraging Deep Object Detection models to identify the location of the Lesion, then using the Segment Anything Model (SAM) and MedSAM for extracting the border of the lesions, then finally using various pre-trained states-of-the-art Deep Convolution Networks for Classification. Comprehensive experiments and evaluations are performed in this research; the results demonstrate the effectiveness of using Zero-Shot Segmentation methods over traditional deep learning architectures in skin cancer classification.

Keywords: Skin Cancer, Computer vision, Cancer classification, image processing, Vision Transformer, image classification, Cancer Cell Segmentation.

I. INTRODUCTION

Excessive exposure to Sunlight's ultraviolet (UV) rays or other sources of UV rays can damage the DNA of the external layer of the skin (epidermis) cells, resulting in abnormal cell growth and the formation of skin cancer [1,2]. Amon many types of skin cancer, Basal Cell Carcinoma (BCC), Squamous Cell Carcinoma (SCC), Melanoma, and Merkel Cell Carcinoma (MCC) are most common types of cancers (Fig 1 shows some examples images of these type of skin cancers). A benign skin tumour is called nevus. Melanoma is the most dangerous type of skin cancer. It develops when melanocytes (the cells that give the skin pigment) start to grow out of control. Fortunately, Melanoma can be cured if detected early. Almost all nevi are not harmful, but some types can become Melanoma.



Basal Cell Carcinoma (BCC) | **Squamous Cell Carcinoma (SCC)** | **Merkel Cell Carcinoma (MCC)** | **Melanoma**

Fig. 1 Example images of BCC, SCC, MCC and Melanoma skin cancers

DOI: 10.17148/IJARCCE.2024.13835

Accurate skin lesion detection requires specialized expertise. Experienced dermatologists are needed for accurate skin cancer detection, which may not be readily available to all patients. As such, computer-aided skin cancer detection methods using computer vision and deep learning algorithms have been studied for skin cancer detection. In this paper, we have done extensive research to leverage state-of-the-art semantic segmentation algorithms to analyse the border of lesions and improve the accuracy of the final diagnosis.

Since predominantly skin cancer appears on the surface of the skin due to UV exposure, using non-invasive, cost-effective methods of using a Dermatoscope or high-resolution smartphone camera is common with dermatologists. As a result, there is increasing attention to developing computer-aided diagnoses of skin cancer by leveraging Image processing, computer vision techniques, and deep learning methods to assist healthcare professionals in increasing the accuracy of diagnosis at the very early stages of cancer while reducing false positive diagnosis [1].

In order to analyze the skin image and consequently improve the clinical diagnosis of skin cancer, several diagnostic algorithms have been developed. One of the most common approaches is the ABCD rule of dermoscopy [2]. In this approach, each letter corresponds to a characteristic feature of Skin Cancer: A stands for Asymmetry (asymmetrical shapes can be an indication of cancer), B for Border irregularity (uneven, ragged, unsmooth borders), C for Colour variation (more range of colours is high risk of cancer), D for Diameter (lesion with larger than 6 millimetres diameter). Skin cancer segmentation is a task focusing on identifying the location and boundary of skin lesions. Which is related to A, B and D in ABCD approach. Knowing the border of the lesion, the automatic algorithm can measure if the lesion is symmetric or asymmetric. Also, if the border is well-shaped or has irregularity, the size of the lesion is measured by identifying the lesion border. Thus, accurate skin cancer segmentation is very critical for identifying skin cancer.

With the rise of deep learning and computer vision methods for medical image analysis in recent years, many approaches have been developed to automate skin lesion analysis and assist skin care professionals in identifying cancer lesions quickly and accurately [3,4,5]. Specifically for lesion segmentation, there are many deep learning methods have been proposed which has shown better performance compared to traditional computer vision-based segmentation algorithms, such as R-CNN [6] and U-Net [7,8]. One of the challenges facing developing deep learning methods is having enough high-quality annotated images labelled and annotated by medical professionals. The costs of creating such databases are very high. Zero-shot learning [9] is a promising alternative approach to overcome such limitations. In zero-shot learning, a model is trained on a different domain and applied to the new type of objects for Segmentation by mapping the visual features of an object to a semantic space. Recently, a zero-shot learning method called the Segment Anything Model (SAM) [10, 11] demonstrated the highest accuracy in most popular benchmark datasets and raised a lot of attention to this field. SAM has been trained on 11 million images and 1 billion masks, enabling it to understand the visual concepts deeply. The model can segment any object by providing a bounding box around the area or specifying points that are inside or outside the area.

In this paper, we developed a skin cancer detection model inspired by the ABCD rule. We used segmentation to identify the border of the lesion, remove background skin, and then classify the extracted image. We adapted SAM [10] and MedSAM [12,13] in the proposed model and compared the performance of each method. For the classification task, various deep-learning convolution neural networks (CNN) were fine-tuned. In the next Section, we review relevant papers and work. In Section III, the proposed method is explained in detail. The benchmark dataset is defined, the experiment setup is explained, and the experiment results are presented. In the last section, we conclude the key learning findings from our experiments.

II. RELATED WORK

Since the rise of developing Machine Learning and Computer Vision for developing diagnosis systems [14,15], various methods have been proposed to identify skin cancer from skin lesion images. For example, Dorj et al. [16] used Deep Convolution Neural Networks for skin cancer classification. Proposed combined method of AlexNet and ECOC SVM. In their approach, AlexNet were used to generate features and SVM for classification. They have reported 94% accuracy. Ameri et al. also used CNN, trained their model on HAM10000 dermoscopy images [17,18]. They used vanilla CNN, without any feature extraction for lesion segmentation, and reported 84% accuracy [19]. Mohapatra et al. applied a MobileNet on the HAM10000 dataset and achieved an accuracy of 80% [20], while Chaturvedi et al. could improve the accuracy of MobileNet by leveraging image augmentation and fine tuning parameters to 83.1% [21]. T Rezvantalab et al. utilized various CNN models: DenseNet, ResNet, Inception v3, and InceptionRes in their experiments; they reported DenseNet on HAM10000 (10015 images) + PH2 (120 images) datasets; consisting of eight cancer classes achieved accuracy of 86.59% [22].

DOI: 10.17148/IJARCCE.2024.13835

K. M. Hosny et al. demonstrated with a customized AlexNet model and image augmentation, the accuracy of CNN for classification on PH2 data set can be improved to 98.61% [23]. T. Emara et al. also used the HAM10000 dataset to develop a customized InceptionV4 model and reported 94.7% accuracy [24].

Yu et al. developed a customized CNN with 50 layers [25]. In their approach they used fully convolutional residual network (FCRN) for segmentation and deep residual networks for classification. In this approach, they could acquire more discriminative features, which eventually demonstrated higher accuracy in the training phase. Their work suggests that a two-stage framework with segmentation can achieve better results than direct classification of the dermoscopic images. Andre Esteva et al. developed deep convolutional neural networks using a dataset of 129,450 clinical images [26]. What is unique in their experiment is validating the result of CNN against 21 board-certified dermatologists on the biopsy-proven clinical images. The results demonstrated that CNN can classify skin cancer at a level of competence dermatologists. The accuracy of their proposed CNN, was 96% for carcinoma images, 94% for melanoma images which is comparable with accuracy of an experienced dermatology. Jinnai et al. applied a faster region-based CNN (FRCNN) and could achieve 91.5% accuracy [27]. Nawaz et al. proposed region-based convolutional neural networks (RCNN), combined with fuzzy k-means clustering (FKM). They used faster-RCNN for feature extraction, followed by FKM for lesion segmentation, they reported using ISBI-2016, ISIC-2017, and PH2 image data sets in their experiments [28].

Garcia studied the importance of fusing non-medical data with skin cancer dermoscopic image. They used a ResNet model pre-trained on non-medical data, then finetuned on dermoscopy images. The study observed a significant performance increase using transfer learning approach. The results suggest that non-medical image feature learning is a critical part of developing an image classification model and does not necessarily require medical images. In the final layers of the network, the classification layer connects the trained image patterns to the label [29].

Most of the effort in this area was proposing using the complete image to CNN for classification, while this approach shows high accuracy, but not efficient. Because CNN is not effectively learning discriminative features of target class versus normal class. The network is exposed to various features in the background (surrounding skin), which is not related to the classification task at all, and is indeed noise for training process. Thus, some researchers proposed identifying the location of the lesion using object detection as an intermediate step first, then removing background skin, before classification. For example, Rehan Ashraf et al. proposed a transfer learning-based lesion boundary detection method for skin cancer detection. Previous deep learning-based methods had used complete images for feature learning which can result in a lack of performance in terms of discriminative feature extraction. An Object detection approach identifies the area of interest, then a classification or segmentation model can focus on discriminative features. Rehan proposed method, extract the lesion from the background, then applied a CNN while using image augmentation to fine tune the CNN on training images. They reported an accuracy of 97.9% , suggesting the efficacy of lesion detection and extracting the lesion from the background image is performing better than applying CNN directly on the entire image. Goyal et al. developed a Faster-RCNN for object localization detection as well [30]. In their approach has two phase, object detection to identify the bounding box around the lesion then segmentation to draw the border of lesion. Their experiment results suggest that this approach can improve the segmentation accuracy. In their paper they reported an accuracy of 94.5% and recall of 94.3%, on the ISBI-2017 benchmark dataset.

Jinnai et al. worked with the National Cancer Center, Tokyo and collected a dataset of 5846 dermoscopic images with six classes. They also used FRCNN and reported accuracies of 86.2%in their experiments [31]. Chaturvedi et al. applied ResNet on the HAM10000 dataset and reported an accuracy of 92.83% [32]. In their experiments they finetuned pre-trained models on ImageNet over HAM10000 data set. Garg et al. also explored ResNet50 and VGG16 on HAM10000 dataset (7 classes), achieving an accuracy of 90.51% [33]. In their study they leveraged fine tuning VGG16 and ResNet50 and combining it with Random Forest, XGBoost, and SVM. Benedetti et al. applied InceptionResNetv2 to the HAM10000 dataset (7 classes) and achieved an accuracy of 78.9% [34]. Hatice Catal Reis et al. also applied GoogleNet CNN model on a combined datasets of International Skin Imaging Collaboration HAM10000 images (ISIC 2018), ISIC 2019, and ISIC 2020 datasets. Their approach results in accuracies of 94.59%, 91.89%, and 90.54% respectively for three datasets [35]. Bechelli et al used Xception, VGG16, and ResNet50 in their work, to identify the benign and skin cancer. In their studies they explored both ISIC archive and HAM10000 [36].

Recently, a zero-shot learning method called the Segment Anything Model (SAM) demonstrated outstanding performance in several benchmark datasets. SAM is trained on the SA-1B dataset [37], which contains over 11 million images and 1 billion masks, empower the model to understand the visual concepts. Hua et al. publish a research paper demonstrating the usage of out of the box SAM for skin lesion segmentation. They tried the model on HAM10000 data sets, using a prompt (bounding box) and without a prompt. The results suggest knowing the location of lesion is essential for higher accuracy of segmentation [38].

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

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13835

In similar study, Ma et al. extended the SAM model on medical images. They retrain the SAM model on a large-scale medical image dataset with 1,570,263 image-mask pairs, covering 10 imaging modalities and over 30 cancer types [39]. In our experiments we use Both SAM and MedSAM implementation for lesion segmentation.

III. METHODOLOGY

III-A Data Sets

This study used two publicly available datasets containing Skin Cancer images. The first data set is HAM10000 which contains 10015 images of 7470 lesions, seven categories: Melanocytic Nevi, Melanoma, Benign Keratosis-like Lesions, Basal Cell Carcinoma, Pyogenic Granulomas and Hemorrhage, Actinic Keratoses and Intraepithelial Carcinomae, Dermatofibroma. The data set has border segmentation as well [40,41]. In our experiments we used this data set to train our object detection algorithm to identify the lesion location.

The second data set is International Skin Imaging Collaboration (ISIC 2018) contains 2594 dermatologic images and associated ground truth segmentation masks [42,43]. Since 2016, ISIC has conducted annual challenges for the computer science community; since 2016 till today ISIC datasets become the largest publicly available collection of quality controlled dermoscopic images of skin lesions. The objective is to improve melanoma diagnosis crowdsourcing the AI and computer vision enhancement; ISIC is sponsored by the International Society for Digital Imaging of the Skin (ISDIS).

III-B Preprocessing

In computer vision, preprocessing is a critical step, especially for dermoscopic images collected from various clinics and different imaging setups. In our experiments first we applied digital hair removal (DHR) algorithm [44]. To avoid removing any critical patterns we avoid using any noise removal filter. All images are resized to 224 × 224 to be consistent with the input layer of our deep learning models. Image augmentation requires generating a good amount of annotated data so we can retrain the deep neural networks. Since annotating medical images requires to be conducted by experienced medical professionals, data acquisition takes time and very expensive. Generating more images from existing annotated data is the best cost-effective way to overcome the situation.

In our experiments, all images were scaled with 1/255. We also allow random rotation between 0 and 45 degrees. The zoom level was between 0.5 and 2; numbers below 1.0 result in zooming out, and numbers bigger than 1.0 will magnify. We also allow random adjustment of brightness. The random noise in brightness will help the network be less sensitive to specific image brightness and try to learn the underlying patterns associated with cancer lesions.

III-C Proposed Method

In this paper, we proposed a multi-step process for skin cancer classification. The first step is to identify the location of the lesion. For this step, we have explored various states of the art deep learning object detection methods. All models have been fine-tuned using the HAM10000 data sets. The data set has segmentation masks, so a bounding box fitted to the lesion is calculated for each image. Among all models, YOLOv8 (You Only Look Once), a highly popular deep-learning object detection model, has demonstrated better performance than others.

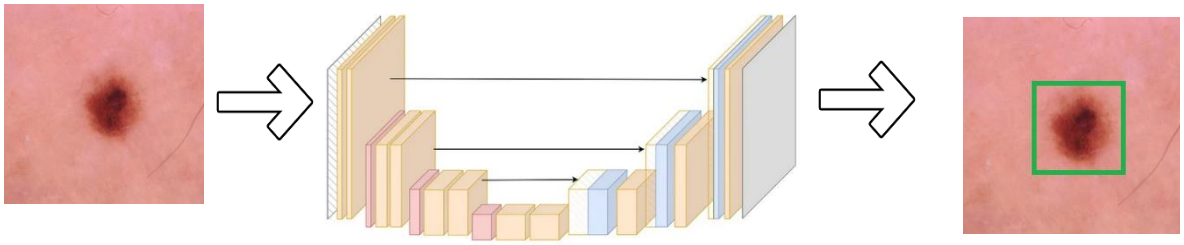


Figure 1: Identifying the location of Skin Lesion using an Object Detection network.

In the second step, we used out-of-the-box Meta's pre-trained Segment Anything Model (SAM) model, as well as Segment Anything in Medical Images (MedSAM) to segment the border of lesions. Both segmentation methods require the bounding box around the object as input. For this reason, we have used Deep Learning Object Detection model in the first step to identify the location of the lesion and the bounding box coordination fit around the lesion. The image, along with the bounding box coordinates, are given to SAM and MedSAM.

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

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13835

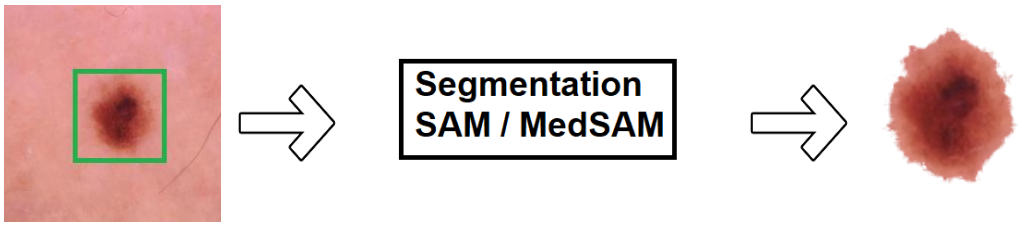


Figure 2: Lesion Segmentation using Segment Anything Model (SAM) and Medical SAM (MedSAM)

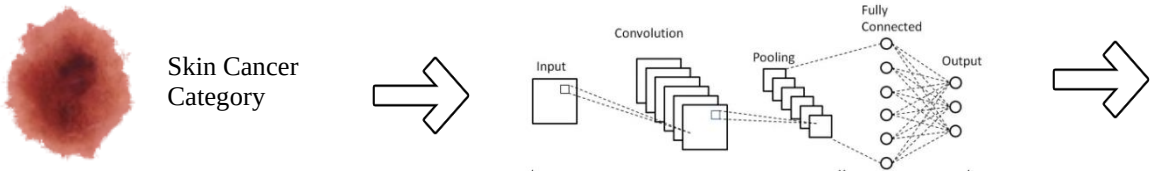


Figure 3: Skin Cancer Classification using CNN on the cropped image of the lesion without background skin.

III-D Evaluation Metrics

To evaluate the performance of the proposed methods, four widely used metrics have been measured in our experiments: Accuracy (1), Precision (2), Recall (3), and F1-Score (4). Please see the formula for each metric below:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = \frac{2 \times Recall \times Precision}{Recall+Precision} \quad (4)$$

TP (True Positive) represents the number of correctly predicted positive cases.
TN (True Negative) represents the number of correctly predicted negative cases.
FP (True Positive) represents the number of incorrectly predicted positive cases.
FN (True Negative) represents the number of incorrectly predicted negative cases.

Accuracy is the most commonly used measurement of the accurate model, defined as the portion of actual positive and negative cases over all measured cases.

Precision defines the portion of positive predictions that have been correctly identified.

Recall, also known as sensitivity or true positive rate, on the other hand, measures the portion of actual positive area that has been correctly predicted as positive.

F1-score is the harmonic mean of the precision and recall.

For Segmentation, we used Intersection Over Union (IoU), a popular metric for measuring the accuracy of a predicted segmentation to a ground truth segmentation. IoU is the ratio of the intersection of the two areas to their combined areas. IoU is calculated as in Formal 5. For Binary class cases it can be also calculated as in Formula 6:

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

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13835

$$IoU = \frac{A \cap B}{A \cup B} \qquad (5)$$

$$IoU = \frac{TP}{TP+FN+FP} \qquad (6)$$

III-E Experiments

In our proposed approach, the first step is developing an accurate Deep Learning Object detection network. For this purpose, we have explored various state-of-the-art Object detection models: YOLO (You Only Look Once) [45], EfficientDet [46], RetinaNet [47], and Faster R-CNN [48]. For training data, we used the first data set explained in Section III-A. All methods have been trained and fine-tuned. However, YOLOv8 demonstrated better performance than all other methods.

The second phase is classification, after lesion segmentation and removing background from cropped images. Many recently introduced CNN models have demonstrated high accuracy in various benchmark data sets. We have selected a couple of widely used models in our experiments. A combination of newly developed models and well-established CNN architecture: InceptionV3 [49], DenseNet [50], Xception [51], VGG16 [52], EfficientNet [53], Coca-ViT [54]. Contrastive Captioner (CoCa) is one of the recently developed classification models that has the highest score on the ImageNet leaderboard. This network has a pretrain image-text encoder-decoder foundation model. CoCa is unifying natural language supervision for representation learning, applying a contrastive loss between unimodal image and text embeddings, and combining multimodal learning such as image and text classification.

IV. RESULTS AND DISCUSSION

In this section, we present the results of the experiment and discuss the effectiveness of the method for each phase of the process.

For Object Detection phase, we conducted comprehensive experiments using 4 popular deep learning object detection models (YOLO, EfficientDet, RetinaNet, Faster R-CNN). The results are presented in Table I. The results suggest that YOLOv8 is performing better than others for object detection. However, the performance of most models are very close, suggesting the object detection is effective and regardless of which model, definitely can help identifying the location of suspicious lesion.

TABLE I THE PERFORMANCE OF VARIOUS DIFFERENT OBJECT DETECTION MODELS TO IDENTIFY THE LOCATION OF SKIN MOLE

Deep Learning Object Detection Model	Accuracy	Precision	Recall	F1-Score
YOLOv8	85.01%	70.34%	81.33%	75.43%
EfficientDet	82.75%	68.26%	80.32%	73.80%
RetinaNet	77.54%	63.88%	77.54%	70.05%
Faster R-CNN	75.42%	73.87%	74.33%	74.09%

The next step is segmentation. We applied SAM and MedSAM, both for ISIC 2018. The Segmentation results are presented in Table II.

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

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13835

TABLE III THE PERFORMANCE OF SAM AND MEDSAM ON ISIC 2018

Deep Learning Segmentation Model	IoU
SAM	70.03%
MedSAM	78.55%

Table 4:

In the classification phase of our experiments, after extracting the lesion by the MedSAM, we used mostly common Deep learning CNNs for classification: InceptionV3, DenseNet, Xception, VGG16, EfficientNetV2 B3, and CoCa. Most of these models are already implemented in Keras, so we used those implementations. In the case of CoCa, we used the CoCa implementation code in OpenCLIP [55]. The results of classification has been presented in Table 4.

TABLE IIII THE PERFORMANCE OF VARIOUS DIFFERENT DEEP LEARNING CONVOLUTION NEURAL NETWORK (CNN) MODELS FOR IMAGE CLASSIFICATION OF SKIN CANCER

Deep Learning Classification Model	Accuracy	Precision	Recall	F1-Score
InceptionV3	78.34%	73.44%	88.37%	80.21%
DenseNet	77.95%	75.36%	83.65%	79.28%
Xception	79.45%	78.23%	84.66%	81.31%
VGG16	75.90%	74.65%	75.14%	74.89%
EfficientNetV2 B3	82.22%	81.23%	90.55%	85.63%
CoCa	85.33%	84.66%	91.56%	87.97%

Table 4:

The results of the classification phase suggest that most state-of-the-art deep-learning classification models are capable of classifying Skin Cancer images with high accuracy. However, CoCa present higher accuracy, F1-Score among all other models, which suggests that the cross modal transfer learning approach such as CoCa which has been trained on image and text labels, has more robust pretrained network and can be fine-tuned to a very specific medical image classification tasks such as Skin Cancer classification.

V. CONCLUSION

In this paper, we utilize a three-step deep learning-based process to classify Skin Cancer Dermoscopic images. The first step uses Deep Object Detection models to identify the location of the lesion. We have explored various Deep Learning Object Detection models, fine-tuned them on HAM10000 data set. The newly enhanced YOLO model (v8) has outperform other models in this task. In the next step, the bounding box surrounding the detected Lesion will be passed to the Semantic Anything Model (SAM) and MedSAM to identify the border of the lesion. By extracting the background image from the lesion itself, we allow the image classification model to focus on the property of the lesion and its patterns. For the classification task, we have fine-tuned various widely used Deep Convolution Neural Network models on our second acquired data set. The results shown CoCa performance on image classification is better than other CNN models for Skin Cancer classification.

DOI: 10.17148/IJARCCE.2024.13835

REFERENCES

[1]. M. Abedini *et al.*, "A generalized framework for medical image classification and recognition," in *IBM Journal of Research and Development*, vol. 59, no. 2/3, pp. 1:1-1:18, March-May 2015, doi: 10.1147/JRD.2015.2390017. keywords: {Feature extraction;Biomedical imaging;Histograms;Data models;Image recognition;Malignant tumors;Image color analysis}.

[2]. W. Stolz, A. Riemann, A. B. Cognetta, L. Pillet, W. Abmayr, D. Holzel, P. Bilek, F. Nachbar, M. Landthaler,and O. Braun-Falco, "ABCD rule of dermatoscopy: A new practical method for early recognition of malignant melanoma," European Journal of Dermatology, vol. 4, pp. 521–527, 1994.

[3]. Codella, N., Cai, J., Abedini, M., Garnavi, R., Halpern, A., Smith, J.R. (2015). Deep Learning, Sparse Coding, and SVM for Melanoma Recognition in Dermoscopy Images. In: Zhou, L., Wang, L., Wang, Q., Shi, Y. (eds) Machine Learning in Medical Imaging. MLMI 2015. Lecture Notes in Computer Science(), vol 9352. Springer, Cham. https://doi.org/10.1007/978-3-319-24888-2_15

[4]. Abedini, Mani & Chen, Qiang & Codella, Noel & Garnavi, Rahil & Sun, Xingzhi. (2015). Accurate and Scalable System for Automatic Detection of Malignant Melanoma. 10.1201/b19107-11.

[5]. Bozorgtabar, B., Abedini, M., Garnavi, R. (2016). Sparse Coding Based Skin Lesion Segmentation Using Dynamic Rule-Based Refinement. In: Wang, L., Adeli, E., Wang, Q., Shi, Y., Suk, HI. (eds) Machine Learning in Medical Imaging. MLMI 2016. Lecture Notes in Computer Science(), vol 10019. Springer, Cham. https://doi.org/10.1007/978-3-319-47157-0_3.

[6]. K. He, G. Gkioxari, P. Dollár et al., "Mask r-cnn." 2961-2969.

[7]. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation." 234-241.

[8]. H. Cao, Y. Wang, J. Chen et al., "Swin-unet: Unet-like pure transformer for medical image segmentation." 205-218.

[9]. W. Wang, V. W. Zheng, H. Yu et al., "A survey of zero-shot learning: Settings, methods, and applications," ACM Transactions on Intelligent Systems and Technology (TIST), 10(2), 1-37 (2019).

[10]. Kirillov, Alexander, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár and Ross B. Girshick. "Segment Anything." *2023 IEEE/CVF International Conference on Computer Vision (ICCV)* (2023): 3992-4003.

[11]. https://segment-anything.com/

[12]. Ma, Jun & He, Yuting & Li, Feifei & Han, Lin & You, Chenyu & Wang, Bo. (2024). Segment anything in medical images. Nature Communications. 15. 10.1038/s41467-024-44824-z.

[13]. https://github.com/YichiZhang98/SAM4MIS

[14]. Abedini, Mani & Bijari, Anita & Banirostam, Touraj. (2020). Classification of Pima Indian Diabetes Dataset using Ensemble of Decision Tree, Logistic Regression and Neural Network. IJARCCE. 9. 10.17148/IJARCCE.2020.9701.

[15]. Afshar, P.; Plataniotis, K.N.; Mohammadi, A. Capsule Networks for Brain Tumor Classification Based on MRI Images and Coarse Tumor Boundaries. In Proceedings of the ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 1368–1372.

[16]. Dorj, UO., Lee, KK., Choi, JY. et al. The skin cancer classification using deep convolutional neural network. Multimed Tools Appl 77, 9909–9924 (2018). https://doi.org/10.1007/s11042-018-5714-1

[17]. Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, *5*.

[18]. https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000

[19]. Ameri A. A deep learning approach to skin cancer detection in dermoscopy images. *Biomed Phys Eng* (2020) 10:801–6. doi: 10.31661/jbpe.v0i0.2004-1107

[20]. Mohapatra, S., Abhishek, N.V.S., Bardhan, D., Ghosh, A.A., Mohanty, S. (2021). Skin Cancer Classification Using Convolution Neural Networks. In: Tripathy, A., Sarkar, M., Sahoo, J., Li, KC., Chinara, S. (eds) Advances in Distributed Computing and Machine Learning. Lecture Notes in Networks and Systems, vol 127. Springer, Singapore. https://doi.org/10.1007/978-981-15-4218-3_42

[21]. Chaturvedi, S.S., Gupta, K., Prasad, P.S. (2021). Skin Lesion Analyser: An Efficient Seven-Way Multi-class Skin Cancer Classification Using MobileNet. In: Hassanien, A., Bhatnagar, R., Darwish, A. (eds) Advanced Machine Learning Technologies and Applications. AMLTA 2020. Advances in Intelligent Systems and Computing, vol 1141. Springer, Singapore. https://doi.org/10.1007/978-981-15-3383-9_15

[22]. Rezvantalab, A., Safigholi, H. and Karimijeshni, S. (2018). Dermatologist Level Dermoscopy Skin Cancer Classification Using Different Deep Learning Convolutional Neural Networks Algorithms. arXiv:1810.10348 [cs, stat]. [online] Available at: https://arxiv.org/abs/1810.10348."

DOI: 10.17148/IJARCCE.2024.13835

[23]. K. M. Hosny, M. A. Kassem and M. M. Foaud, "Skin Cancer Classification using Deep Learning and Transfer Learning," 2018 9th Cairo International Biomedical Engineering Conference (CIBEC), Cairo, Egypt, 2018, pp. 90-93, doi: 10.1109/CIBEC.2018.8641762.

[24]. T. Emara, H. M. Afify, F. H. Ismail and A. E. Hassanien, "A Modified Inception-v4 for Imbalanced Skin Cancer Classification Dataset," 2019 14th International Conference on Computer Engineering and Systems (ICCES), Cairo, Egypt, 2019, pp. 28-33, doi: 10.1109/ICCES48960.2019.9068110.

[25]. Yu L, Chen H, Dou Q, Qin J, Heng PA. Automated melanoma recognition in dermoscopy images via very deep residual networks. IEEE Trans Med Imaging (2017) 36:994–1004.

[26]. Esteva A, Kuprel B, Novoa R, Ko J, Swetter SM, Blau HM, et al.. Correction: Corrigendum: Dermatologist-level classification of skin cancer with deep neural networks. *Nature* (2017) 546:686. doi: 10.1038/nature22985

[27]. Jinnai S, Yamazaki N, Hirano Y, Sugawara Y, Ohe Y, Hamamoto R. The development of a skin cancer classification system for pigmented skin lesions using deep learning. *Biomolecules* (2020) 10:1–13. doi: 10.3390/biom10081123

[28]. Nawaz, M., Mehmood, Z., Nazir, T., Naqvi, R. A., Rehman, A., Iqbal, M., & Saba, T. (2022). Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering. Microscopy Research and Technique, 85(1), 339–351. https://doi.org/10.1002/jemt.23908

[29]. Garcia, S. I. (2021). Meta-learning for skin cancer detection using Deep Learning Techniques. https://doi.org/10.48550/arxiv.2104.10775

[30]. Goyal M, Yap MH. *"Region of interest detection in dermoscopic images for natural data-augmentation,"*. United States: arXiv; (2018) p. 1–8.

[31]. Jinnai, S.; Yamazaki, N.; Hirano, Y.; Sugawara, Y.; Ohe, Y.; Hamamoto, R. The Development of a Skin Cancer Classification System for Pigmented Skin Lesions Using Deep Learning. Biomolecules 2020, 10, 1123. https://doi.org/10.3390/biom10081123

[32]. Chaturvedi, S.S., Tembhurne, J.V. & Diwan, T. A multi-class skin Cancer classification using deep convolutional neural networks. Multimed Tools Appl 79, 28477–28498 (2020). https://doi.org/10.1007/s11042-020-09388-2

[33]. Garg, R., Maheshwari, S., Shukla, A. (2021). Decision Support System for Detection and Classification of Skin Cancer Using CNN. In: Sharma, M.K., Dhaka, V.S., Perumal, T., Dey, N., Tavares, J.M.R.S. (eds) Innovations in Computational Intelligence and Computer Vision. Advances in Intelligent Systems and Computing, vol 1189. Springer, Singapore. https://doi.org/10.1007/978-981-15-6067-5_65

[34]. Benedetti, P., Perri, D., Simonetti, M., Gervasi, O., Reali, G., Femminella, M. (2020). Skin Cancer Classification Using Inception Network and Transfer Learning. In: , et al. Computational Science and Its Applications – ICCSA 2020. ICCSA 2020. Lecture Notes in Computer Science(), vol 12249. Springer, Cham. https://doi.org/10.1007/978-3-030-58799-4_39

[35]. Reis HC, Turk V, Khoshelham K, Kaya S. InSiNet: a deep convolutional approach to skin cancer detection and segmentation. *Med Biol Eng Comput* (2022) 60:643–62. doi: 10.1007/s11517-021-02473-0

[36]. Bechelli S, Delhommelle J. Machine learning and deep learning algorithms for skin cancer classification from dermoscopic images. *Bioengineering* (2022) 9(3):97. doi: 10.3390/bioengineering9030097

[37]. A. Kirillov, E. Mintun, N. Ravi et al., "Segment anything," arXiv preprint arXiv:2304.02643, (2023).

[38]. Hu, Mingzhe & Li, Yuheng & Yang, Xiaofeng. (2023). SkinSAM: Empowering Skin Cancer Segmentation with Segment Anything Model.

[39]. Ma, J., He, Y., Li, F. *et al.* Segment anything in medical images. *Nat Commun* **15**, 654 (2024). https://doi.org/10.1038/s41467-024-44824-z

[40]. Tschandl, P., Rinner, C., Apalla, Z. et al. Human–computer collaboration for skin cancer recognition. Nat Med (2020). https://doi.org/10.1038/s41591-020-0942-0

[41]. Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci Data 5, 180161 (2018). https://doi.org/10.1038/sdata.2018.161

[42]. Noel Codella, Veronica Rotemberg, Philipp Tschandl, M. Emre Celebi, Stephen Dusza, David Gutman, Brian Helba, Aadi Kalloo, Konstantinos Liopyris, Michael Marchetti, Harald Kittler, Allan Halpern: "Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC)", 2018; https://arxiv.org/abs/1902.03368

[43]. Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci. Data 5, 180161 doi:10.1038/sdata.2018.161 (2018).

[44]. Koehoorn J, Sobiecki A, Rauber P, Jalba A, Telea A. Effcient and effective automated digital hair removal from dermoscopy images. *Math Morphol - Theory Appl* (2016) 1:1–17. doi: 10.1515/mathm-2016-0001

[45]. Glenn Jocher and Ayush Chaurasia and Jing Qiu, "Ultralytics YOLOv8", 2023, https://github.com/ultralytics/ultralytics

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13835

[46]. M. Tan, R. Pang and Q. V. Le, "EfficientDet: Scalable and Efficient Object Detection," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 10778-10787, doi: 10.1109/CVPR42600.2020.01079.

[47]. R. Del Prete, M. D. Graziano and A. Renga, "RetinaNet: A deep learning architecture to achieve a robust wake detector in SAR images," 2021 IEEE 6th International Forum on Research and Technology for Society and Industry (RTSI), Naples, Italy, 2021, pp. 171-176, doi: 10.1109/RTSI50628.2021.9597297.

[48]. R. Girshick, "Fast R-CNN," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 1440-1448, doi: 10.1109/ICCV.2015.169.

[49]. Szegedy, Christian & Vanhoucke, Vincent & Ioffe, Sergey & Shlens, Jon & Wojna, ZB. (2016). “Rethinking the Inception Architecture for Computer Vision”. 0.1109/CVPR.2016.308.

[50]. Huang, Gao & Liu, Zhuang & van der Maaten, Laurens & Weinberger, Kilian. (2017). “Densely Connected Convolutional Networks”. 10.1109/CVPR.2017.243.

[51]. Chollet, Francois. (2017). “Xception: Deep Learning with Depthwise Separable Convolutions”. 1800-1807. 10.1109/CVPR.2017.195.

[52]. Simonyan, Karen & Zisserman, Andrew. (2014). “Very Deep Convolutional Networks for Large-Scale Image Recognition”. arXiv 1409.1556.

[53]. Tan, Mingxing & Le, Quoc. (2019). “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks”.

[54]. Yu, Jiahui & Wang, Zirui & Vasudevan, Vijay & Yeung, Legg & Seyedhosseini, Mojtaba. (2022). “CoCa: Contrastive Captioners are Image-Text Foundation Models”.

[55]. CoCa model in OpenCLIP. https://colab.research.google.com/github/mlfoundations/open_clip/blob/master/docs/Interacting_with_open_coca.ipynb