Automatic diagnosis of melanoma using linear and nonlinear features from digital image

Conference Paper *in* Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference · July 2017

DOI: 10.1109/EMBC.2017.8037802

4 authors:

Tamanna Tabassum Khan Munia
Michigan State University
28 PUBLICATIONS **214** CITATIONS
SEE PROFILE

Md Alam
University of North Dakota
7 PUBLICATIONS **116** CITATIONS
SEE PROFILE

Jeremiah Neubert
University of North Dakota
54 PUBLICATIONS **320** CITATIONS
SEE PROFILE

Reza Fazel-Rezai
University of North Dakota
140 PUBLICATIONS **2,196** CITATIONS
SEE PROFILE

Automatic Diagnosis of Melanoma from Dermoscopic Image Using Real-Time Object Detection

Shudipto Sekhar Roy, Akkas Uddin Haque, and Jeremiah Neubert

Abstract— **Among all types of skin cancer, a melanoma is the deadliest. Melanoma is typically a small, usually black or brown colored mole, which can develop anywhere on the skin. Detecting melanoma in its earliest stages is among the most important factors in improving the outcome of a melanoma diagnosis. In this paper, a real-time object detection technique is used to automatically detect melanoma in dermoscopic images. For detecting melanoma in real-time a state-of-the art detection model named YOLOv2 (You Only Look Once: version 2) is used. YOLOv2 uses a single neural network to the full image, enabling real-time performance. It is capable of processing images at 40-60 fps using a Titan X GPU. Our proposed model predicts the diagnosis of a mole with an accuracy of 86.00%, sensitivity = 86.35% and specificity = 85.90%. In addition, the proposed method is shown to be invariant to the presence of hair in the image.**

I. Introduction

Though Melanoma covers only 1% of all skin cancers, it is responsible for the highest rate of death among all skin cancer related casualties [1]. A statistic from The Skin Cancer Foundation shows that melanoma accounts for approximately 10130 deaths in the US annually [2]. For 2017, the American Cancer Society estimates 87,110 new incidences of melanoma resulting in 9,730 deaths [1]. To prevent these deaths, a feasible solution for rapid and early detection of melanoma has become a research topic. Like other cancerous cells, melanoma can rapidly spread to other parts of the body causing severe damage. However, the probability of melanoma cells becoming metastasized remains low in early stages. Physicians can take proper steps to keep this cancer under control, if diagnosed and treated in these early periods.

One common way of diagnosing skin cancer is through dermoscopic images. Dermoscopy is a technique of imaging an area of skin using a high-quality magnifying lens, and lighting system. It can capture images with enough detail that skin structure and patterns can be examined through the image [3]. Dermatologists use this type of images to apply one of the several melanoma diagnosis methods. One of these methods is ABCD (Asymmetrical Shape, Border, Color, Diameter) rule where a scoring process takes place based on the presence of those four different features [4]. A combined score higher than 5.45, diagnoses a lesion positive for melanoma. Menzies method for detecting melanoma is also practiced by dermatologists [5]. This is also a scoring method where dermoscopic features are divided into two types, namely, positive and negative features. Positive features consist of 9 identifications such as, blue-white veil, multiple brown dots, radial streaming, etc. On the other hand, negative features include symmetry of pattern and presence of single color. A lesion is diagnosed as melanoma if it has at least one or more of the positive features and neither of both negative features. Another method of diagnosing melanoma is the seven-point checklist method [6]. This widely used scoring method divides seven dermoscopic features, such as atypical pigment network, irregular streaks, atypical vascular pattern, etc., into two main groups, namely, major and minor. Major group features are responsible for a score of 2 and minor group features are bound for 1. A combined score higher than 3 detects a lesion as melanoma.

Though, dermoscopy is the most effective technique for detecting melanoma, the reliability of the detection also depends on the operating skill of the dermatologists. As the detection depends on human vision and previous experience, making it automatic is an encouraged research topic to pursue. One work shows that, a bag-of-features classification method can be used on dermoscopic images for automatically detect melanoma [7]. Two methods were presented as Global and Local for the classification. The Global method was performed by automatic segmentation, followed by an extraction of color and texture features for training the classifier. The Local method was inclined towards image analysis and recognition. Recently, to make this detection process more approachable for the general population, digital images are being used. One work with digital image used a combination of Otsu and k-means clustering segmentation methods for detecting the affected area, and extracting several linear and non-linear features from the lesion portion [8]. These features were then evaluated with a machine learning model consisting of five different classifiers. Two web applications are available which can perform analysis on digital images for classifying melanoma mole from benign one. However, the lower accuracy of these applications makes it difficult to rely for making clinical decision [9]. Another work shows that, an automatic analysis of images is often impaired due to the presence of bodily hair. Using percolation algorithm, clusters of connected points can be analyzed in an image and a linear shape of the cluster, along with lower pixel intensity can indicate the presence of hair [10]. However, making the system fully automatic with a satisfactory accuracy rate, and run in real-time is an open research topic.

In this paper, we propose a method which can automatically detect the presence of melanoma characteristics in a mole from dermoscopic image and provides a real-time

S. Sekhar Roy is with the Mechanical Engineering Department, University of North Dakota, Grand Forks, ND 58202, USA (e-mail: shudipto.sekharroy@und.edu).

A. U. Haque is with the Mechanical Engineering Department, University of North Dakota, Grand Forks, ND 58202, USA (e-mail: akkasuddin.haque@und.edu).

J. Neubert is with the Mechanical Engineering Department, University of North Dakota, Grand Forks, ND 58202, USA (e-mail: jeremiah.neubert@engr.und.edu).

prediction percentage about possible condition of that mole, using a real-time object detection technique (YOLOv2) [11].

II. Material And Methodology

For facilitating the computer-aided diagnosis related work on melanoma many datasets can be found online. In this work, PH2 dataset is used, which is a dermoscopic image database [12]. The samples used in this database are acquired from the Dermatology Service of Hospital Pedro Hispano, Matosinhos, Portugal. This database consists of a total of 200 dermoscopic 8-bit RGB color images of melanocytic lesions, together with 80 common nevi, 80 atypical nevi, and 40 melanomas. These images are captured using a magnification of 20x with a resolution of 768x560 pixels. For convenience of data extraction, a binary mask of the segmented lesions is also given, along with the original dermoscopic images. The classification of all images can be found within the dataset package.

Our work can be divided into four main steps including, Data Extraction, Data Augmentation, Training, and Testing as described in the following sections:

A. Data Extraction

This section describes the process of extracting lesion area information from the dataset. As mentioned earlier, along with each dermoscopic image, a binary masked version was also present in the dataset package. This binary masked image was of same size as the original where the pixels of lesion area were white and rest of the image was black. From this binary masked image, a rectangular area containing the lesion was determined as a bounding box location. By obtaining the minimum position values of the white pixels along the two axes, top-left corner of the bounding box was found. Similarly, the maximum position values along the two axes gave the bottom-right corner of the lesion confining bounding box. For each image set, the calculated bounding box from the binary masked image was drawn on the original image to justify the detection of lesion. These coordinates were used later for training the system. Fig. 1 provides a sample dermoscopic image, along with its binary mask, and the extracted region on the image.

(a) (b) (c)

Figure 1. Sample of a dermoscopic image showing region extraction from binary masked version: (a) dermoscopic image, (b) binary image, (c) extracted region.

Clinical classification of each image was provided in the classification text file found in the dataset. Finally, for each image a new data file was created for labeling purpose where the clinical type was presented by a numeric value (e.g. "1" for melanoma and "0" for other types) and the normalized center coordinates, width, and height of the lesion bounding box were included.

B. Data Augmentation

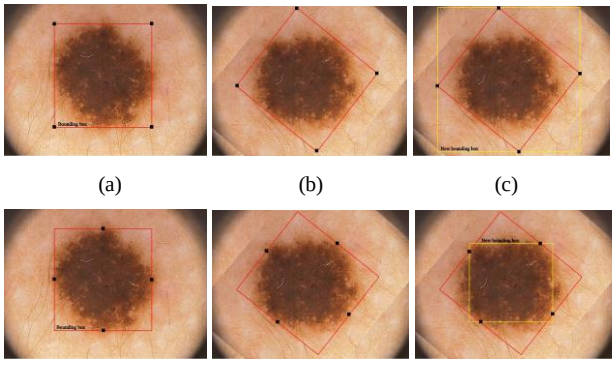
Using machine learning for detecting objects requires a large dataset of images to provide desired performance. As the dataset utilized only 200 images, augmentation of the image data was necessary. The first method of data augmentation was performed through rotation. Rotation of the images was necessary as the system could be trained with the same image features as viewed from several different angles. Each image was rotated by ten degrees from angular positions zero to 350. As a result, from every single image, another 35 images were generated. To rotate the image, a 2x3 transformation matrix was calculated. Given the center of the image, angle of rotation and, scale value, the transformation matrix becomes:

$$\begin{bmatrix} \alpha & \beta & (1-\alpha).center.x - \beta.center.y \\ -\beta & \alpha & \beta.center.x + (1-\alpha).center.y \end{bmatrix}. \quad (1)$$

Equation (1) represents the transformation matrix, where, α=scale.cos(angle), β=scale.sin(angle). Scale value was kept as 1 because the images were not gone through any zooming process. Applying this transformation matrix in the following equation provided the desired rotation in the output image:

$$dst(x, y) = src(M_{11}x + M_{12}y + M_{13}, M_{21}x + M_{22}y + M_{23}). \quad (2)$$

Equation (2) is used for getting the desired rotation, where, M=Transformation Matrix, src=input image, dst=output image. When an image got rotated, the bounding box containing the lesion area was also needed to be re-calculated. The computed transformation matrix was used again to obtain the shifted bounding box defining coordinates after rotating an image. Then a new bounding box, confining the lesion, was calculated using the shifted coordinates of the previous bounding box. At first, the corner points of the lesion bounding box from the original image were being used for getting a new bounding box on a rotated image. But after getting a new bounding box it was observed that some unnecessary background was getting covered by this new box, which was misleading for the system while training. So, edge centers of the original bounding box were used instead of corners for getting the new bounding box. By doing this the new lesion bounding box on the rotated image could cover the maximum portion of the lesion without taking any needless background. Fig. 2 demonstrates both rotation methods by rotating a sample image by 50 degrees.



(a) (b) (c)

(d) (e) (f)

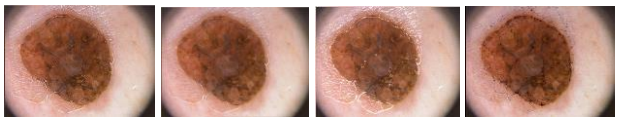
Figure 2. Sample of a dermoscopic image undergoing rotation methods: (a) using corner points of the bounding box, (b) rotating the image about 50 degrees, (c) a new bounding box determined using previous corner points, (d) using edge centers of the bounding box, (e) rotating the image about 50 degrees, (f) a new bounding box determined using previous edge centers.

For maintaining the size of the rotated image as of the original one, a pixel extrapolation method was used. Every rotated image was patched over the clone of the original one, so the missing pixels outside the edge of the rotated image can be compensated by the original ones.

Another method of augmentation was performed by blurring the whole image set. This will provide the invariance to possible blurred photos. The gaussian filter implemented to blur the image can be described as:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad . \qquad (3)$$

Equation (3) represents a 2D Gaussian, where, σ =standard deviation of the distribution and assuming the distribution has a mean of zero. Two other morphological augmentation processes were performed as "dilation" and "erosion" to the image set. But the accuracy of the system dropped by around 11.2% from the present value. After dilating and going through erosion the image features were getting modified, such as the boundary line, separating the lesion from the background was getting reshaped, texture as well as color combination of the lesion were getting changed and so on. So, these two augmenting processes were tested and got rejected for the betterment of system performance. Fig. 3 shows a sample of dermoscopic image and its different versions after modifying with image processing techniques.



(a) (b) (c) (d)

Figure 3. Sample of a dermoscopic image along with its processed versions for data augmentation: (a) original image, (b) image after blurring, (c) image after dilation, (d) image after erosion.

C. Training

For each training stage, one-fold was removed from training and separated as a validation set for the testing purposes. This was performed to ensure each image in the data set was tested once. For training the system a state-of-the-art real-time object detection system named YOLOv2 (You Only Look Once: version 2) was used. YOLOv2 applies a single neural network on an image, and divides the full image into regions. The network then predicts bounding boxes and applies class probabilities for each region. We ran YOLOv2 on an NVIDIA TITAN X GPU allowing it to process images at 40-90 FPS. To train YOLOv2 two sets of information were required. These two sets of information included the images to train with, and a corresponding label for each image. The labelling document comprised the clinical type referred by a numeric digit along with the normalized center coordinates, followed by the width and height of the region of interest on that image as a bounding box for the lesion. Using an optimal learning rate helps neural network to be quick enough to get trained with significant features. In this work, a learning rate of 0.0001 was utilized. The CNN model developed was fine-tuned using the pretrained weights as found on the YOLO website [13]. The purpose of using a pretrained model is that it provides well differentiated weights for each class, leading to faster and more accurate training results. While training, the convolutional neural network fine-tunes these weights by adjusting them to minimize a loss function to classify the object as the annotated training images through a supervised learning process. This process of adjusting the weights is called backpropagation. One iteration is represented by a single completion of this process. For this paper, each fold of the five-fold cross validation was trained at 35,000 iterations. The final weight for each of the five-folds was used to cross validate the data kept separate for testing.

D. Testing

After obtaining the final weights from each training fold, it was used for testing in the cross validation. The set of images not included while training was used for testing with that finalized set of weights. The output of the testing provided the detected type, along with the predicted probability for each tested image. Fig. 4 presents two sample detections from testing.

(a) (b)

Figure 4. Detection samples with prediction percentage: (a) detection of melanoma, (b) detection of non-melanoma.

Threshold for detection is an important parameter to determine while testing the dataset. A specific value of threshold allows a certain range of predicted probability to be claimed as detected. Here, we ran the test by varying the threshold from 0.0 to 1.0 for each 0.05 interval to generate a Receiver Operating Characteristic (ROC) curve [14]. The ROC curve was used to evaluate the response of the binary classifier as either melanoma or non-melanoma (benign). The ROC curve was generated by plotting the true positive rate (TPR) against the false positive rate (FPR) at different thresholds. The TPR, or, Sensitivity = Σ True positive/Σ Condition positive. The FPR, or, Fall-out = Σ False positive/Σ Condition negative. Accuracy (ACC) = (Σ True positive + Σ True negative)/Σ Total population. And, the area under the curve (AUC) is a measurement of how well the classifier can distinguish between two groups. An acceptable threshold value was found through analyzing these curves.

III. Result And Analysis

The diagnostic ability of this binary classifier model was evaluated by using ROC curve analysis. The observation was included with two population sets, one with melanoma and other without melanoma. The ROC curve of each cross-validation fold showed that the approximated optimal threshold value can be used as 0.7. Fig. 5 shows the ROC curves obtained for each fold of five-fold cross validation.

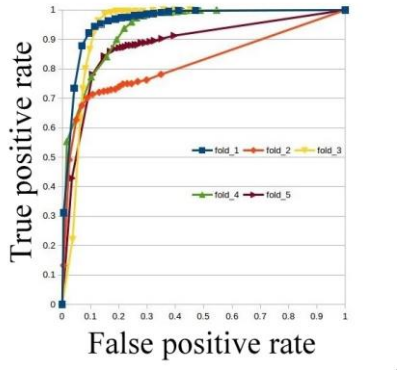


Figure 5. ROC curves for five-fold cross validation

TABLE I: Data for five-fold cross validation with threshold = 0.7

Validation Set	*Specificity*	*Sensitivity*	*Accuracy (%)*	*AUC*
01	0.8845	0.9444	89.65	0.96
02	0.9115	0.7014	86.94	0.82
03	0.9028	0.8663	89.55	0.94
04	0.7808	0.9375	81.22	0.94
05	0.8155	0.8680	82.60	0.88
Average	**0.8590**	**0.8635**	**86.00**	**0.91**

From Table 1, we can confirm that the proposed model was able to diagnose Melanoma effectively in real time. It was also observed that the model was invariant to the presence of any kind of bodily hair while detecting a mole. Fig. 6 provides a sample detection which was clinically correct, while being unaffected by the presence of hair.

Figure 6. Correct detection in the presence of hair.

In comparison, previous works utilizing the same database introduced a global method with Sensitivity = 96%, Specificity = 80%, and a local method with Sensitivity =100%, Specificity = 75% [7]. Where, our model achieves a result of Sensitivity = 86.35% and, Specificity = 85.90% which shows a clear improvement in specificity value.

IV. Conclusion

The proposed work was aimed to develop a method for detecting melanoma using an object detection technique which can run in real time, while working with dermoscopic images. To provide an efficient system, the YOLOv2 convolutional neural network was trained on melanoma data. An available dataset with dermoscopic images of both malignant melanoma and benign moles was used for this purpose. For executing the training, every image from that dataset was annotated with the clinical classification of the mole. For making the system robust, the data was augmented for training. To augment the data, each image was rotated, blurred, and then annotated. Dilation and erosion were also applied for augmentation but rejected because of the poor performance of the system. The system was trained using a state-of-the art object detection technique YOLOv2. The trained system was tested using a five-fold cross validation technique and the results provided an accuracy of 87% while showing invariance to detection of problematic hair. The proposed system can run on a desktop computer utilizing a Titan-X GPU. Future work could include development of a mobile application where the user can take images of a suspicious mole, and the mobile app may remotely run online to image and present the predicted result in real time.

References

[1] "Key Statistics for Melanoma Skin Cancer." [Online]. Available: https://www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics.html

[2] "Melanoma - SkinCancer.org." [Online]. Available: http://www.skincancer.org/skin-cancer-information/melanoma

[3] "Dermoscopy Tutorial." [Online]. Available: http://www.dermoscopy.org/atlas/base.htm

[4] O. B.-F. Stolz, W, A. Riemann, Armand B. Cognetta, "ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma," *EJD. Eur. J. dermatology*, vol. 4, pp. 521–527, 1994.

[5] S. W. Menzies, C. Ingvar, K. A. Crotty, and W. H. McCarthy, "Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features.," *Arch. Dermatol.*, vol. 132, no. 10, pp. 1178–82, 1996.

[6] G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, and M. Delfino, "Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis.," *Arch. Dermatol.*, vol. 134, no. 12, pp. 1563–1570, 1998.

[7] C. Barata, M. Ruela, M. Francisco, T. Mendonca, and J. S. Marques, "Two Systems for the Detection of Melanomas in Dermoscopy Images Using Texture and Color Features," *IEEE Syst. J.*, vol. 8, no. 3, pp. 965–979, Sep. 2014.

[8] T. T. K. Munia, M. N. Alam, J. Neubert, and R. Fazel-rezai, "Automatic Diagnosis of Melanoma Using Linear and Nonlinear Features from Digital Image," In Press; 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2017.

[9] J. A. Wolf et al., "Diagnostic Inaccuracy of Smartphone Applications for Melanoma Detection," JAMA Dermatology, vol. 149, no. 4, p. 422, 2013.

[10] A. Afonso and M. Silveira, "Hair detection in dermoscopic images using Percolation," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2012, pp. 4378–4381.

[11] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," Dec. 2016.

[12] "ADDI - Automatic computer-based Diagnosis system for Dermoscopy Images." [Online]. Available: https://www.fc.up.pt/addi/ph2 database.html. [Accessed: 01-Sep-2017].

[13] "YOLO: Real-Time Object Detection." [Online]. Available: https://pjreddie.com/darknet/yolo/

[14] D. J. Hand, "Measuring classifier performance: A coherent alternative to the area under the ROC curve," Mach. Learn., vol. 77, no. 1, pp. 103–123, 2009.