Melanoma Skin Cancer Detection Using Deep Learning and Advanced Regularizer

Md. Arman Hossin
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
mdarmanhossin99@gmail.com

Farhan Fuad Rupom
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
rupom.cse16@gmail.com

Hasibur Rashid Mahi
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
mahialhasib@gmail.com

Anik Sarker
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
anik.brac.2016@gmail.com

Farshid Ahsan
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
arnobahsan2016@gmail.com

Sadman Warech
Computer Science and Engineering
BRAC University
Dhaka, Bangladesh
ridom555@gmail.com

***Abstract*—Melanoma cancer Detection System is a predictive model that dynamically anticipates melanoma skin cancer by evaluating dermoscopic images with the help of deep learning. The fundamental goals behind this research are to identify skin cancer at early stages by achieving swift results with greater accuracy. The reason behind the goal signifies the problem of increment in skin cancer patients worldwide, high medical costs and exponential increment of death risk for not starting the diagnosis at early stages which is a result of late detection. Our presented research work proposes a solution to the problem of higher medical costs behind diagnosis, lower accuracy rate in detection and portability problem of the manual detection system. In this system, dermoscopic images are classified to predict skin cancer using a multi-layered CNN approach with multiple regularization techniques named dropout and batch normalization. As a result, our system has provided an accuracy of 93.58% which is higher than most other conventional approaches.**

***Index Terms*—Convolutional Neural Network, skin cancer, regularization**

I. Introduction

In today's age, computer aided diagnosis systems through using machine learning and deep learning has become a requirement for early detection and diagnosis of many fatal diseases. Our system also represents a computer aided diagnosis system, where the system classifies and detects if an image of the damaged skin is melanoma cancer or not. It has been done by the help of deep layered convolutional network algorithms. The deep layers of CNN train the data set and extract the features more easily and accurately. Image processing has been done also to remove noise from image data and to prepare it more convenient and suitable to train in the model.

Unfortunately, for the last 40 years, melanoma occurrences have risen worldwide [1]. 1% of skin tumours tends to be malignant melanoma and causes about 60 percent of skin cancer deaths. Presently, Australia has been crowned with the highest incidence of melanoma cancer as each year 40 cases have been found within ten thousand people [1]. The occurrences of malignant melanoma in Europe and the USA have increased by three times for the past three decades [2]. In terms of cost in skin cancer diagnosis and treatment, it has been skyrocketing. Per patients' average cost on diagnosis had risen from $1000 in 2006 to $1600 in 2011 [3]. Medicare payment databases on skin cancer management provided information from 880,000 medical providers receiving healthcare costs of $77 billion in 2012 [3]. The estimated annual average expenses of skin cancer treatment between 2002 and 2006 was $3.6 billion; but from 2007 to 2011, the total yearly expenses were $8.1 billion which was a rise of 126.2% [3]. By comparison, all other cancers have only risen by 25.1% in medical cost [3]. In terms of casualty rate, about 65155 individuals die from non-melanoma cancer and 60712 casualties occurred globally from melanoma disease according to [4]. The overall survival rate is 92 percent throughout all levels of melanoma. This survival rate of five years tends to be prolonged up to 99% if the skin cancer is detected in early stage [5]. As a result, the sooner the skin cancer gets identified, the better the chances of skipping complex treatments like surgery and avoiding the risks of potential disfigurement or death [6]. Even melanoma cancer is also curable if its detected early [6], as this type of cancer spreads very quickly beyond one's skin than other types of cancer to other organs including brain and bones [7]. If it gets spread throughout other body parts rather than only skin, treatments become more complex and the percentage of curability almost becomes zero [7]. For this reason, melanoma cancer is needed to be detected accurately as early as possible. At present though there are many conventional systems that exist which are able to detect melanoma cancer but lack in higher accuracies. For detecting the cancer, the system has to be accurate and more efficient in classification. For these reasons, a convenient and efficient system is needed to classify the melanoma cancer accurately in order to prevent and cure the cancer as early as possible.

The rest of the paper has been described in the following manner. Objectives behind this research has been presented in section II and the section III describes about the related works. The methodology and implementation of the proposed model has been discussed at the section IV. Additionally, Section V represents the result and analysis. Lastly, section VI concludes our system.

II. Research Objective

The objective behind this work is to develop a system which can automatically detect melanoma more precisely only by examining an image of the region of interest. In order to detect melanoma more accurately than most other conventional systems, our proposed system has been built with convolutional neural network including deep layers for better feature recognition and detection. The presented system can detect melanoma skin cancer with higher accuracy of about 93.58% that serves the objective behind this system. Moreover, the manual diagnosis and detection test for melanoma skin cancer is much costlier for financially back lagged people. Even after being so high priced, people tend to rely on manual diagnosis and detection tests as manual diagnosis usually provides accurate result. So, another objective behind the indagation is to build a methodology which is much less costly than manual diagnosis with higher accuracy than any other manual and conventional diagnosis systems. Additionally, developing a software system which can be implemented through a rapid test kit for skin cancer detection has always one of the dedicated objectives behind this research as it will be more compact and faster in providing results.

III. Related Works

Hemalatha N et al. [8] proposed an advancement in skin cancer categorization through using CNN that managed to classify the cancer with the help of TensorFlow as well as Keras and determines whether its Benign or Malignant. The precision of his proposed Convolution two-dimensional layer technique was around 78 percent.

Chanki Yu et al. [9] implemented the dermoscopy of acral melanoma including healthy nevi pictures in the hands along with feet through convolutional neural networks for its early detection. To ensure system accuracy, 2-fold cross validation was performed where fifty percent of the dataset were selected for training and the other half for testing. Their System precision rate of true positive and true negative had been shown about 83.51% and 80.23 respectively by implementing convolutional neural network on whole dataset of images.

Swati Srivastava et al. in [10] have experimented with an artificial neural network approach implementing BNN (Back-propagation neural network) along with ANN (Auto-associative neural network).

In paper [11], J Abdul Jaleel et al. represented a technique in which their system's cancer identification process includes image filtering for skin hair removal, image segmentation through implementing Maximum Entropy Threshold, extraction of features with the help of Gray Level Co-occurrence Matrix and Artificial Neural Network(ANN) used in classification of test data. Not only ANN but also Back-Propagation Neural (BPN) Network has been implemented in order to categorize more accurately.

Another modified automated method of segmentation has been presented in [12] which is an exceptional approach of implementing Convolutional Neural Network (CNN) algorithm. CNNs were implemented in the system in order to extract features and ANNs were used in order to categorize the extracted features. A noteworthy topic they mentioned about not using any additional classifier like support vector machine or k-nearest neighbors. These algorithms were not needed for the CNN, as they trained the classification model through utilizing three inter connected convolutional layers.

Besides that, Ridell and Spett [13] published research where convolutional networks have been trained utilizing the Google Inception v3 in order to categorize and detect cancer in skin. Researchers analyzed system's efficiency in the categorization between benign melanoma and malignant melanoma were determined by their composition of the training data set dimension along with the size of the data set.

IV. Proposed Model

To begin with, the dataset of images has been implemented into the system. After that, the pre-processing starts, in which the model has performed through resizing the images, creating train, validation and test sets from the dataset, normalization, removing noise of the images and finally implementing data augmentation technique.

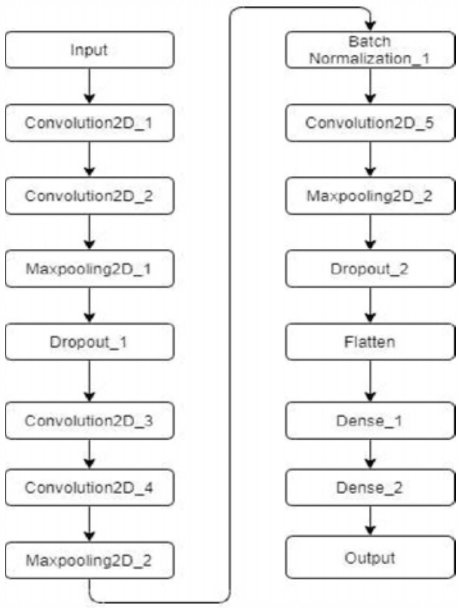

Fig. 1. Model Overview

After the completion of pre-processing, two dimensional convolutional neural has been used with the help of TensorFlow and Keras. Subsequently, another convolution layer has been added in order to operate each of the feature maps located in first convolutional layer instead of operating images directly. Again, in order to lessen the parameter numbers along with fitting the proposed model, max pooling layer has been implemented. Additionally, a dropout layer has also been used in order to prevent over fitting. In the second phase, after two convolution layers and max pooling, batch normalization has been introduced which re-scales output of a particular layer so that it always has 0 mean and standard deviation of 1 in every batch of training. After that, the 5th convolutional layer has been applied which includes max pooling dropout and output flattening of the 5th layer respectively. Subsequently, the acquired flattened output gets passed through two fully connected dense layers to achieve the final output.

A. Dataset

CNN model has been trained and applied on the skin lesions photos in order to assess the proposed approach. For this reason, an image dataset of skin lesions has been acquired through ISIC (International Skin Image Collaboration) archive which includes 1497 images with malignant marked moles and 1800 images with benign moles [14]. Images, that have been acquired from that dataset have been resized into lower resolution of 224x224. After that, each component of the dataset has been trained in order to implement the model accurately and the overall accuracy of the validations has been measured accordingly. Figure [14] displays some of the images from the dataset as an example.

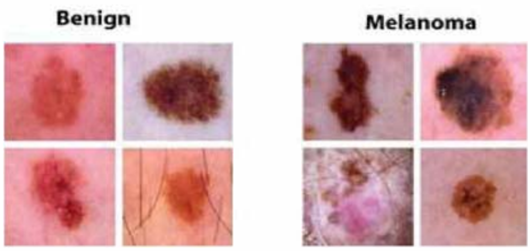

Fig. 2. Example of dataset

B. Pre-processing

Pre-processing is a technique that has been used to improve image data that eliminates any unnecessary defects or enhances other image characteristics that are necessary for further processing.

1) *Initializing*: Images have been loaded and converted by using their RGB values into NumPy arrays.
2) *Creating train, validation, and test sets*: Set of images has been generated with their respective labels where 70% of it for the training and 30% for the validation purpose.
3) *Normalization*: The RGB values of the images have been divided by 255 to have values from 0 to 1.
4) *One-hot encoding*: Categorical variables have been converted into 0 1 vectors where 0 has been labelled benign and 1 has been malignant.
5) *Removing noise*: The noise from the image has been reduced and denoised from the restoration module using total variation filter through applying weights of 0.1.
6) *Data augmentation*: A fair improvement of data in the system's training set has produced additional added data. The goal is to add additional information to the system's workspace, which is close to the information that exists already but it is updated to a certain extent and not exactly the same. Some of the additional parameters are spontaneous rotation around (-45, 45) degrees, shifts, shears, and flips.

C. CNN

CNN is one of the variants of the neural networks used extensively in the computer vision field. This is an algorithm that initializes an image as input, assigns importance to many elements of the image and can discriminate between them. CNNs consist a number of convolutional layers with linear operators as well as nonlinear operators. A three-dimensional sequence of dimensions, where spatial dimensions are height as well as width and 3rd dimension is numerical count of channels or measurements, represents feature map along with individual layer of CNN. In the process of implementing CNN, input composed of an inflexible size of 224 x 224 x 3 has been transmitted across a convolutional layers' stack having sixty-four filters and kernel structure of 3x3, in which a rectified linear unit (ReLU) activation function has been followed by them individually. After that, a method called max-pooling has been utilized on a window of 2×2-pixel including a stride of 2. After that, a sequence of two convolution layers followed by max-pooling and dropout of 0.2 has been applied. In the similar fashion, third and fourth convolution layers have also been implemented followed by max pooling as well as batch normalization. Thenceforth, the acquired flattened output gets passed through two fully connected dense layers with the help of ReLU along with softmax activation in order to achieve the final output.

D. Pooling

A large number of parameters have always been one of the problems of fitting neural networks. The performance of convolution layers can be summarized in a succinct manner. Pooling has been used in order to do that task. Moreover, a collection of pixels has been synthesized on the basis of their total pooling values. Also, after convolution, the pooling layer has been added. The pooling window size refers to any consecutive 2 to 2 blocks. This ensures that pooling takes a maximum number from input designated for each individual layer in output over a range composed with two to two pixels. The pooling layer tends to be $x_h \times x_w \times x_c$, where the result is presented through equation 1.

$$[\{(x_h - y)/z + 1\} \times \{(x_w - y)/z + 1\} \times x_c] \quad (1)$$

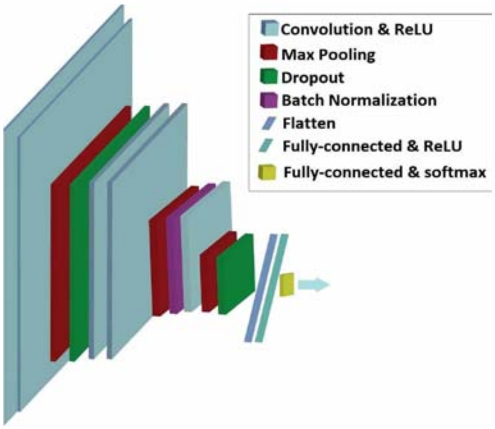

Fig. 3. CNN architecture

Using the pooling operation, the number of parameters has been dramatically reduced in the model.

E. Regularization

Regularization is a strategy by which the complexity of the classifier gets managed. To implement this strategy, using dropout in neural network is one of the optimal options which uses the regularizers L1 and L2, data augmentation, batch normalization and early stopping. Overfitting is a major problem in a deep neural network. Large networks are often hard to use, making overfitting impossible to cope with when integrating the results of several different large neural networks at the time of evaluation [15].

1) Dropout: Dropout technique has been also used in our proposed model. In this technique, a random subset of units in a layer has been chosen to be ignored in each learning step. This unit group also gets ignored on both forward pass through the network as well as back propagation state later on [15]. Dropout of 0.2 has been applied throughout the system.

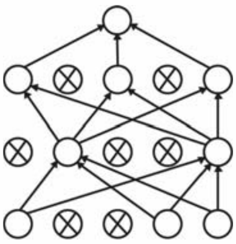

Fig. 4. Applying dropout

The method allows several different networks to be trained on various data pieces. Each time a qualified network gets randomly chosen from the whole network. That way, if some of the data is overly sensitive to any noise in the data, other sections compensate for this, as of that noise they have not seen the samples. It also prevents the various parts of the network from becoming overly correlated in their operations.

2) Batch Normalization: In this model, batch normalization has been used as well after the convolution layer. This operation takes the output of a particular layer as well as re-scales in order to ensure 0 as mean and 1 as standard deviation in every batch of training consistently. This solves the problem where different batches of input produce widely different distributions of output in any given layer in the network. The adjustment to the weight through back-propagation depends on the activation of the units in every step of learning. The Batch Normalization Transform algorithm has been described below with these equations [16] :

Input: Values of x over mini-batch $B = \{x_{1\ldots m}\}$;
parameters to be learned: γ, β
Output: $\{y_i = BN_{\gamma,\beta}(x_i)\}$;
Mini-batch mean:

$$\mu_B \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i \tag{2}$$

Mini-batch variance:

$$\sigma_B^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m} (x_i - \mu_B)^2 \tag{3}$$

Normalize:

$$\hat{x}_i \leftarrow \frac{(x_i - \mu_B)}{\sqrt{\sigma_B^2 + \epsilon}} \tag{4}$$

Scale and shift:

$$y_i \leftarrow \gamma\hat{x}_i + \beta \equiv BN_{\gamma,\beta}(x_i) \tag{5}$$

Here, B and m has been taken as a mini-batch and the size of that batch respectively. It describes that there are m values for activation of this mini batch B. Though normalization is applied to each activation individually, here only the activation of x_i has been described where x_i(i=1 to m) is a dimensional layer input. So, the algorithm that has been shown through equations 1 to 4 have been expressed based on applying to activation x only over a mini-batch B. Additionally, β and γ are expressed as a parameter pair, that is able to do scaling and shifting of normalized value. Here, in the presented algorithm, σ has been identified as a constant appended with variance of mini-batch in order to ensure numerical stability. Here, μ_B, σ_B^2, $\hat{x}_i$, y_i determines the mean, variance, normalized as well as scaled and shifted values respectively. Through the notation $y_= BN_{\gamma,\beta}(x)$, γ and β have been indicated as learning parameters. Though BN transform has not been able to process the activation in each training example independently. Rather, it has depended on the training as well as other samples in mini-batch. The value of shifting and scaling which is denoted by y gets passed to the other layers of network. Batch standardization helps SGD to denormalize it only by adjusting all weights for each activation instead of by modifying all the weights along with sacrificing the reliability of the network.

F. Training Process

Our dataset comprises 3297 images where related labels have been specified and rescaled to 224 × 224 pixels. In addition, the training data set has been artificially increased to expand the variation robustness of the geometric transformation on this proposed CNN approach. Additional increased data have been formed by the rotation of the original training set into(-45, 45) degrees and flipping the images. Thirty percent of the training data have been chosen randomly as a validation set where the rest of the data as a training initiation. The validation data have been used in order to avoid overfitting problem of the training data as well as to dispense guidance for stopping the network training. The network has been trained using a sequential technique, with a sixty-four-batch size. Parameters of momentum, learning rate, along with decay of weight have been settled to 0.9, 0.0001, and 0.005 sequentially. The input of 1st layer consists of RGB image with a dimension of 224x224 pixels which has been converted with sixty four trained 3x3-kernel filters. As a result, 64 feature maps have been generated with a size of 224x224. On the next level, the full pooling has been implemented in the size of 2 x 2 after adding the two convolution layers in order to preserve the invariant translation features in output results as well as to add the drop-out layer. The implementation of pooling helped to generate a high precision performance. After applying all five convolution layers, the outputs have flattened as well as transmitted through two linked layers in order to generate the final output. At the end of architecture, the model for 50 epochs has been educated.

G. Metrics

The model's accuracy, recall, specificity and F1 score have been determined to assess the model's efficiency.
Here, precision describes how accurate one's model is from these positive predictions; that determines about the number of actual positives. The expression of precision has been shown in equation 6.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \qquad (6)$$

The recall which has been used in this system, evaluates if the proportion of actual positives are identified correctly or not. It has been represented through equation 7.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \qquad (7)$$

F1 Score has been determined in this research to measure and check the balance between precision and recall as well as uneven distribution of classes. The expression of F1 score determination has been presented by equation 8.

$$F1 = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \qquad (8)$$

V. Result and Evaluation

The CNN model has been assessed through 990 photos in process of training. All these images have been pre-processed using the same techniques as described in the early sections. The variable parameters of the convolution layers have been fixed during the training process according to the results obtained. Whether if learning process has been progressing as expected or whether the network has learnt enough, have been determined through observing the changes in the curve.

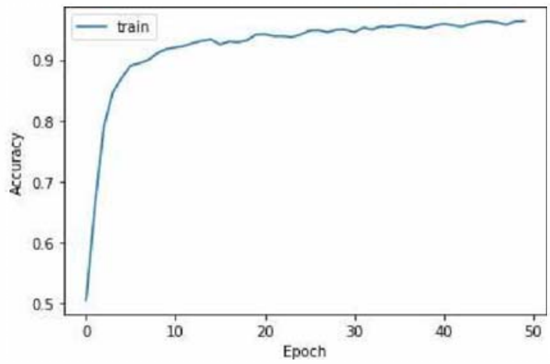

Fig. 5. Model Accuracy (Train)

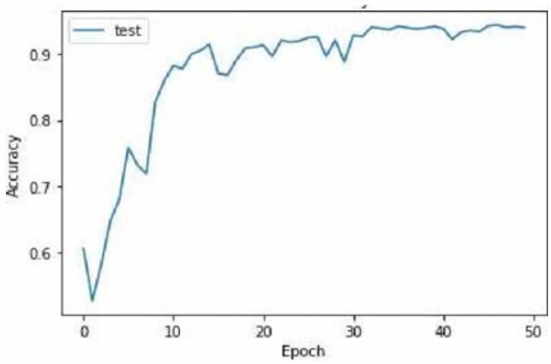

Fig. 6. Model Accuracy (Test)

The overall average accuracy of this method with respect to 50 epochs is 93.58 percent. The variations in the accuracy of the test and train during the training has been presented in Fig 5 and Fig 6.

The loss function tends to keep going down as long as the learning processing keeps going well. The variations in the loss of the test and train during the training have been described in Fig 7 and Fig 8.

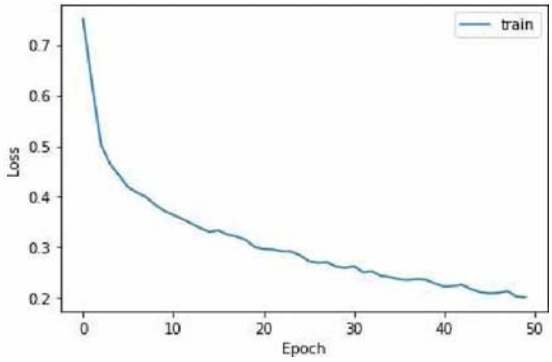

Fig. 7. Model Lost (Train)

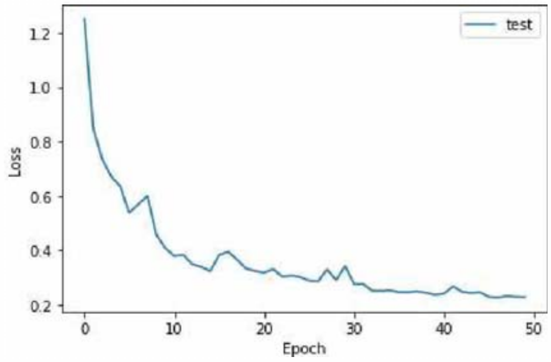

Fig. 8. Model Loss (Test)

Here, the curve in the Fig 8 shows the categorical cross-entropy loss in the network. The overall average loss of this method with respect to 50 epochs is 21.5 percent.

After training the model, the values of accuracy, recall, precision, F1 score have been shown in Table 1 to represent the efficiency of the model.

TABLE I
RESULT OF ACCURACY, PRECISION, RECALL AND F1 SCORE

Parameter	Result in %
Accuracy	93.58
Precision	88.84
Recall	91.70
F1 Score	88.76

While applying ResNet 50 architecture on this same dataset, 83% accuracy was achieved in another research [14].

VI. Conclusion

To recapitulate, it can be said that the death rate behind melanoma cancer tends to be much larger if it's not detected in early stages. As a result, we have proposed this system of melanoma cancer detection as it has higher accuracy of about 93.58% in performance than most other conventional detection systems. Moreover, the effects and changes in accuracy of using multi-layered CNN in melanoma cancer detection systems which has been observed in this research implementation. Additionally, it also reduces the time in finding the result, as our system can provide the output almost instantly. The accuracy might get improved by applying hair removal techniques which are planned to implement later on. This system has high preciseness in performance and cost effective as well which can fill up the need of an affordable rapid test software that helps all walks of people to detect melanoma cancer at early stages.

References

[1] V. Madan, J. T. Lear, and R.-M. Szeimies, "Non-melanoma skin cancer," *The lancet*, vol. 375, no. 9715, pp. 673–685, 2010.

[2] A. Katalinic, A. Waldmann, M. A. Weinstock, A. C. Geller, N. Eisemann, R. Greinert, B. Volkmer, and E. Breitbart, "Does skin cancer screening save lives? an observational study comparing trends in melanoma mortality in regions with and without screening," *Cancer*, vol. 118, no. 21, pp. 5395–5402, 2012.

[3] J. T. Chen, S. J. Kempton, and V. K. Rao, "The economics of skin cancer: an analysis of medicare payment data," *Plastic and Reconstructive Surgery Global Open*, vol. 4, no. 9, 2016.

[4] Z. Khazaei, F. Ghorat, A. Jarrahi, H. Adineh, M. Sohrabivafa, and E. Goodarzi, "Global incidence and mortality of skin cancer by histological subtype and its relationship with the human development index (hdi); an ecology study in 2018," *World Cancer Res J*, vol. 6, no. 2, p. e13, 2019.

[5] American Cancer Society, "Cancer facts figures 2020," 2020. [Online]. Available: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2020/cancer-facts-and-figures-2020.pdf

[6] A. C. Geller, S. M. Swetter, and M. A. Weinstock, "Focus on early detection to reduce melanoma deaths," *Journal of Investigative Dermatology*, vol. 135, no. 4, pp. 947–949, 2015.

[7] "Skin cancer," 2019. [Online]. Available: https://www.webmd.com/melanoma-skin-cancer/understanding-skin-cancer-basics1

[8] H. N, N. B.S, A. K.P, and Navaneeth, "Detection of skin cancer using deep cnn," *International Journal of Recent Technology and Engineering Special Issue*, vol. 8, no. 5S, p. 22–24, Jan 2020. [Online]. Available: https://www.ijrte.org/wp-content/uploads/papers/v8i5s/E10050285S20.pdf

[9] C. Yu, S. Yang, W. Kim, J. Jung, K.-Y. Chung, S. W. Lee, and B. Oh, "Correction: Acral melanoma detection using a convolutional neural network for dermoscopy images," *PloS one*, vol. 13, no. 4, p. e0196621, 2018.

[10] D. Sharma and S. Srivastava, "Automatically detection of skin cancer by classification of neural network," *International Journal of Engineering and Technical Research*, vol. 4, no. 1, pp. 15–18, 2016.

[11] R. Amelard, J. Glaister, A. Wong, and D. A. Clausi, "Melanoma decision support using lighting-corrected intuitive feature models," in *Computer Vision Techniques for the Diagnosis of Skin Cancer*. Springer, 2014, pp. 193–219.

[12] M. ur Rehman, S. H. Khan, S. D. Rizvi, Z. Abbas, and A. Zafar, "Classification of skin lesion by interference of segmentation and convolotion neural network," in *2018 2nd International Conference on Engineering Innovation (ICEI)*. IEEE, 2018, pp. 81–85.

[13] P. Ridell and H. Spett, "Training set size for skin cancer classification using google's inception v3," 2017.

[14] C. Fanconi, "Skin cancer: Malignant vs. benign," 2019. [Online]. Available: https://www.kaggle.com/fanconic/skin-cancer-malignant-vs-benign

[15] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[16] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.