Melanoma Cancer Cell Classification Using Deep Learning

Hasan Naqvi[1], Ashir Wahid[2], & Kazim Raza Rizvi[3]
[1] [2] Student, Electrical Engineering, School Of Science & Engineering, Habib University, Pakistan, sn00564@st.habib.edu.pk
[3] Student, Electrical Engineering, School Of Science & Engineering, Habib University, Pakistan, mw00899@st.habib.edu.pk
Student, Electrical Engineering, School Of Science & Engineering, Habib University, Pakistan, sr00728@st.habib.edu.pk

***Abstract*— Melanoma is known to be a dangerous and harmful form of cancer. The early detection of the danger behind melanoma cancer cell makes the recovery and survival rate chances to increase dramatically. However, being able to accurately detect the possibilities of a melanoma cancer cell being malignant as compared to benign is a challenge that we haven't solved yet. The reason being that both benign and malignant cells have high similarities, their features are extremely similar to the human eye, and the presence of artifacts such as skin colors, hair, and others make it difficult to identify the risk involved. For this very reason, in order to develop a reliable system to detect and classify melanoma cells, large amounts of research is being conduct in the areas of computer vision in terms of skin lesion analysis. The International Skin Imaging Collaboration (ISIC) has made public a dataset of melanoma images and announced challenges in segmentation, feature extraction, and classification. In this paper, we propose a classification method based on Convolutional Neural Networks. Prior to training our model, we propose a preprocessing unit to enhance the quality of the given dataset. We apply techniques such as region of interest extraction, artifact removal, and data augmentation. Our results have been promising in terms of our model classifying the test images. We have been able to obtain an accuracy of 0.8214 for our proposed system.**

***Index Terms* –** melanoma cancer cells, skin lesion analysis, melanoma recognition, convolutional neural networks, melanoma classification, melanoma segmentation.

I. Introduction

Melanoma , contributing to around 75% of deaths in the world, is known to be one of the most dangerous forms of cancer that humanity is up against [1]. Identifying the danger at the earlier stages help prevents fatal conclusions. However, in order to successfully detect and identity a risk associated with melanoma, we require well-trained specialists who are able to tell whether a cell is cancerous or not. This is highly impractical as there are a large number of variations and uncertainties present within the human eye observation. In order to tackle the issue at hand regarding the efficient detection and classification of melanoma cancer cells, what we need is an automated and reliable system in place which can do the task for us.

In order to improve the identification of melanoma cancer cells, technology advancements have led us to a technique known as dermoscopy. Dermoscopy is a skin imaging technique that takes an illuminated and enlarged image of the cancel cell region to result in clearer identification of the area [2]. However, regardless of such a development, the recognition of melanoma cancer cells from such images still seems to be quite challenging. This is due to the fact that both benign (not cancerous) and malignant (cancerous) cells have extremely high similarities making it difficult to identify the difference using the human eye. Moreover, artifacts on these images, such as hair, skin color, and others make it difficult to ensure accurate identification of the state of the cancer cell.

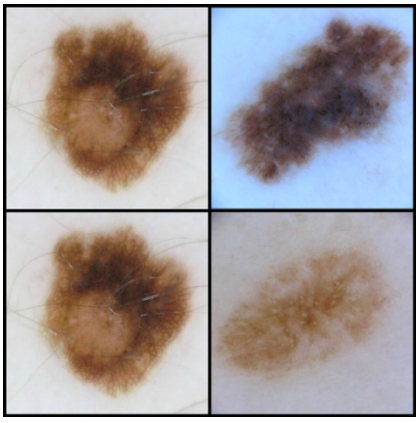

Fig. 1. Examples of images of melanoma cancer cells under different varying conditions such as artifacts, skin color, lighting, etc.

While humans may not be able to successfully classify melanoma images, we can use high computational powers to train a machine learning model on a large dataset of melanoma images. Recent developments in the areas of computer vision have led us to progress in this very field. Li et al. proposed a deep learning framework along with a segmentation and feature extraction platform to classify melanoma cancer cells [2]. They were able to obtain high accuracies while competing in the ISIC 2017 competition. Similarly, Yu et al. proposed a very deep fully convolutional residual network for both segmentation and classification of images through a limited dataset [3]. They were able to obtain highly competitive results in the ISIC 2016 competition with an accuracy of around 85.55%. Codella et al. also proposed a system using recent developments along with a deep machine learning network

to achieve the task [4]. Gutman et al. worked on an overview of general techniques used to segment and classify melanoma cancer cells [5].

While a lot of work has been done and continues to be done in the area of melanoma cancer cell segmentation and detection, there is still enough room for continuous improvement. The International Skin Imaging Collaboration (ISIC) has been organizing challenges in the areas of skin imaging techniques. They have made public a dataset for melanoma images which different groups from all over the world work on. The challenge has been divided into three different tasks, where task I is about lesion segmentation, task II is on feature extraction from dermoscopy images, and finally task III is on lesion classification.

The idea is to be able to first segment the region of interest, being the cancel cell itself, and then move on to different techniques to extract in-depth features which can be used to train a machine learning model. In this paper, we propose a research methodology that can be summarized as follows:

1) While existing methods may proposed and set the importance of preprocessing at different levels, we ensure that most of our energies are spent on preprocessing the given images and coming up with a large and strong dataset. Therefore, we focus on the preprocessing of images by first applying segmentation, following by region of interest extraction, artifact removal, and data augmentation.

2) We then propose a Convolutional Neural Network that is 14 layers deep in order to train our preprocessed dataset on. We vary our given parameters in the neural network in order to achieve the highest possible accuracy on a Central Processing Unit (CPU).

3) Finally, we implement a detailed analysis on our results to understand how our model works and tests the given testing data. Moreover, we look at methods and areas which can be improved in order to increase the given accuracy.

II. Methodology

In this section of our paper, we aim to introduce our research methodology and how we went ahead with using our preprocessing techniques along with a Convolutional Neural Network (CNN) to achieve the task at hand. Our preprocessing methodology can be broken down into four independent steps, starting from segmentation to region of interest extraction, artifact removal, and finally data augmentation. Our techniques were applied on the dataset provided by ISIC and at the last stage, our data was ready to be trained on the neural network we built.

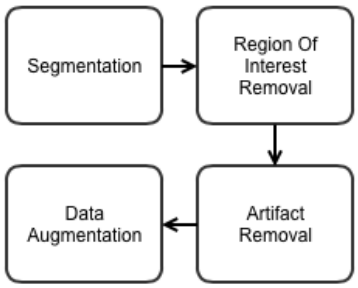

Fig. 2. Preprocessing techniques using to improve the given dataset of images

A. Preprocessing

1) Segmentation

The ISIC dataset came along with segmented images for the training data. These images were segmented on a rough region with an estimated shape in order to take into account the complete cancer cell. We went on to develop our own techniques for segmentation using Otsu Thresholding in order to extract the region of interest. An example of our segmentation technique is shown in Fig. 3.

Now that we have both the options for segmented images. We used these to be able to extract the required region of interest in order to avoid training our model with unnecessary background images. This provides us with a mask image of where our cancer cell is located. These features could disturb our learning abilities and cause us problems at the testing stages.

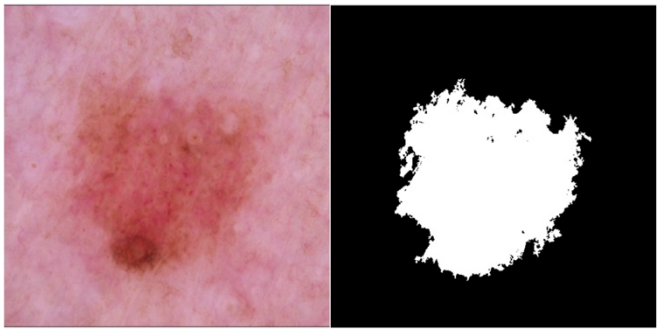

Fig. 3. Segmentation technique to create a mask which identifies the location of the cancerous cell

2) Region Of Interest Extraction

Once we had our segmented images along with our original images containing the melanoma cancer cells, we went on to extract the region of interest, being the cancer cell only, through our mask image we created with our segmentation technique. This was done using simple binary operations where we wrote a program to break the original image into its RGB constituents and then apply logical operations on each channel to extract the cancerous cell region while putting the rest of the background to null. Finally, we put together the channels to produce our output image which contained only the region of interest. This can be practically shown in Fig. 4.

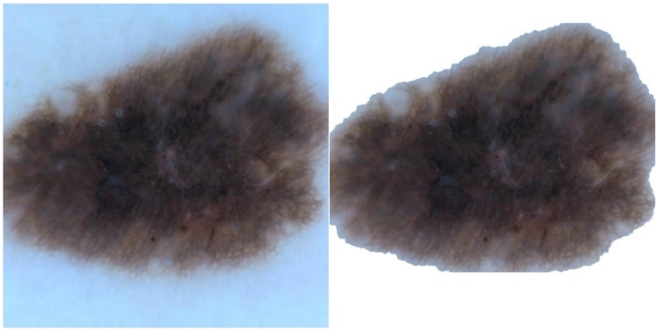

Fig. 4. Region of interest extraction using the segmented image to obtain the cancer cell region without its background

3) Artifact Removal

Now that we had our region of interest in place, we needed to take into account the different artifacts that may be in the view of the cancer cell. A prominent example of this is the presence of hair on top of the cancer cell which becomes an artifact in the dermoscopy images.

In order to remove such artifacts, we applied basic image processing techniques on our given images. We used a disk structuring element and applied the bottom hat transform along with a dilation process to remove objects of a certain size. This allowed us to remove all hair artifacts from our images in the dataset such that they will not interfering with our feature extraction process during the training of the dataset on our machine learning model.

A practical example of hair removal can be seen in Fig. 5.

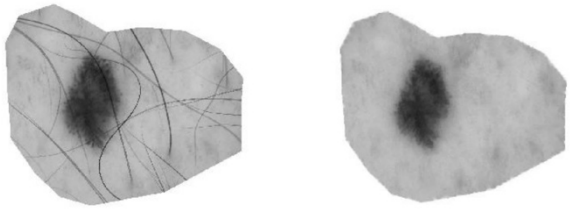

Fig. 5. Artifact removal using image processing techniques

4) Data Augmentation

With the given dataset, we had an imbalance with the benign and malignant images which would have caused us trouble during our training process. Therefore it was necessary to balance out the categories prior to training our Convolutional Neural Network. For this stage in our preprocessing module, we applied data augmentation to balance our dataset.

Initially, we had a much lower count on the malignant cancer cell images as compared to the benign ones. Therefore, we applied augmentation techniques to the malignant dataset in order to ensure our data is balanced out prior to training. This was done through operations such as rotation, mirroring, etc.

After our data augmentation process, we had a total of 1154 images which we used to train our machine learning model. This can be seen in Table 1.

TABLE I
INFORMATION ON DATA AUGMENTATION

	Benign	Malignant
Original	564	118
After Data Augmentation	564	590

B. Training The Classifier

a) Network Architecture

In order to train our model, we set up a Convolutional Neural Network based on around 14 layers in total. The idea was to be able to develop a network deep enough to provide us with a decent accuracy and one that could also be trained in an acceptable time period on a CPU. The model was trained on the dataset generated following the preprocessing tasks that were explained in the last section.

Initially, we trained our model on the dataset without preprocessing and noticed that our results were unacceptable. Our accuracy was in the range of 60-70% and continued to remain there when the layers were varied from 7 layers to 14 layers. Similarly, even when we had applied preprocessing our model did not produce decent results if we trained it on 7 layers. It was only when our model depth was increased to 14 layer that it began to generate sufficient result. Therefore, it was necessary to be able to find the common ground between all parameters in order to generate the desired accuracy.

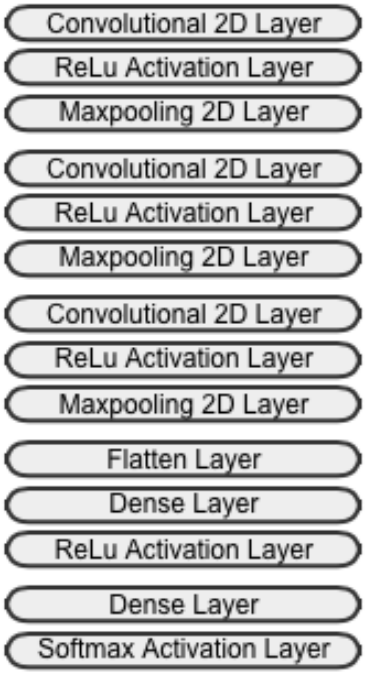

Fig. 6. Convolutional Neural Network Architecture (14 layer deep).

b) Implementation

For our network implementation we used Keras, which is Python's deep learning library. Keras allows for the implementation of Convolutional Neural Networks using Python code. Through our implemented model on Keras, we use 75% of the dataset for training data and the remaining 25% for validation data. Our preprocessing module was individually implemented using MATLAB through which we created the dataset by going through the four steps independently. We then trained our Convolutional Neural Network on a CPU with the dataset that was created using the preprocessing steps mentioned above. For our epochs, we set the value to around 50 epochs given the size of our data.

III. Performance Analysis

We tested our trained model on a total of 292 images of melanoma cancer cells. The images went through the preprocessing unit in order to be free from any background features and artifacts. Once we had our results, we put together a confusion matrix to analyze our data. This can be visualized in Table 2.

TABLE 2
CONFUSION MATRIX

		Actual	
		Benign	**Malignant**
Predicted	**Benign**	133	32
	Malignant	13	114

Our trained model provided us with a validation accuracy of 82.08% After testing our model on the test images, we were able to classify most of the given data into benign and malignant successfully, as it can be seen from Table 2.

The International Skin Imaging Collaboration (ISIC) defines a number of evaluation metrics on the basis of which we are able to identify and evaluate our performance. These metrics include the accuracy (AC), dice coefficient (DI), sensitivity (SE), jaccard index (JA), and specificity (SP). Let TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative respectively. Then our metrics can be defined as:

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

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

$$DI = \frac{2 \cdot TP}{2 \cdot TP+FN+FP} \quad (3)$$

$$SE = \frac{TP}{TP+FN} \quad (4)$$

$$SP = \frac{TN}{TN+FP} \quad (5)$$

Fig. 7. Evaluation Metrics to analyze performance

Based on our tested data, we put together our results an observed our metrics as accuracy 0.85, jaccard index 0.75, dice coefficient 0.86, sensitivity 0.91 and specificity 0.78. These are extremely competitive metrics based on the ISIC 2017 competition results. Some of our test cases can be shown in Fig. 8.we

IV. Comparison with other groups

During the ISIC 2017 competition, around 15 teams participated and put together their classification results based on the dataset provided by the competition as well as the performance metrics that were defined.

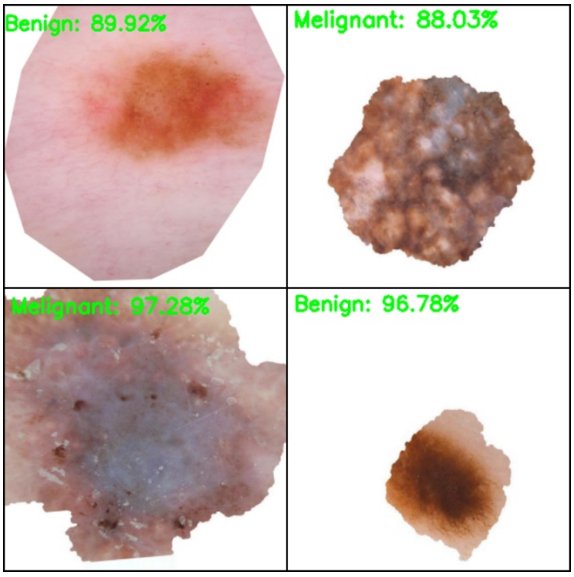

Fig. 8. Test cases on different melanoma cancer cells

Method	AUC	AC	AP	SE	SP
CSUJT	**0.911**	0.816	0.748	**0.856**	0.812
MPG-UCIIIM	0.910	0.849	0.747	0.140	**0.998**
RECOD Titans	0.908	0.883	**0.752**	0.451	0.970
USYD-BMIT	0.896	**0.888**	0.732	0.508	0.970
IHPC-NSC	0.886	0.873	0.665	0.568	0.940
UoG-MLRG	0.886	0.879	0.703	0.453	0.971
icuff	0.851	0.819	0.578	0.524	0.893
icuff	0.850	0.817	0.579	0.524	0.890
USYD-BMIT	0.836	0.850	0.569	0.210	0.989
CVI	0.829	0.863	0.593	0.460	0.950
UoD	0.825	0.849	0.557	0.591	0.907
INESC TEC Porto	0.823	0.657	0.566	0.814	0.615
UFMG	0.823	0.830	0.567	0.488	0.900
LIN (ours)	0.823	0.852	0.476	0.504	0.930
IPA	0.811	0.811	0.542	0.362	0.901

Fig. 9. Results of the ISIC Classification Challenge in 2017 [2]

Fig. 9 is taken from [2] and it shows the results of the classification challenge from ISIC 2017. As we can see, our performance can be seen to outrank at least 8 different groups in terms of the accuracy and various others based on different evaluation metrics.

The reason why our false positive rate (actual: malignant, predicted: benign) is so high is because of the lack of

diverse dataset in the malignant category. We had to augment the data for malignant images which resulted in the same images being reflected or rotated. Moreover, our limited in terms of the CPU when it came to training also prevented us from using a very deep Convolutional Neural Network. However, our results turned out to still be quite competitive.

V. Conclusion

In this paper, we proposed a method to classify melanoma cancer cell images using a Convolutional Neural Network that was 14 layers deep along with a preprocessing module that involved segmentation, region of interest extraction, artifact removal, and finally data augmentation. Our aim was to be able to classify images using a model trained on CPU with a decent enough accuracy. Moreover, due to the presence of a limited dataset, we had to ensure our preprocessing techniques were outstanding in order to generate the desired results.

If we would be able to train our model on GPU, by increasing the number of layers and moving towards a very deep neural network, then our accuracy might further increase given that we can increase our dataset as well.

Our performance was evaluated based on various different metrics and we observed that our results were quite competitive as compared to the ones present in the International Skin Imaging Collaboration 2017. Our evaluation metrics were calculated to be: accuracy 0.85, jaccard index 0.75, dice coefficient 0.86, sensitivity 0.91 and specificity 0.78. For this very reason, our proposed model demonstrates its abilities to address the challenge at hand in a quite competitive manner.

References

[1] A. F. Jerant, J. T. Johnson, C. D. Sheridan, and T. J. Caffrey, "Early detection and treatment of skin cancer," American Family Physician, vol. 62, no. 2, pp. 381-382, 2000.

[2] Y. Li and L. Shen, "Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network," *Sensors*, vol. 18, no. 2, p. 556, Nov. 2018.

[3] Yu, Lequan & Chen, Hao & Dou, Qi & Qin, Jing & Heng, Pheng-Ann. (2016). Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks. IEEE Transactions on Medical Imaging. PP. 10.1109/TMI.2016.2642839.

[4] Codella, Noel & Nguyen, Quoc-Bao & Pankanti, S & Gutman, David & Helba, Brian & Halpern, Allan & R. Smith, John. (2017). Deep Learning Ensembles for Melanoma Recognition in Dermoscopy Images. Ibm Journal of Research and Development. 61. 10.1147/JRD.2017.2708299.

[5] Gutman, David & Codella, Noel & Celebi, M. Emre & Helba, Brian & Marchetti, Michael & Mishra, Nabin Kumar & Halpern, Allan. (2016). Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC). arXiv:1605.01397 [cs.CV].

[6] Codella, Noel & Gutman, David & Celebi, M. Emre & Helba, Brian & Marchetti, Michael & W. Dusza, Stephen & Kalloo, Aadi & Liopyris, Konstantinos & Mishra, Nabin Kumar & Kittler, Harald & Halpern, Allan. (2017). Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC).

[7] Mishra, Nabin Kumar & Celebi, M. Emre. (2016). An Overview of Melanoma Detection in Dermoscopy Images Using Image Processing and Machine Learning. arXiv e-print. arXiv:1601.07843.