


Squamous Cell Carcinoma of Skin Cancer Margin Classification From Digital Histopathology Images Using Deep Learning

Cancer Control
Volume 29: 1–16
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10732748221132528
journals.sagepub.com/home/ccx


Beshatu Debela Wako^{1,2} , Kokeb Dese^{1,3} , Roba Elala Ulfata^{4,5} ,
Tilahun Alemayehu Nigatu⁶, Solomon Kebede Turunbedu⁴, and Timothy Kwa^{1,7,8}

Abstract

Objectives: Now a days, squamous cell carcinoma (SCC) margin assessment is done by examining histopathology images and inspection of whole slide images (WSI) using a conventional microscope. This is time-consuming, tedious, and depends on experts' experience which may lead to misdiagnosis and mistreatment plans. This study aims to develop a system for the automatic diagnosis of skin cancer margin for squamous cell carcinoma from histopathology microscopic images by applying deep learning techniques.

Methods: The system was trained, validated, and tested using histopathology images of SCC cancer locally acquired from Jimma Medical Center Pathology Department from seven different skin sites using an Olympus digital microscope. All images were preprocessed and trained with transfer learning pre-trained models by fine-tuning the hyper-parameter of the selected models.

Results: The overall best training accuracy of the models become 95.3%, 97.1%, 89.8%, and 89.9% on EfficientNetB0, MobileNetv2, ResNet50, VGG16 respectively. In addition to this, the best validation accuracy of the models was 94.7%, 91.8%, 87.8%, and 86.7% respectively. The best testing accuracy of the models at the same epoch was 95.2%, 91.5%, 87%, and 85.5% respectively. From these models, EfficientNetB0 showed the best average training and testing accuracy than the other models.

Conclusions: The system assists the pathologist during the margin assessment of SCC by decreasing the diagnosis time from an average of 25 minutes to less than a minute.

Keywords

histopathological margins, squamous cell carcinoma, deep learning, transfer learning, classification, recurrence rate, reconstruction surgery

Received May 25, 2022. Received revised September 17, 2022. Accepted for publication September 26, 2022.

¹School of Biomedical Engineering, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia

²Center of Biomedical Engineering, Jimma University Medical Center, Jimma, Ethiopia

³Artificial Intelligence and Biomedical Imaging Research Lab, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia

⁴Department of Pathology, Jimma Institute of Health, Jimma University, Jimma, Ethiopia

⁵Department of Pathology, Adama General Hospital and Medical College, Adama, Ethiopia

⁶Department of Biomedical Sciences (Anatomy Course Unit), Jimma Institute of Health, Jimma University, Jimma, Ethiopia

⁷Department of Biomedical Engineering, University of California, 451 Health Sciences, Davis, CA, USA

⁸Medtronic MiniMed, 18000 Devonshire St., Northridge, Los Angeles, CA, USA

Corresponding Authors:

Kokeb Dese, Department of Biomedical Engineering, Jimma University, Jimma 378, Ethiopia.

Email: kokebdese86@gmail.com, dese.gebremeskel@ju.edu.et

Timothy Kwa, Department of Biomedical Engineering, Jimma University, Jimma 378, Ethiopia.

Email: tkwa@ucdavis.edu



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and

Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Introduction

Skin cancer is the most common type of cancer that affects humans worldwide. According to the literature out of three people diagnosed with cancer, there is a possibility of one patient with skin cancer.¹ It is a common type of cancer that starts to grow in the epidermis layer of the skin.^{2,3} The number of people affected by skin cancer will be expected to exceed 13.1 million by 2030.^{2,4} In the United States, the occurrence of skin cancer is reported to be 22.1 per 100 000 people. The number of new patients yearly predicted is expected to be more than 63 000, and skin cancer is now rated as the sixth most common of all cancers.⁴ Skin cancer is generally classified into two major groups; melanoma and non-melanoma. The frequency of non-melanoma skin cancer (NMSC), including basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) has increased from 3.4 to 4.9 million cases per year.⁵ Nevertheless, they can be fatal when it is left undiagnosed and untreated early.⁶ SCC accounts for most NMSC-related metastatic cancer diseases and death. According to Ref. 1 it is the second most frequent kind of skin cancer. Generally, scaly red spots, open sores, raised growths with a central depression, or warts are frequent signs of SCC. Nevertheless, there are three differentiation stages of SCC. These are (see Figure 1); (1) well-differentiated SCC, (2) poorly differentiated SCC, and (3) undifferentiated/or invasive SCC. The well-differentiated SCC is characterized by having one property of grade I SCC. The poorly differentiated property of cells indicates grade II and grade III SCC. Grade IV SCC can be characterized as an invasive/undifferentiated type. Visual inspection,, and histopathology are the current diagnostic methods for surgeons to differentiate between tumor and normal tissue for skin cancer including SCC. Of these techniques, histopathology diagnosis is the gold standard method used not only to identify its type but also for grading and diagnosing/assessing the tumor margins.⁷ The differential diagnosis between SCC histologic grades is crucial, as it will further determine the therapeutic approach and follow-up of the tumor.⁸ However, in this research, since we are focusing on the marginal diagnosis of the tumor, we consider all grades as malignant and the cancer-free margin as normal/or benign.

As shown in Figure 1 above, in well-differentiated tumors, the cells are organized and have a shape that has been usually seen in normal tissue images. The poorly differentiated cells are looking disorganized when seen under the eyepiece of the microscope, and tend to grow and spread faster than grade I tumors, ie, the well-differentiated ones. Those SCC tumor cells, which are not differentiated, look highly disorganized and spread more tremendously than the poorly differentiated categories.

Therefore, early detection of skin cancer margin is required to prevent the progression of cancer to advanced stages and reduce cancer fatality. Nowadays, SCC is clinically diagnosed using dermoscopic examination and tissue biopsy followed by Mohs micrographic surgery (MMS).^{6,9} Among these, biopsy tests are the gold standard method in the diagnosis procedure of SCC. After diagnosis, for treatment planning, the surgical excision is the routinely used method for all SCC treatments, followed by a histopathological margin assessment of all ribs of the tumor. This would help for the confirmation of the total removal of the tumor cells.¹⁰ A cancer margin, as defined by the National Cancer Institute (NIH), is “the edge or border of the tissue removed in cancer surgery”.¹¹ If the margin is assessed correctly, this border surrounds the cancerous tissue as well as a rim of normal tissue to later confirm a successful resection. Histopathological assessment of surgical margin is performed by analyzing by taking sample tissue from all margin and examining it under the microscope. Surgery can cure ~45% of all patients with cancer,⁵ however, in 40% to 50% of cases a remaining tumor cell is found at the margins,⁷ and extra surgery is required, which results in sophisticated treatment, high cost, greater morbidity, infection risk, and late therapy.¹² Unfortunately, up to 39% of the patients who experience, surgery leave the operating room without a complete resection due to positive or close margins.

The manual histopathology, which is based on the conventional microscope margin assessment method, is a time-consuming and tedious process. The accurate decision of the margin diagnosis needs an experienced pathologist. Sometimes it may require the decision of two or more experts to provide a reliable pathology report, which directly affects the delay of the treatment plan and cure rate. The current

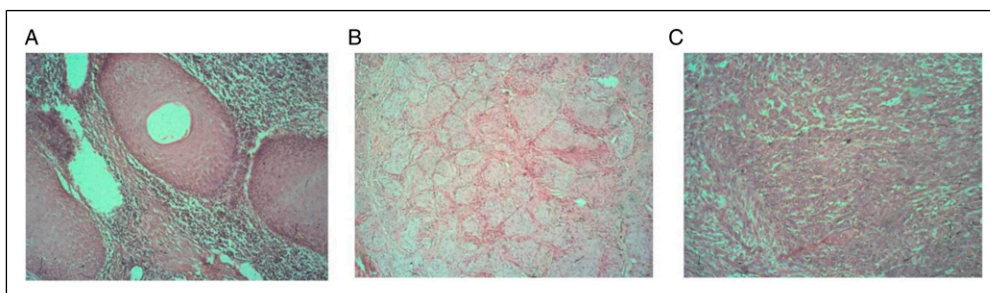


Figure 1. Sample histopathology SCC images acquired from Jimma University Medical Center (a) Well-differentiated SCC, (b) Poorly differentiated SCC, (c) Undifferentiated/or invasive SCC. Abbreviation: SCC, squamous cell carcinoma.

procedural protocol for any skin cancer-related treatment in Ethiopia is the removal of the tumor part and waiting for a pathology report for the complete removal of cancer. The report took more than a month.¹²⁻¹⁴ A current topic of research focuses on creating computer-aided diagnostic (CAD) systems for skin lesions, intending to help dermatologists by reliably analyzing histopathology images of skin lesions for automated identification of SCC.

Related Works

To date, various image processing and machine learning techniques have been used to diagnose the SCC margin. However, the accuracy of the developed system was not sufficient most probably due to the use of few data sets only from online sources and use of Most recently, M. Halicek et al¹⁵ proposed the studies on hyperspectral imaging (HSI) and fluorescent imaging of head and neck SCC in fresh surgical samples from 102 patients/293 tissue samples. HIS was captured using Maestro spectral imaging system. The autofluorescence images were acquired from 500 to 720 nm in 10 nm increments to produce a hypercube of 23 spectral bands using autofluorescence-imaging modality. They used a deep learning method of Inception V4 transfer learning to classify the whole tissue specimens into cancerous and normal. In this study two experiments were performed. The first experiment consisted of training the CNN on the primary tumor (T) and all normal (N) tissues while testing on T and N tissues from other patients. The second experiment consisted of training on the primary tumor (T) and all normal (N) tissues while testing only tumor-involved cancer margin (TN) tissues from other patients. HSI detected conventional SCC in the larynx, oropharynx, and nasal cavity with .85-.95 AUC score, and autofluorescence imaging detected HPV+ SCC in tonsillar tissue with .91 AUC score for different organ sites. Generally, the result shows that AUCs upwards of .80-.90 were obtained for SCC detection with HSI-based. Again another study in Ref. 16 which was written by M. Halicek et al shows the ability of HSI-based cancer margin detection for oral cancer of thyroid cancer and oral SCC. The CNN-based method classifies the tumor-normal margin of oral squamous cell carcinoma (SCC) vs normal oral tissue with an area under the curve (AUC) of .86 with 81% accuracy, 84% sensitivity, and 77% specificity. In the same study, thyroid carcinoma cancer normal margins were classified with an AUC of .94 for interpatient validation, performed with 90% accuracy, 91% sensitivity, and 88% specificity. This study compared support vector machine (SVM) with radial basis function (RBF) type kernel and CNN deep neural network model to classify SCC, and .80 and .85 AUC were achieved by the models respectively. In Ref. 7 L. Ma et al proposed, that a fully convolutional network (FCN) model based on U-Net architecture was implemented and trained for tissue classification in hyperspectral images (HIS) of 25 ex vivo SCC surgical specimens from 20 different patients. They used only patches containing the

tumor-normal margin to train the model, while the patches with only tumor or only normal tissue were not used in the training process. The model was evaluated per patient and achieved pixel-level tissue classification with an average area under the curve (AUC) of .88, as well as .83 accuracy, .84 sensitivity, and .70 specificity. Kassem. M.A et al¹⁷ proposed Skin Lesions Classification Into Eight Classes for ISIC 2019 Using Deep Convolutional Neural Network and Transfer Learning. This paper proposes a model for highly accurate classification of skin lesions. The proposed model utilized the transfer learning and pre-trained model with GoogleNet. The proposed model successfully classified the eight different classes of skin lesions, namely, melanoma, melanocytic nevus, basal cell carcinoma, actinic keratosis, benign keratosis, dermatofibroma, vascular lesion, and Squamous cell carcinoma. The achieved classification accuracy, sensitivity, specificity, and precision percentages are 94.92%, 79.8%, 97%, and 80.36%, respectively. They used online datasets to train and test their models.

L. Zhang et al¹⁴ proposed a deep learning-based stimulated Raman scattering (SRS) microscope of laryngeal squamous cell carcinoma on fresh surgical specimens using a 34-layered residual convolutional neural network (ResNet34) to classify 33 fresh surgical samples into normal and neoplasia to diagnosis the abnormality of the samples. Even though they modeled the system with high accuracy (100%) for the classification of samples into normal and neoplasia, margin assessment was not addressed. On the other hand, Khalid M et al in Ref. 18 proposed Classification of Skin Lesions into Seven Classes Using Transfer Learning with AlexNet. The parameters of the original model are used as initial values, where they randomly initialize the weights of the last three replaced layers. The proposed method was tested using the most recent public dataset, ISIC 2018. Based on the obtained results, they could say that the proposed method achieved great success where it accurately classifies the skin lesions into seven classes. These classes are melanoma, melanocytic nevus, basal cell carcinoma, actinic keratosis, benign keratosis, dermatofibroma, and vascular lesion. The achieved percentages were 98.70%, 95.60%, 99.27%, and 95.06% for accuracy, sensitivity, specificity, and precision, respectively. In Ref. 19 B. Fei et al proposed a machine learning-based quantification method for HIS data from 16 patients, who underwent head and neck surgery used for binary classification as cancer normal tissues. They used normal and tumor tissues for training and the model were evaluated on the histopathology of tumor-normal interface from the same patients. The study classifies the normal and cancer tissues but not on the boundary of the tumor margin. They got distinguished of $90\% \pm 8\%$ accuracy, $89\% \pm 9\%$ sensitivity, and specificity of $91\% \pm 6$. The above-mentioned studies used hyperspectral imaging (HSI) modalities for the peripheral margins, which has a limitation on the deep penetration of the deep margins where the most positive margin cases were reported. Starting with the primary clinical samples obtained

from the Jimma Medical Center (JMC), Department of Pathology, histopathology images tainted with typical artifacts such as fringing dust, and non-collimated lighting were acquired using a locally available microscope. Our setup closely resembled a clinical microscope that is often seen in resource-poor hospital settings. The images were then preprocessed to remove the artifacts and increase the number of trained data sets. Different transfer learning and deep learning artificial intelligence-based models were applied and their classification performance was compared.

Proposed Models

The acquired microscopic histology images often contained artifacts from diverse sources that needed to be rectified using appropriate preprocessing methods. Therefore, this section explicates the details of the image acquisition and image processing techniques required for the margin classification, followed by a brief discourse on the transfer learning methods used in this work. The overall workflow/block diagram used for developing the system is outlaid in Figure 2.

In this research, four models have been selected and trained with the locally collected SCC data sets. These models were selected due to their outperforming in related works. These were, VGG16, ResNet-50, MobileNetV2, and EfficientNetB0. A detailed explanation of each model is found in Supplementary Material 1.

Experimental Design

Data Collection/Image Acquisition

In collaboration with the Pathology, Histology, and Dermatology Departments at Jimma University Medical Center, tissue samples were obtained from skin cancer surgical resection. The tissues were obtained from different skin parts (legs, feet,

hands, toes, eyes, face, and neck) of the patient (see Table 1) for SCC, which is the most abundant and most frequently diagnosed skin cancer type in Jimma University Medical Center (JUMC). The tissue images were acquired using a digital compound light microscope (Olympus, CX21FS1, Guangzhou, China) equipped with a $\times 100$ oil immersion objective and a $\times 10$ eyepiece magnification integrated with a camera of 5MP digital resolution (see Figure 3(a)). For a given slide (see Figure 3(b)), a magnification of $\times 10$ was used in the image acquisition of the histopathology image (see Figure 3(c) and (d)). To do this, the tissue biopsies were processed via formaldehyde Xing and paraffin embedding (FFPE) and cut into thin sections. Finally, it was stained with hematoxylin and eosin (H&E) to observe the structure of the cells (see Figure 3(b)).

The safest margin for surgical resection of different cancer types is different based on the tumor resection margin standards of the providers.^{20–23} For the oral tongue, a negative margin was proposed to be 2.2 mm. Another study found cuts within 1 mm of oral cavity tumor margins are associated with significantly increased recurrence rates. Negative resection margins are the primary prevention of disease relapse of the cancer cells.^{16,24} For this study, based on the JUMC standard of care for skin cancer histopathology margin assessments, more than 1 mm surgical margin is considered as a margin negative, and less than 1 mm is considered as a margin positive. Taking^{19,22} as a reference three regions of interest were selected and images were acquired in this study: the tumor, normal, and tumor-normal interface regions.

The collected slides (see sample slides in Figure 3(b)) were from 50 patients. The number of patients distributed for each organ was: 12 patients with SCC of the legs, 8 on hands, 3 on the eyes, 14 on feet, 6 on toes, 4 on the neck, and 3 on the face. Regarding histologic grading, 17 patients with well-differentiated SCC and 15 patients with poorly differentiated SCC, 18 patients were Invasive SCC as stipulated in

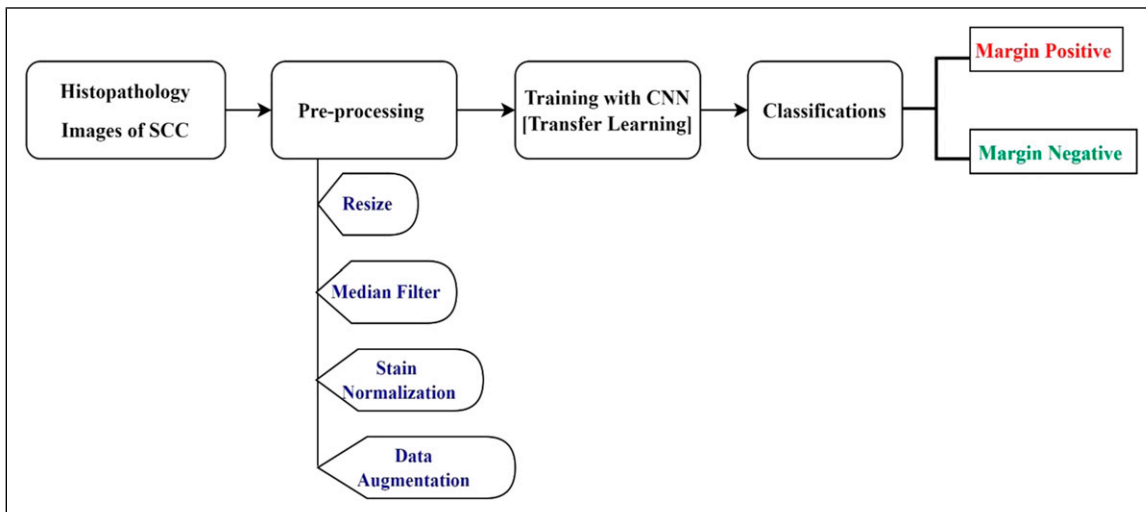


Figure 2. The general diagram of the proposed system.

Table 1. Squamous Cell Carcinoma Data Set the Information of Patients and Whole Slide Images.

Site of Dataset Information	Number of Patients	Normal (WSI)	Tumor (WSI)	Tumor-Normal (WSI)
Leg	12	104	80	56
Hand	8	60	58	36
Foot	14	101	88	56
Toe	6	46	4	24
Eye	3	17	18	13
Neck	4	8	4	8
Face	3	9	9	6
Total	50	345	284	199
Based on histological grading				
Well-differentiated	17	110	82	67
Poorly-differentiated	15	112	95	60
Invasive	18	123	10	72
Total	50	345	284	199

Abbreviation: WSI, whole slide images.

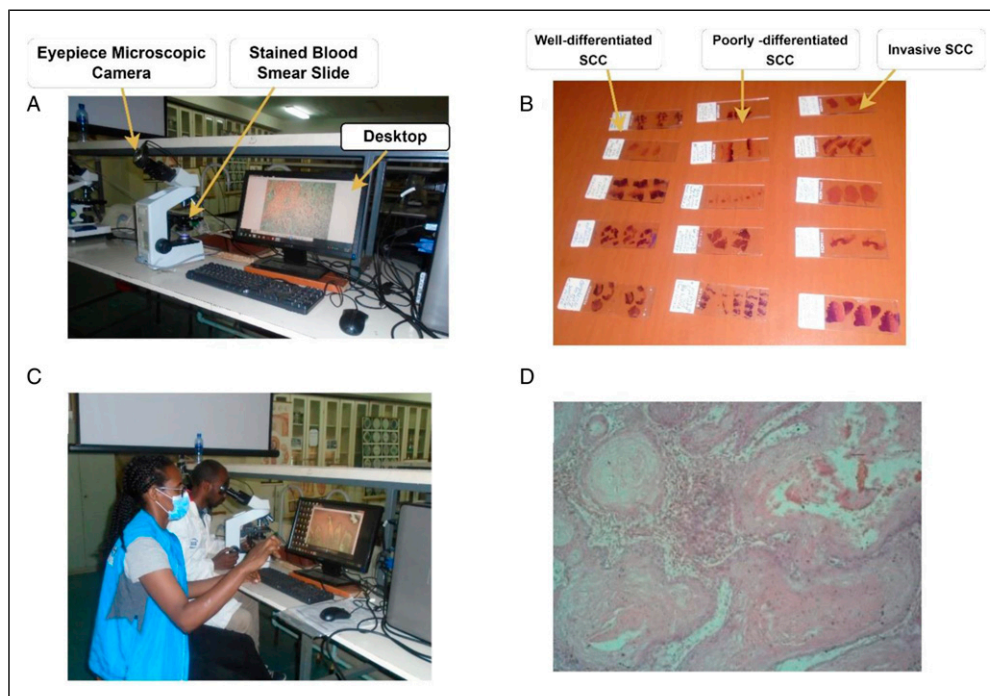


Figure 3. Data acquisition procedure in Jimma University Medical Center pathology department. (a) The setup used for image acquisition, (b) Shows sample slides with SCC, (c) during the image acquisition, (d) sample acquired well-differentiated SCC histopathology image. Abbreviation: SCC, squamous cell carcinoma.

Table 1. Tissue samples that are entirely normal were used as Margin Negative and the sample that contains tumor-normal margins and entire tumor were used as Margin Positive category. All H&E-stained histopathology images were labeled as margin negative and margin positive and confirmed by 2 (two) pathologists for histopathologic assessment. Finally, both pathologists and histologists validated the correct labeling of the captured slide images, which were used as our acquired data used for developing our model. In this research, a total of three 345

normal, 284 images for tumor, and 199 for a tumor-normal section of histopathology images were originally acquired (see sample acquired image in [Figure 3\(d\)](#)).

From [Table 1](#) above, out of 50 patients originally 345 margin negative and 483 margin positive (the combination of pure tumor and tumor-normal section) histopathology images were acquired. Seven different skin organs and three histologic grades of SCC were used aiming to use the models for most skin parts of the body. As the research did not involve the

direct use of humans, animals or other subjects, a formal ethics approval was not required for this study. This was checked and confirmation for this was received from the Jimma University's institutional review board.

Image Preprocessing

The acquired images usually contained noise due to excessive irregularities arising from the staining procedure. On the other hand, the number of originally acquired images could be not enough to train our model. Thus, the purpose of preprocessing is to improve image quality by removing unwanted objects and noise from histopathology images and increasing the number of images by applying different image augmentation techniques.^{25,26} In the preprocessing step, the following methodology was adopted.

1. **Resize:** Deep learning models are computationally expensive and require all input images to have the same size. Therefore, to decrease the computational

time,^{20,27} the original Red Green Blue (RGB) image (2048×1536) was reduced to 224×224 pixels (see [Figure 4](#)).

2. **Image Smoothing:** during image capturing of microscopic images, it could be susceptible to different noises, such as additive, random, impulsive, and multiplicative are normally associated with any image. Noise deletion is most important in medical image analysis.²⁸ The most frequently affected noises in the medical images are Gaussian, pepper, speckle, and Poisson noises. As compared with other filters, in this research, a median filter was used to remove the salt and pepper noise in the whole slide image. One of the major advantages of the median filter is that it strongly preserves the edges of an image²⁹ (see [Figure 5](#)).
3. **Stain Normalization:** color normalization is an important preprocessing task in the whole-slide image (WSI) of digital pathology.^{30,31} It refers to standardized color distribution across input images and focused on hematoxylin and eosin (H&E) stained slides. Color

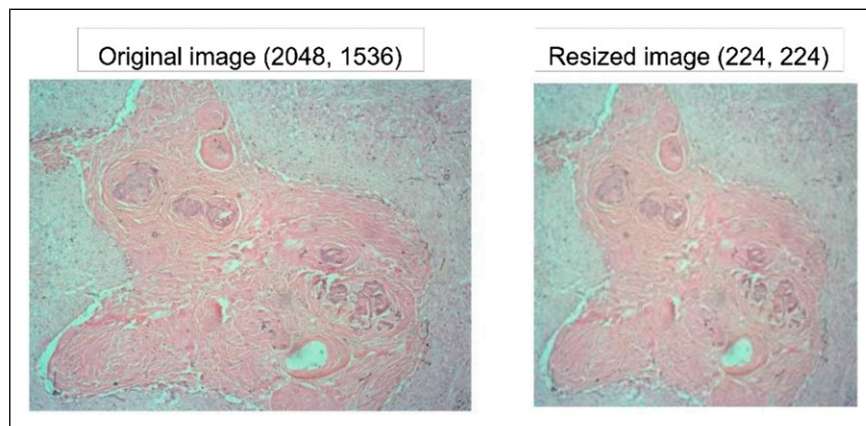


Figure 4. Original and resized image.

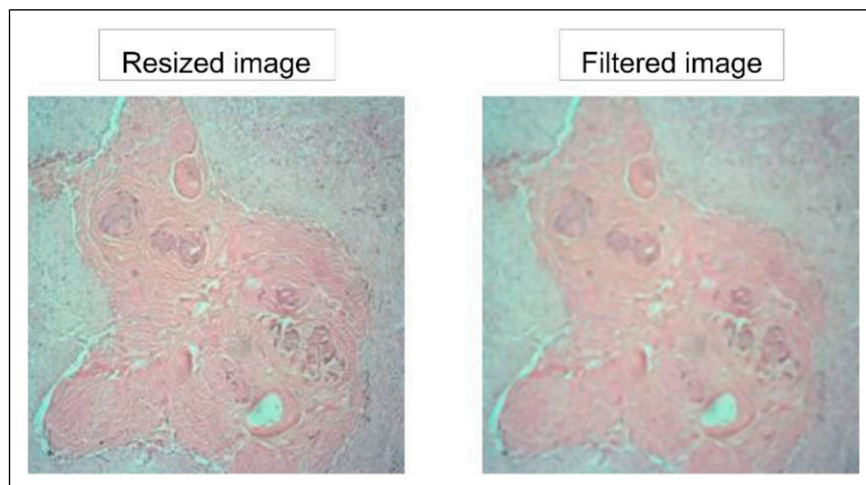


Figure 5. The original resized image and the median filtered image.

normalization techniques like stain normalization are an important processing task for computer-aided diagnosis (CAD) systems³² which is achieved by normalizing the stains for enhancement and reducing the color and intensity variations present in stained images from different laboratories, consequently, increasing the estimation accuracy of CAD systems.³⁰ In this study, a Macenko stain normalization algorithm, which was popular in histopathology slides³²⁻³⁴ was used (see Figure 6).

4. Data Augmentation: It is a method used to significantly increase the amount and variety of data available for training models.^{28,35,36} Data augmentation was performed by rotating the images in 90°, 180°, 270°, horizontal flip, and vertical flip to increase the available data without affecting their features. As result, the number of data was increased by six times.

Model Training

The obtained original data was split into 80% for training, 10% for validation, and 10% for testing through a stratified cross-validation method. This means out of 828 originally acquired images, 662 were used for training, 82 for validation, and 84 for testing purposes. After augmentation of 6× (with 90°, 180°, 270°, horizontal flip, and vertical flip), the number of images in each class becomes 1656 for Margin Negative, and 2316 for Margin Positive excluding the testing data set, which needs to be the original dataset and is 84 (35 for MN and 49 for MP) images. Therefore, the training, validation, and testing data classes contain 3972, 492, and 84 images, respectively. To train the models for the SCC classification task, utilizing the concept of transfer learning,^{37,38} the actual classifier was replaced (1000 nodes) in each pre-trained model with a new

one (sigmoid layer with 1 node) for binary classification of SCC images.

During training, the bottom layers were kept fixed (frozen) and not retrained (using the weight values from a pre-trained model or it was already trained), while a few top layers (dense layers or fully connected layers) and the appended classifier (activation function (sigmoid) that delivers an output classification and sigmoid is mostly used for binary classification). Since training from scratch is computationally expensive and requires a large amount of data to achieve high performance we applied the concept of transfer learning by adjusting the parameters such as a learning rate, the number of epochs, and the optimizer, to achieve the best possible results (see Tables 2 and 3).

Taking a pre-trained deep neural network (VGG 16, Resnet 50, Mobile net v2, Efficient net B0) as a feature extractor and freezing the weights for the convolutional layers in the network. The last three layers have been replaced with a new fully-connected, sigmoid, and 2 classification output layers on top of the body of the network.

After operating on several trials and testing with different transfer learning pre-trained models, we have selected four models and compared their results. These were (1) the visual geometry group (VGG16), (2) Residual Network (ResNet50), (3) EfficientNetB0 and MobileNetV2.

The network architecture of VGG16 is a sixteen-layer deep CNN. It consists of thirteen convolution layers arranged into five blocks, each followed by a pooling operation. The network uses filters of size 3×3 for convolution and 2×2 size windows for pooling operation. The convolutional stack is followed by two fully connected layers, each consisting of 4096 nodes. The final layer is a SoftMax layer that assigns a class to each image.³⁷ The residual network (ResNet50): has a depth of fifty (50) layers, forty-eight (48) convolutions, one

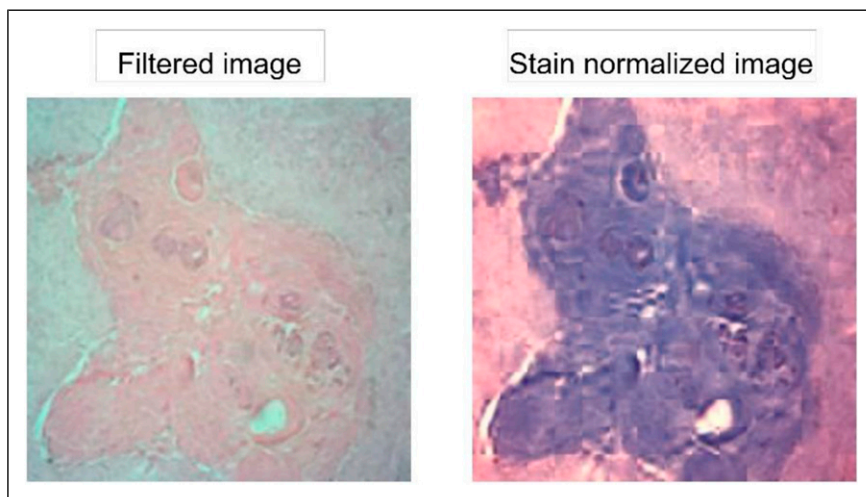


Figure 6. The median filtered image and stained normalized image.

Table 2. Fine-Tuning Made on the Layers of the Model.

Models	Frozen Convolutional Layers (Fixed Layers)	New-Top Layer	Output Features Extracted	Input Features for the Classifier	Classifier Output
VGG 16	13 convolutional layers	Last three layers	25 088	256	2
Resnet 50	48 convolutional layers	Last three layers	2048	256	2
Mobile net v2	52 convolutional layers	Last three layers	1280	256	2
Efficient net B0	81 convolutional layers	Last three layers	1280	256	2

Table 3. Functions and Parameters Used for Each Model During the Training.

Function/Parameter	EfficientNetB0	MobileNet V2	ResNet 50	VGG16
Classification function	Sigmoid (binary)	Sigmoid	Sigmoid	Sigmoid
Optimizer	Adam	Adam	Adam	Adam
Loss function	Binary-cross entropy	Binary-cross entropy	Binary-cross entropy	Binary-cross entropy
Epochs	30	50	100	70
Early stop	10	10	10	10
Learning rate	10^{-3}	10^{-3}	10^{-3}	10^{-3}
Batch size	64	64	64	64

max-pooling, and one average pooling and 3 times deeper than VGG-16, having less computational complexity.³⁷ The residual addresses the problem of training a really deep architecture by introducing an identity skip connection, which is also called a shortcut jump over layer.³⁹ On the other hand, an EfficientNetB0, which is an Efficient Net family a newly developed classifier, uses a compound scaling approach with fixed ratios in all three dimensions to maximize speed and precision and shows enormous results in this study⁴⁰ and does not change the layer operation in the baseline network while scaling. Furthermore, MobileNetV2 is having bottleneck layer in the residual connections. Lightweight depth-wise convolutions are used by the intermediate expansion layer to filter features as the source of nonlinearity. MobileNetV2 is having 32 filtered initial fully connected convolutions.³⁹

In this research, different hyper-parameters of the model were fine-tuned to increase the performance of our developed module while it was trained with the modified models. These include choosing the right optimizer, adjusting the learning rate, and choosing the appropriate activation and loss function. The following Table 3 shows the functions and parameters used for the models during the training.

As an optimizer, the Adam optimizer was chosen for its best performance in terms of speed to converge faster and accuracy.³⁷ The number of epochs used was different based on the models, while the learning rate was set to .0001 and the activation function used was ReLu. The loss function for binary class classification was binary cross-entropy.

Performance Evaluation Metrics

To evaluate the performance, we calculated accuracy, precision, recall, F1-score, specificity, and AUC value. These statistical metrics are based on True Positives (TP), False Negatives (FN), False Positives (FP), and True Negatives (TN). Here, TP and TN represent the number of correctly identified margin positive and margin negative images, respectively, while FP and FN denote the number of margin negative images wrongly classified or accepted as margin positive and the number of margins positive images incorrectly classified as margin negative respectively.^{27,37} All equations from equations (1)-(5) were taken from Ref. 41.

1. Accuracy: the accuracy scores tell how often the models produced correct results and it is calculated using equation (1) below

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

2. Precision: it simply shows “what number of selected data items are relevant”. In other words, out of the observations that the algorithm has predicted to be positive, how many of them are positive is calculated by precision. In other words, precision reflects a model’s consistency concerning margin positive outcomes. Precision is calculated based on the following equation (2)

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

3. Recall: it presents “what number of relevant data items are selected”. It indicates out of the positive observations, how many of them have been predicted by the algorithm. According to equation (3), the recall equals the number of true positives divided by the sum of true positives and false negatives: Recall calculates the ratio of correctly identified Margin Positive images to all Margin Positive images in the test data (see equation (3)).

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

4. Specificity: determines how much it classifies the Margin Negative images correctly (see equation (4)).

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (4)$$

5. F1 score: The F1 score represents a weighted average of precision and recall (equation (5)).

$$\text{F1 Score} = 2 * \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

6. ROC-AUC score: This metric is calculated using the ROC curve (receiver operating characteristic curve) which represents the relation between the true positive rate (sensitivity or recall) and false positive rate (1-specificity). Area Under ROC Curve or ROC-AUC is used for binary classification and demonstrates how good a model is in discriminating between positive and negative target classes. Especially, in our case, the importance of margin positive (reduced recurrence) and margin negative (organ conservation) classes are equal for us, ROC-AUC score can be a useful performance metric.^{37,42} Receiver Operating Characteristic (ROC) plots TP rate (equation (6)) vs FP rate (equation (7)) and helps us understand the relationship between correctly classified Margin Positive and misclassified Margin Negative images. The area under the curve (AUC) is a scalar value ranging between 0 and 1 and represents how well our model differentiates between Margin Negative and Margin Positive images.

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$\text{FPR} = \frac{\text{FP}}{\text{TN} + \text{FP}} = 1 - \text{Specificity} \quad (7)$$

An excellent model has an AUC near 1 which means it has a good measure of separability. A poor model has an AUC near 0, which means it has the worst measure of separability.

Results

Training and Validation Results

In this study, a binary classification for the Histopathology Margin of SCC was established. As per the data split ratio used, the amount of data for training the models was 1656 for Margin Negative (MN) and 2316 for Margin Positive classes (MP). Totally, 3972 images have taken as a training set and 490 for validation (204 for Margin Negative (MN) and 288 for Margin Positive (MP), and 84 for testing (34 images for Margin Negative (MN), 48 images for Margin Positive (MP)) were used.

During training, the performance of the validation group was calculated and monitored. The optimal operating threshold was calculated for the validation group for generalizable results, and it was used for generating performance evaluation metrics for the testing group. The early stop trigger would activate when validation loss did not improve for 10 consecutive epochs. In this case, the training phase would stop. Therefore, the best loss value saved and best validation loss would be achieved for the optimal operating threshold. Generally, the training process is monitored by ‘best loss’ which quantized the error between algorithm output and a given target value, and the validation accuracy and training accuracy in this best loss would be gained. After the end of the training, the best model or best checkpoint is saved based on this the saved model is loaded and can be tested using a testing dataset that is independent of the training and validation data set. In this study, we used stratified cross-validation. So, a 10-fold cross-validation was performed, splitting all datasets into 80% for training, 10% for validation, and 10% for the testing group. To reduce bias in the experiment, the fully independent testing group was only classified a single time at the end of the experiment with 84 images, after all, network optimization had been determined using the validation set. Different models (as shown in Figure 7) were trained and tested. From those models, Four (4) models with higher accuracy and AUC were selected. VGG16, ResNet 50, Mobile Net v2, and Efficient Net B0 were the selected models. Finally, the learning and generalizability performance of the models was measured using a learning curve.

The experimental results demonstrate that the application of Efficient Net B0 to the dataset of SCC considerably improves the overall performance and thus achieves the best outcome compared to other convolutional neural networks.

The following Figure 8 shows the training and validation accuracy for the four (4) selected models (VGG16, ResNet 50, Mobile Net v2, and Efficient Net B0).

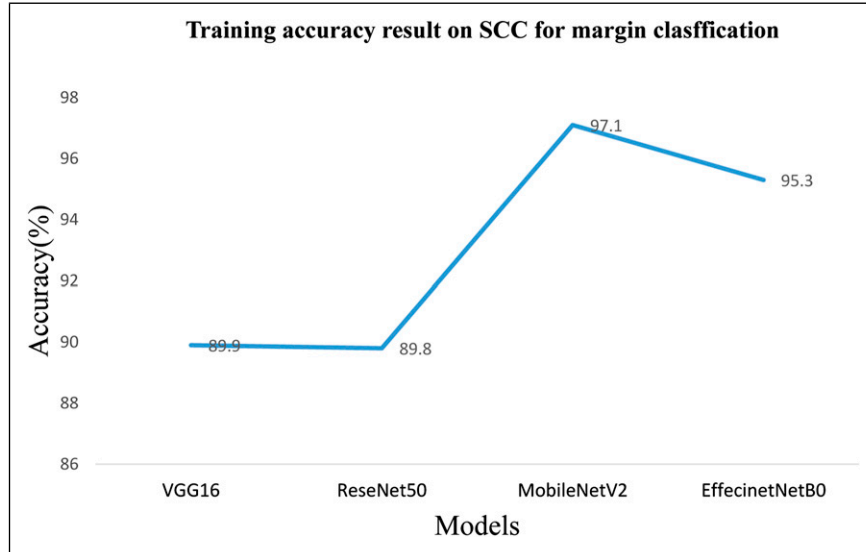


Figure 7. Different models' training accuracy on squamous cell carcinoma data set.

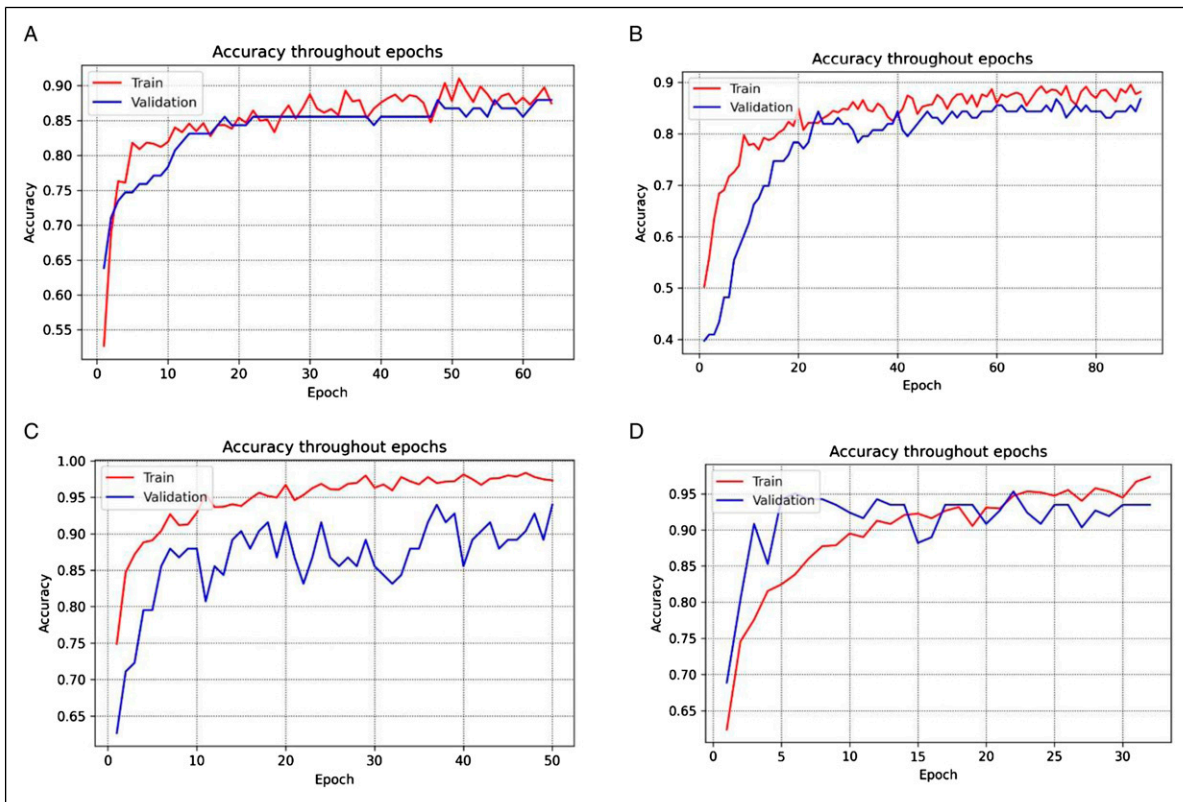


Figure 8. Training and validation accuracy for (a) VGG16, (b) ResNet 50, (c) Mobile Net v2, (d) Efficient Net B0.

The train learning curve is calculated from the train data set. It shows how well the model is learning while the validation learning curve is calculated from a hold-out validation data set to see how well the model is generalizing. For the selected models, their trained learning curves are good for all of them

and the validation learning curves were good for the three models (VGG16, ResNet 50, and Efficient Net B0). Mobile Net v2 has less generalization on the validation data set. Table 4 shows the testing and validation best accuracy of the models' weight values acquired at different epochs.

Testing Results

The performance of the models was tested on 84 images; with 35 margins negative and 49 margins positive, respectively, obtained from the originally collected data. The confusion matrix in Figure 9 shows the performance of each model on the test data.

Once the confusion matrix is done, the TP, TN, FP, and FN values are easily known. From those values, the overall precision, recall, specificity, f1-score, and test accuracy were calculated and their result is seen in Table 5 below. The following table shows the overall training results for the selected network architectures for SCC margin classification.

As depicted in Table 4 above, among the four (4) models used, the EfficientNetB0 model achieved the best performance.

On the other hand, the performance of the model can be evaluated using receiver operator characteristic (ROC) Curves, which are a useful tool to predict the probability of binary outcomes and describe how well the model is at distinguishing the classes. The Area Under the Curve (AUC) is a measure of the ability of a classifier to distinguish between Margin Negative and Margin Positive and is used as a summary of the ROC curves. Figure 10 illustrates the ROC curve generated using SCC histopathology images for histopathology margin classification with average values of AUC, 90.5%,94%,95%,100% for VGG16, ResNet 50, Mobile Net v2, Efficient Net B0, respectively.

As indicated in Figure 10 above, for all models used in this research, EfficientNetB0 outperforms with the highest AUC and the best performance of the model in distinguishing the margin positive and margin negative classes with 100%.

Table 4. The Models Have Saved the Best Weight Values Acquired at the nth Epoch.

Models	Validation Loss (nth Epoch)	Validation Loss (Value)	Validation Accuracy (%)	Training Accuracy (%)
VGG16	54	.297	86.7	89.9
ResNet-50	70	.278	87.8	89.8
Mobile Net v2	38	.17	91.8	97.1
Efficient Net B0	22	.159	94.7	95.3

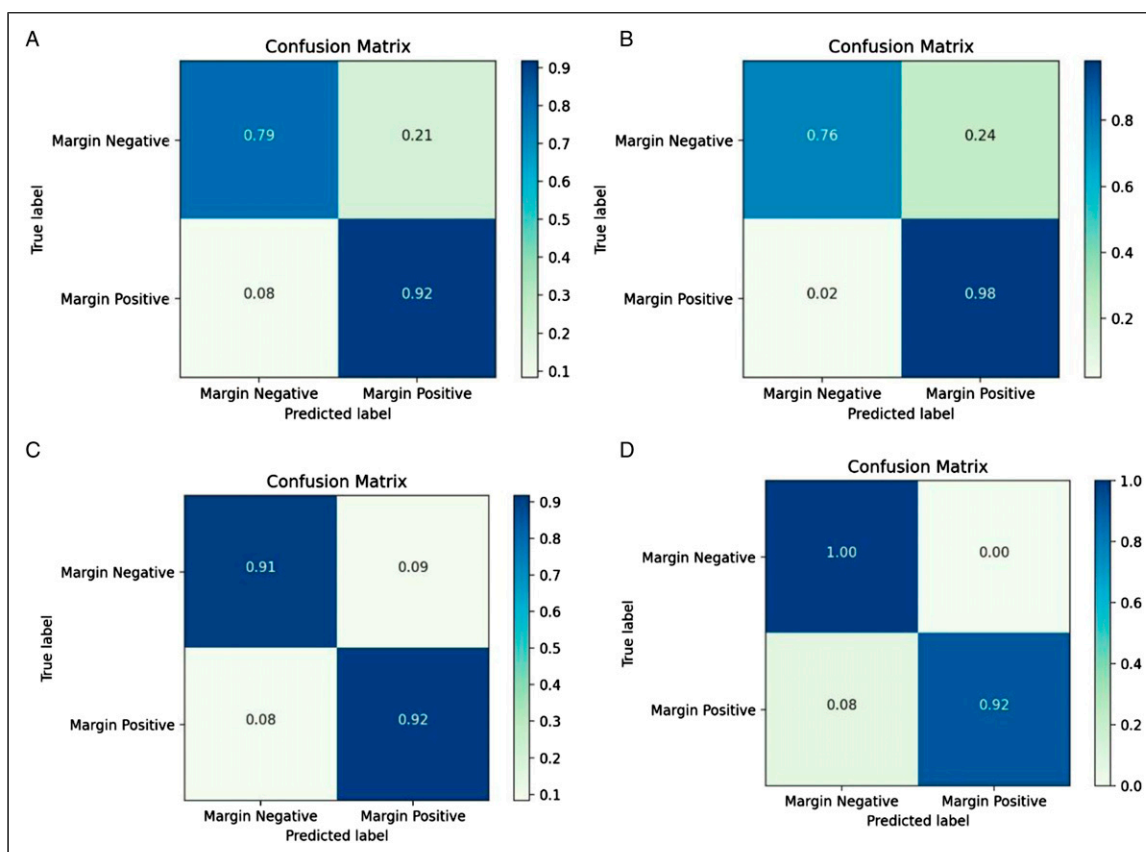
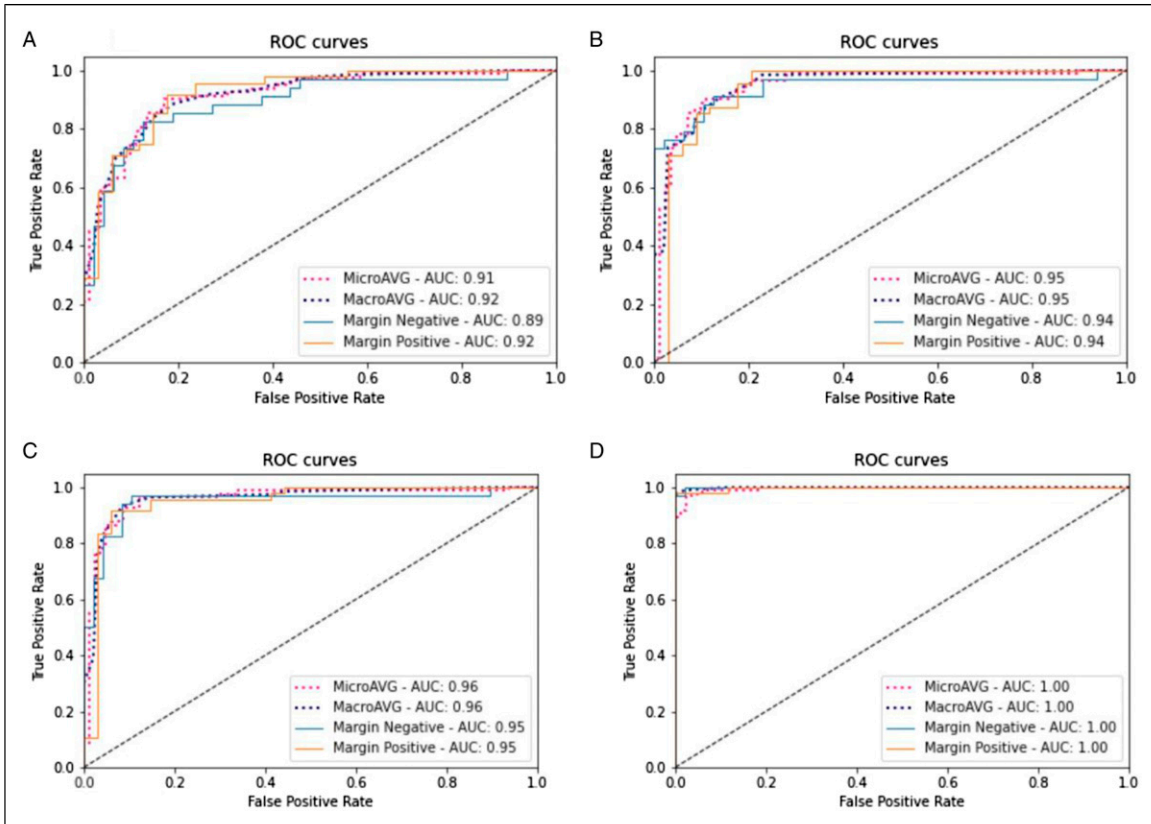


Figure 9. The normalized confusion matrix for the (a) VGG16, (b) ResNet 50, (c) Mobile Net v2, (d) Efficient Net B0 models.

Table 5. Models |Testing Performance Results Summary.

Models	Accuracy (%)	Precision (%)	Recall (%)	FI-Score (%)	Specificity (%)	Area Under the Curve%
VGG16	85.5	87	86	86	86	90.5
ResNet 50	87	91	87	88	87	94
MobileNetV2	91.5	91.5	91.5	91.5	95	95
EfficientNetB0	95.2	95	96	95	96	100

**Figure 10.** Receiver operating characteristic curve and area under the curve value for (a) VGG16, (b) ResNet 50, (c) Mobile Net v2, (d) Efficient Net B0 models.

Discussion

This work focuses on a deep learning-based SCC diagnosis system. The developed system shows the promising result of replacing the currently existing manual diagnosis methods with an automated system. Skin cancer SCC can be diagnosed by clinical examination, including visualization,⁶ optical imaging technique, and histopathology (biopsy) tests. Among these, the histopathology test is the gold standard and the most common technique used to identify cancer types and classify the grade, and margin status of the tumor margin in low resource settings.⁹ The most preferable treatment for SCC is the surgical removal of the entire tumor tissue, followed by margin assessments¹⁰ which can help the surgeon repeatedly

operate the margin removal process until margin free report is gained and proceed to the next step for reconstruction surgery, which is depending on pathologist margin status reports. Unfortunately, there is a shortage of pathologists in most developing countries and health care providers, including Ethiopia. The complexity of margin assessments and their subjective decision, which depends on the expert's experience, leads to misdiagnosis and local recurrence of the cancer cells.

The major aim of this study was to classify SCC histopathological images as Margin Negative and Margin Positive to classify the histopathological surgical margin. To achieve this, four different models were developed. The best result was achieved by fine-tuning the pre-trained model of EfficientNetB0.

Table 6. Comparing the Proposed Method With Others.

Authors	Preprocessing	Data Size and Site	Model Used	Modality/Output Results	Accuracy (%) / AUC
Proposed method	-Median filter -Stain normalization -Normalization	828 images/seven sites, foot, leg, eye, hand, toe, face, and neck	VGG16, ResNet-50, MobileNetV2, EfficientNetB0	Compound light microscope/ binary classification	95.3% training and 95.2% testing accuracy with EfficientNetB0 model
L. Ma et al (2021) ⁷		Squamous cell carcinoma/ hypopharynx, larynx	U-net architecture	Maestro spectral imaging/binary classification	AUC of 88% accuracy, 83% sensitivity 84%, specificity 70%
A. R. Triki et al (2017) ¹²	-Sobel edge detector -Gaussian filter	Breast	LeNet (CNN)	OCT/Binary classification	90% accuracy
J. D. Dorm et al (2019) ¹⁵	—	293 tissues samples/ head and neck	Inceptionv4	Fluorescent imaging/Binary classification	80-90% AUC
M. Halicek et al (2018) ¹⁶	—	—	CNN-based method	Maestro spectral imaging/Multi-class classification	SCC: (AUC) of 86% with 81% accuracy, thyroid: AUC of 94% 90% accuracy
E. Kho et al (2019) ⁴³	Spectral normalization	18 patients	SVM	Maestro spectral imaging/Binary classification	88% accuracy
B. Fei et al ¹⁹	Data normalization was to remove the spectral nonuniformity	16 patients/head and neck	—	Maestro spectral imaging/binary classification	Average accuracy of 90% ± 8%

Abbreviations: SVM, support vector machine; AUC, area under the curve.

As shown in the testing result confusion matrix in [Figure 9](#), ResNet50 classifies the margin positive 98% with the best results, and Efficient Net B0 equally classifies the margin positive as that of ResNet50. VGG16 is about 92% for margin negative, ResNet50 classified worthily, which is 76%. However, the margin negative data is 100% classified by both MobileNetV2 and Efficient Net B0 models. As shown in [Table 4](#), the best overall training and validation accuracy achieved by Efficient Net B0 was 95.3% and 94.7% respectively, which is on average greater than the other models used in this work. Moreover, as depicted in [Table 5](#) the overall testing performance of the system achieved by Efficient Net B0 were 95% (at 22 epoch) accuracy, 95% precision, 96% recall, 95% F1 score, 96% specificity, and 100% AUC. This result shows the EfficientNetB0 model outperformed the other models in classifying the SCC.

In this work, a histopathological dataset of SCC and implement a state-of-the-art EfficientNetB0 CNN architecture for margin classification with the best results. To the best of the authors' knowledge, this is the first work to investigate SCC margin classification of skin cancer disease in digitized whole-slide histological images for seven different skin parts and on the three histologic grades of SCC and with such much-improved accuracy. This is the first attempt to design and develop a deep learning computer-aided diagnosis of SCC margin classification system using whole slide images using

locally acquired data sets. We can conclude that the developed system can classify the whole slide of SCC histopathology images with good classification accuracy. Moreover, the developed model has overcome the gap in margin classification of histopathology images in margin-free results during skin cancer surgical treatment of SCC.

In the following [Table 6](#), our proposed system was compared with some previous studies. Almost all studies were focused on only one skin organ location for margin classification, ie, oral. However, for the proposed method, seven different skin organ locations were collected and classified with good accuracy results.

Nevertheless, this study focuses only on the SCC type of skin cancer margin classification and was limited due to financial and time constraints to acquire more datasets to study for other types of cancer cells. Moreover, the current module not able to grade the SCC levels other than classification of the tumor.

Algorithm Demonstration

The developed graphical user interface (GUI) using EfficientNetB0 (with the highest testing accuracy model ~95.2%) was tested with respect to response time and ease of use. It is found to be easy to use and convenient for users. Once initialized, the result can be achieved within less than 10 seconds.

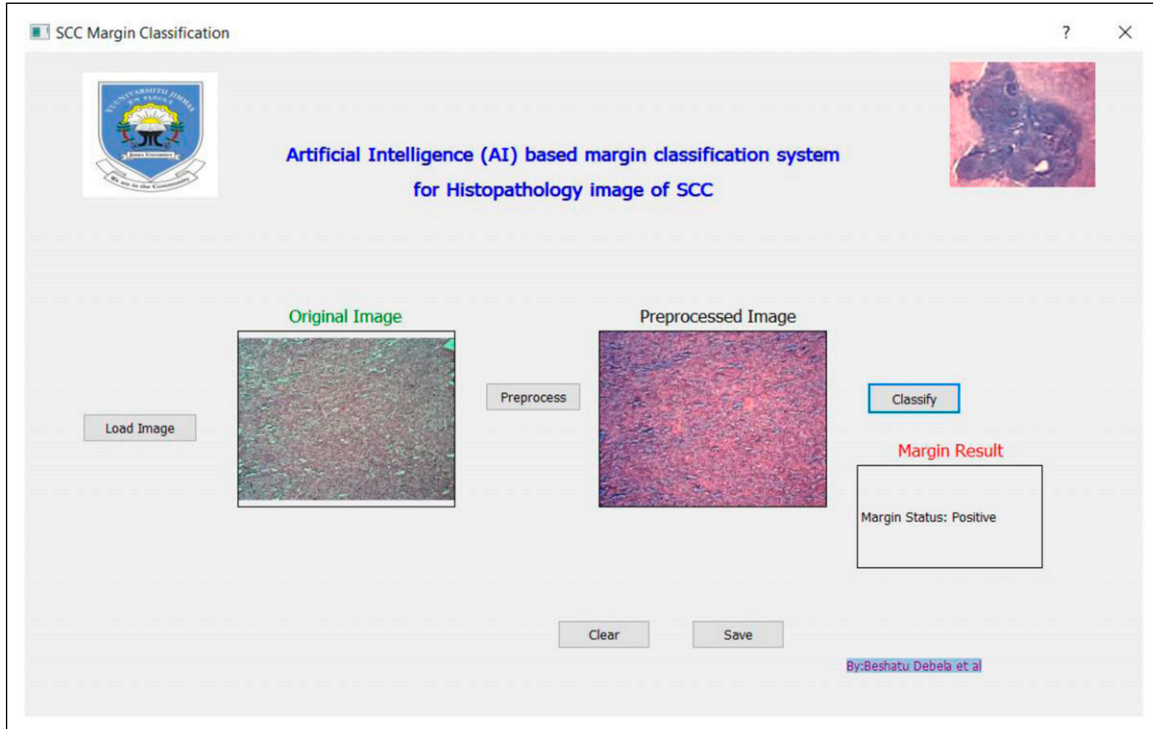


Figure 11. The developed graphical user interface.

As shown in Figure 11, the GUI has a button to load an image and preprocess it and display/classify the diagnosing result. Moreover, the result obtained can be saved using the “save” button, and possible to continue analyzing more images while the “clear” button is used.

Conclusions

The existing manual histopathology margin assessment for the SCC method requires experienced experts, and it is time-consuming, tedious, and depends on the knowledge and experience of the pathologist, which may sometimes require two or more experts to provide a reliable pathology report, which directly affects the treatment plan and cure rate. In this research, we used whole slide images of clinical data collected from Jimma University Medical Center, Pathology Department and trained, validate, and test different selected models by fine-tuning the hyperparameter of four different models, and got significant accuracy. The novel module of our dataset and the promising results of this work demonstrates the potential of such methods that could help to create a tool to increase the efficiency and accuracy of pathologists performing margin assessment on histological slides for the guidance of skin cancer resection operations, especially in low resource settings. The developed system provides the margin classification result within a minute, which shows much improvement from 20 to 30 minutes manual diagnosing methods. For the future, concatenating models of ResNet 50 which had more advantage on

margin positive, which benefit the patients with reduction of recurrence rate of cancer cells, and Efficient Net B0 which had more advantage on margin negative guaranty organ preservation and increases the module performance.

Appendix

Abbreviations

AUC	Area Under the Curve
BCC	Basal Cell Carcinoma
CCPDMA	Complete Circumferential Peripheral and Peep Margin Assessment
FN	False Negative
FP	False Positive
H & E	Hematoxylin and Eosin
HPV	Human Papilloma Virus
HFUS	High-Frequency Ultrasonography
IPC	Intraoperative Pathologist Consultant
JUMC	Jimma University Medical Center
OCT	Optical coherence tomography
MMS	Mohs Micrographic Surgical
NMSC	NonMelanoma Skin Cancer
RCM	Reflectance Confocal Microscopy
ROC	Receiver Operator Characteristic
SCC	Squamous Cell Carcinoma
TN	True Negative
TP	True Positive
WSI	Whole Slide Image

Acknowledgments

We would like to thank the Jimma Institute of Technology, School of Biomedical Engineering for the research funding support and all collaborators from Jimma University Medical Center.

Author Contributions

BDW, KD: Conceptualization, Methodology, Software, Formal analysis, Validation, Writing - original draft, and Data curation. KD, TK: Supervision, Methodology, Writing - original draft, Visualization, Software, Writing - review & editing. BDW, RE, TA, and SK: Data acquisition, data labeling, and analysis. All authors read and approved the final manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Jimma University (JU/2011), Jimma Institute of Technology, School of Biomedical Engineering.

Ethical Approval

This research did not involve direct humans, animals, or other subjects. According to Jimma University's institutional review board (IRB), we had checked that no formal ethical approval was required in this particular case.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding authors on reasonable request.

Guarantor

Jimma University

ORCID iDs

Beshatu Debela Wako  <https://orcid.org/0000-0003-2449-8491>

Kokeb Dese  <https://orcid.org/0000-0003-3591-9570>

Roba Elala Ulfata  <https://orcid.org/0000-0002-9559-896X>

Supplemental Material

Supplemental material for this article is available online.

References

1. Kassem MA, Hosny KM, Damaševičius R, Eltoukhy MM. Machine learning and deep learning methods for skin lesion classification and diagnosis: A systematic review. *Diagnostics*. 2021;11(8):1390. doi:10.3390/diagnostics11081390
2. Hasan MK, Elahi MTE, Alam MA, Jawad MT. DermoExpert: Skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning, and augmentation. medRxiv; 2021. doi:10.1101/2021.02.02.21251038
3. Hasan M, Das Barman S, Islam S, Reza AW. Skin cancer detection using convolutional neural network. Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence, Bali Indonesia, April 19-22, 2019, pp. 254–258. doi:10.1145/3330482.3330525
4. CDC and United States Cancer Statistics (USCS). *United States Cancer Statistics. Highlights from 2019 Incidence*; 2019.
5. Liu Y, Walker E, Iyer SR, et al. Molecular imaging and validation of margins in surgically excised nonmelanoma skin cancer specimens. *J Med Imaging*. 2019;6(1):16001. doi:10.1117/1.JMI.6.1.016001
6. Dildar M, Akram S, Irfan M, et al. Skin cancer detection: A review using deep learning techniques. *Int J Environ Res Public Health*. 2021;18(10):5479. doi:10.3390/ijerph18105479
7. Ma L, Shahedi M, Shi T, et al. Pixel-level tumor margin assessment of surgical specimen with hyperspectral imaging and deep learning classification. Proc. SPIE 11598, Medical Imaging, 15 February 2021, 1159811. doi:10.1117/12.2581046
8. Combalia A, Carrera C. Squamous cell carcinoma: An update on diagnosis and treatment. *Dermatol Pract Concept*. 2020;10(3):e2020066. doi:10.5826/dpc.1003a66
9. Sun C-K, Kao C-T, Wei M-L, Liao Y-H. Slide-free histopathological imaging of hematoxylin-eosin-stained whole mount tissues using Cr:forsterite laser-based nonlinear microscopy. Proc. SPIE, 30 April 2019, 11026. doi:10.1117/12.2520417
10. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: A review. *IEEE Rev Biomed Eng*. 2009;2:147-171. doi:10.1109/RBME.2009.2034865
11. Marsden M, Weyers BW, Bec J, et al. Transactions on biomedical engineering intraoperative margin assessment in oral and oropharyngeal cancer. *IEEE Trans Biomed Eng*. 2020;68(3):857-868. doi:10.1109/TBME.2020.3010480
12. Triki AR, Blaschko MB, Jung YM, et al. Intraoperative margin assessment of human breast tissue in optical coherence tomography images using deep neural networks. *Comput Med Imaging Graph*. 2017;69:21-32.
13. Robbins KT, Triantafyllou A, Suarez C. Clinics in oncology surgical margins in head and neck cancer: Intra- and postoperative considerations. *Clin Oncol*, 3; 2018:1-8.
14. Zhang L, Wu Y, Zheng B, et al. Rapid histology of laryngeal squamous cell carcinoma with deep-learning based stimulated Raman scattering microscopy. *Theranostics*. 2019;9(9):2541-2554. doi:10.7150/thno.32655
15. Halicek M, Dormer JD, Little JV, et al. Hyperspectral imaging of head and neck squamous cell carcinoma for cancer margin detection in surgical specimens from 102 patients using deep learning. *Cancers*. 2019;11(9):1367. doi:10.3390/cancers11091367
16. Halicek M, Little JV, Wang X, et al. Tumor margin classification of head and neck cancer using hyperspectral imaging and convolutional neural networks. *Proc SPIE Int Soc Opt Eng*. 2018;10576:1057605. doi:10.1117/12.2293167

17. Kassem MA, Hosny KM, Fouad MM. Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning. *IEEE Access*. 2020;8:114822-114832. doi:[10.1109/ACCESS.2020.3003890](https://doi.org/10.1109/ACCESS.2020.3003890)
18. Hosny KM, Kassem MA, Fouad MM. Classification of skin lesions into seven classes using transfer learning with AlexNet. *J Digit Imag*. 2020;33(5):1325-1334. doi:[10.1007/s10278-020-00371-9](https://doi.org/10.1007/s10278-020-00371-9)
19. Fei B, Lu G, Wang X, et al. Label-free reflectance hyperspectral imaging for tumor margin assessment: A pilot study on surgical specimens of cancer patients. *J Biomed Opt*. 2017;22(8):1. doi:[10.1117/1.jbo.22.8.086009](https://doi.org/10.1117/1.jbo.22.8.086009)
20. You S, Sun Y, Yang L, et al. Real-time intraoperative diagnosis by deep neural network driven multiphoton virtual histology. *NPJ Precis Oncol*. 2019;3(1):33. doi:[10.1038/s41698-019-0104-3](https://doi.org/10.1038/s41698-019-0104-3)
21. Paolino G, Donati M, Didona D, Mercuri SR, Cantisani C. Histology of non-melanoma skin cancers: An update. *Biomedicines*. 2017;5(4):71. doi:[10.3390/biomedicines5040071](https://doi.org/10.3390/biomedicines5040071)
22. Halicek M, Shahedi M, Little JV, et al. Head and neck cancer detection in digitized whole-slide histology using convolutional neural networks. *Sci Rep*. 2019;9(1):14043. doi:[10.1038/s41598-019-50313-x](https://doi.org/10.1038/s41598-019-50313-x)
23. Kiss K, Sindelarova A, Krbal L, et al. Imaging margins of skin tumors using laser-induced breakdown spectroscopy and machine learning. *J Anal At Spectrom*. 2021;36(5):909-916. doi:[10.1039/D0JA00469C](https://doi.org/10.1039/D0JA00469C)
24. Thomas Robbins K, Triantafyllou A, Suarez C, et al. Surgical margins in head and neck cancer: Intra- and postoperative considerations. *Auris Nasus Larynx*. 2019;46(1):10-17. doi:[10.1016/j.anl.2018.08.011](https://doi.org/10.1016/j.anl.2018.08.011)
25. Van Eycke Y-R, Allard J, Salmon I, Debeir O, Decaestecker C. Image processing in digital pathology: An opportunity to solve inter-batch variability of immunohistochemical staining. *Sci Rep*. 2017;7(1):42964. doi:[10.1038/srep42964](https://doi.org/10.1038/srep42964)
26. Anitha S, Radha V. Comparison of image preprocessing techniques for textile texture images. *Int J Eng Sci Technol*. 2010;2:12.
27. Dese K, Raj H, Ayana G, et al. Accurate machine-learning-based classification of leukemia from blood smear images. *Clin Lymphoma Myeloma Leuk*. 2021;21(11):e903-e914. doi:[10.1016/j.clml.2021.06.025](https://doi.org/10.1016/j.clml.2021.06.025)
28. Rani RU and Amsini P. Image processing techniques used in digital pathology imaging: An overview. *Int J Eng Res Comput Sci Eng*. 2018;5(1):1-4.
29. Kanmani P, Rajivkannan A, Deepak kumar P, et al. Performance analysis of noise filters using histopathological tissue images. *Int Res J Pharm*. 2017;8(1):50-54. doi:[10.7897/2230-8407.080110](https://doi.org/10.7897/2230-8407.080110)
30. Anghel A, Stanisavljevic M, Andani S, et al. A high-performance system for robust stain normalization of whole-slide images in histopathology. *Front Med*. 2019;6:193. doi:[10.3389/fmed.2019.00193](https://doi.org/10.3389/fmed.2019.00193)
31. Jose L, Liu S, Russo C, Nadort A, Di Ieva A. Generative adversarial networks in digital pathology and histopathological image processing: A review. *J Pathol Inf*. 2021;12:43. doi:[10.4103/jpi.jpi_103_20](https://doi.org/10.4103/jpi.jpi_103_20)
32. Lakshmanan B, Anand S, Jenitha T. Stain removal through color normalization of haematoxylin and eosin images: A review. *J Phys Conf Ser*. 2019;1362:012108. doi:[10.1088/1742-6596/1362/1/012108](https://doi.org/10.1088/1742-6596/1362/1/012108)
33. Smith B, Hermsen M, Lesser E, Ravichandar D. Developing image analysis pipelines of whole-slide images: Pre- and post-processing. *J Clin Transl Sci*. 2020;5(1):1-33. doi:[10.1017/cts.2020.531](https://doi.org/10.1017/cts.2020.531)
34. Macenko M, Niethammer M, Marron JS et al. A method for normalizing histology slides for quantitative analysis. Proceedings of the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Boston, MA, June 28-July 1, 2009, pp. 1107–1110. doi: [10.1109/ISBI.2009.5193250](https://doi.org/10.1109/ISBI.2009.5193250)
35. Öztürk Ş, Akdemir B. Effects of histopathological image pre-processing on convolutional neural networks. *Procedia Comput Sci*. 2018;132:396-403. doi:[10.1016/j.procs.2018.05.166](https://doi.org/10.1016/j.procs.2018.05.166)
36. Mahbod A, Schaefer G, Wang C, Ecker R, Ellinge I. Skin lesion classification using hybrid deep neural networks. ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, May 12-17, 2019, pp. 1229–1233. doi:[10.1109/ICASSP.2019.8683352](https://doi.org/10.1109/ICASSP.2019.8683352)
37. Amin I, Zamir H, Khan FF. Histopathological image analysis for oral squamous cell carcinoma classification using concatenated deep learning models. medRxiv, 2021. doi:[10.1101/2021.05.06.21256741](https://doi.org/10.1101/2021.05.06.21256741)
38. Raheem MA. A deep learning approach for the automatic analysis and prediction of breast cancer for histopathological images using A Webapp. *Int J Eng Res Technol*. 2021;10(6):996-1001.
39. Praveen Gujjar J, Prasanna Kumar HR, Chiplunkar NN. Image classification and prediction using transfer learning in colab notebook. *Glob Transitions Proc*. 2021;2(2):382-385. doi:[10.1016/j.gltip.2021.08.068](https://doi.org/10.1016/j.gltip.2021.08.068)
40. Rafi TH. An efficient classification of benign and malignant tumors implementing various deep convolutional neural networks. *Int J Comput Sci Eng Appl* 2020;9(2):152-158.
41. Hosny KM, Kassem MA, Fouad MM. Skin melanoma classification using ROI and data augmentation with deep convolutional neural networks. *Multimed Tool Appl*. 2020;79(33):24029-24055. doi:[10.1007/s11042-020-09067-2](https://doi.org/10.1007/s11042-020-09067-2)
42. Rannen Triki A, Blaschko MB, Jung YM, et al. Intraoperative margin assessment of human breast tissue in optical coherence tomography images using deep neural networks. *Comput Med Imaging Graph*. 2018;69:21-32. doi:[10.1016/j.compmedimag.2018.06.002](https://doi.org/10.1016/j.compmedimag.2018.06.002)
43. Kho E, de Boer LL, Van de Vijver KK, et al. Hyperspectral imaging for resection margin assessment during cancer surgery. *Clin Cancer Res*. 2019;25(12):3572-3580. doi:[10.1158/1078-0432.CCR-18-2089](https://doi.org/10.1158/1078-0432.CCR-18-2089)