Tariq et al LGURJCSIT 2018

LGURJCSIT
ISSN: 2519-7991

LGU R Jounral for
Computer Sciences IT

Vol. 2 Issue 3, July - September 2018

A Review on Skin Cancer Data Using Image Processing

Muhammad Arslan Tariq, Rafaqat Alam Khan
Department of Computer Science Lahore Garrison University (LGU), Pakistan

Abstract

The wireless industry is going very fast nowadays. We can easily see the evolution from 2G to 3G and now advance to the 4G and 5G network. Before wireless networks, wired networks were commonly used in every field. But there were some disadvantages regarding mobility, quality of service and connectivity. Wired network bounded the region of the working area for the internet and it requires multiple wires to connect computer from one device to another. While on the other hand wireless network is an open source for everyone to use the internet. There is no limitation of the region and no issue regarding connectivity because data is transfer through signal which includes frequency in the form of waves. But there are also some disadvantages of wireless network regarding cost, speed, coverage, bandwidth etc. If we talk about the better network so it depends on the situation and problem.
Keywords: Wired, Wireless, Network, Security, Internet

1. INTRODUCTION

Skin cancer is the main source of human casualty these days with 9,000 passing's every year. In United States more than 5 million individuals are determined to have skin cancer consistently. Numerous scientists are working here for robotized examination yet so far, nothing near or unified framework is being executed up until now. For better order of Skin sore for cancer discovery, scientists are utilized dermoscopic pictures. Dermoscopic pictures are gotten through imaging system in which the impressions of the skin surface are dispensed. Because of which skin further area are appropriately imagined and improved for better characterization. Then again dermatologist additionally preferred dermocospic pictures when contrasted with the standard photos for indicative exactness and furthermore extensive number of dermoscopic connections which are going to achieve the market for advanced cell clients and that would prompt expand the impact better patient care.

2. LITERATURE REVIEW

In 2009, G. Capdehourat et al. introduced a machine adapting approach to manage and arrange melanocytic sores in unsafe form of dermatoscopic pictures. The photo database is made out of 433 generous wounds and 80 injuries. After a photo handling stage that fuses hair departure separating, each photo is thus divided using understood picture division calculations. On that point, each sore is portrayed by a part of that vector which contains shape, shading and association data, and besides neighborhood and overall parameters that undertake to reflect structures utilized as a bit of restorative conclusion. The learning and portrayal organizing is performed to utilizing AdaBoost.M1 with C4.5 decision trees. For the subsequently separated database, then arranged and passed on a false positive rate of 8.75% for an affectability of 95%. A comparative portrayal philosophy related with physically separated pictures by a specialist dermatologist yielded a false positive rate of 4.62% for an affectability of 95%.

These results are confirmed and seems more better than which are discussed in this literature. Execution appraisal is sensitive since each and every revealed outcome were gotten using unmistakable databases. Starting at this moment, improvement of a tremendous database of dermatoscopic pictures that could be

used as reference test bed that shows up of being a basic issue

The point of this research is to grow such a better picture robotized determination device of melanoma from dermoscopic pictures. The undertaking is divided in two 3 noteworthy stages and are clarified as below.

2.1. Segmentation of Lesion area

The main objective of this task is, the automatic segmentation of dermoscopic images boundaries would be obtained.

Original Image Segment on Mask

The data consists of original image along with segmented lesion area and the doctor report related to the patient diagnosis would be extracted using natural language processing.There are 2000 images which are available in JPEG format. These images data would be further provided to classifiers for classification. The different classifiers used for classification are SVM, Deep Learning and Neural Network. The different metrics that would be obtained and on the basis of which our results would support scientific completion are as follows.

- Accuracy
- Dice Coefficient
- Jaccard Index
- Specificity
- Sensitivity
- ROC Curve
- Average precision

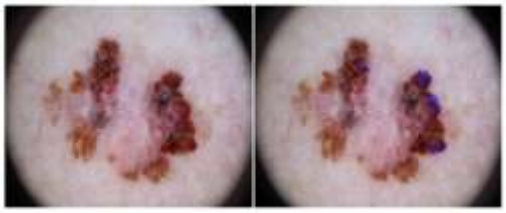
The training would be done on binary mask lesion images having PNG format. In these binary masks lesion images “0” would represent the background area i.e. the area outside the lesion area and “255” would represent the area inside the lesion.

2.1. Localization and Detection of Dermoscopic Visual Patterns/ Features

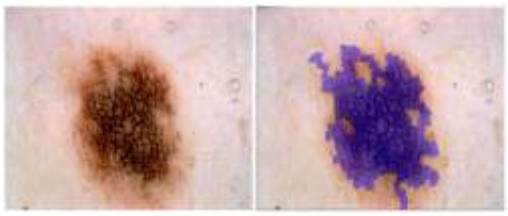
In this part of Skin lesion, on the basis of super pixel tile automated prediction of dermoscipic features would be obtained. The feature data would compromise of original lesion image, Super-pixel mask and super-pixel mapped annotation consists of the following different features.

- Streaks
- Negative Network
- Milia-like Cyst
- Network

Streaks



Pigment Network



The SLICO algorithms would be used to subdivide the lesion images in to super-pixels to minimize the variability and dimensionality. The training data consists of 2000 lesion images in JPEG format and the 2000 super-pixel masks are in PNG format. The ground truth annotation consisting of 2000 dermoscopic features are in JSON format. The JSON number would be either "0" or "1", where "0" would represent absence of dermoscopic feature while "1" would represent presence of dermoscopic feature. The different metrics that would be used for evaluation purpose are given below.

- Accuracy
- Specificity
- Sensitivity
- ROC curve
- Average Precision

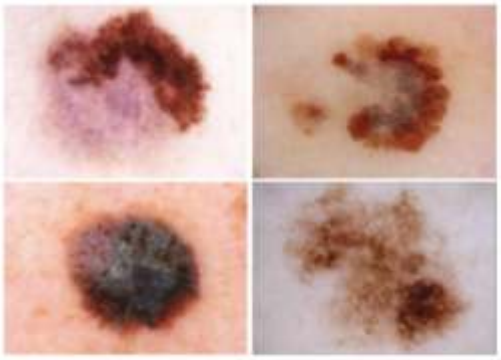
3. CLASSIFICATION OF DISEASES

In this part the classification of two different binary classifications would be performed for diagnosis of skin lesion images i.e. The binary classification would be done between Melanoma (1) and Nevus, Seborrheic Keratosis (2, 3), and Melanoma, Nevis (1,2) and Seborrheic Keratosis (3) respectively.

3.1. Melanoma (Malignant Skin Tumor)

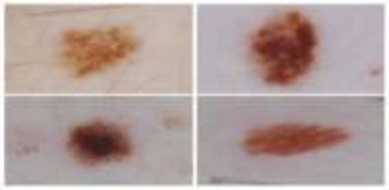
Melanoma is the deadliest form or last stage of the skin cancer if it is not diagnosed in his early stages. They appear as a mole on the skin and they make pigments inside skin cells known as melanocytes.

Melanomma



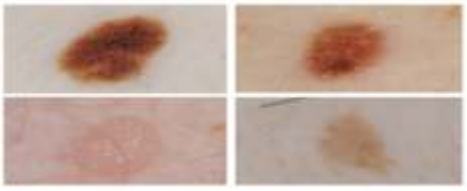
3.2. Nevus (Benign Skin Tumor)

Nevus



3. Seborrheic Keratosis (Benign Skin Tumor)

Seborrheic eratosis



The training data consists of 2000 images out of which melanoma cases are 374, Seborrheic Keratosis are 254 and the remaining 1372 are nevi cases. The data is in CSV format consists of three columns i.e. image id, age approximate and sex. The ground truth data is in CSV format compromise of three columns in which first column contains images id and in second column consists of first binary classification task in which "1" is used for lesion melanoma and "0" is used for lesion Nevi or Seborrheic Keratosis, while in third column of CSV the second binary classification in which "1" is used lesion Seborrheic Keratosis and "0" for melanoma or Nevi.

The binary classification metrics used for evaluation purpose are as given below.

- Accuracy at 0.5 threshold
- Sensitivity at 0.5 threshold
- Specificity at 0.5 threshold
- Average precision
- AUC
- ROC curve

For all the three tasks discussed above the same dataset would be applied to get the desired results.

4. RESULTS AND DISCUSSION

The medical reporting such as dermatologist report consists of specific wording (features) related to the particular disease organized into report related section. For machine learning, it requires to organize unstructured text into structured text information for understanding and learning the machine. In this regard, NLP based methods are applied to extract information from human text report automatically and machine learning approaches are used for association of extracted text information with images automatically.

In this paper we explore how to extract the features from human written dermascopic reports using NLP based methods? How to automatically associate dermascopic images with report? How to automatically classify and predict skin cancer based on extracted information.

5. RELATED WORK

Several works have been reported regarding

segmentation, localization and classification of medical images. [3] Kisilev, P et al. proposed a semi-automated system that helps radiologists in diagnosing the medical images. They extract features from textual description and mapped them to the images. The work is done based on diagnosis descriptions (features) of radiology images. In another work Kisilev, P et al. [3] constructed a feature-based classifier for medical images. This classifier uses features obtained from image segmentation to classify images into categories. Zhang et al. [5] proposed MdNet for automatic report generation by mapping radiology images and textual diagnosis reports. Shin, H. et al. [9] proposed method that used disease name instead features for labeling the images.

In this field of study most of the images in a dataset consists of diseased images. However, all cases are not diseased cases. Therefore, diagnosing the disease in an unbalanced dataset in which most images are normal is more challengeable.

References

[1] Bird, Steven, and Edward Loper. "NLTK: the natural language toolkit." Proceedings of the ACL 2004 on Interactive poster and demonstration sessions. Association for Computational Linguistics, 2004.

[2] German Capdehourat, Andres Corez, Anabella Bazzano, and Pablo Mus, Pigmented Skin Lesions Classification Using Dermatoscopic Images, Springer Verlag Berlin Heidelberg, 2009, pp. 537-544

[3] Kisilev, P., Walach, E., Hashoul, S. Y., Barkan, E., Ophir, B., & Alpert, S. (2015). Semantic description of medical image findings: structured learning approach. In BMVC (pp. 171-1).

[4] Kisilev, P., Walach, E., Barkan, E., Ophir, B., Alpert, S., & Hashoul, S. Y. (2015). From medical image to automatic medical report generation. IBM Journal of Research and Development, 59(2/3), 2-1.

[5] Zhang, Z., Xie, Y., Xing, F., McGough, M., & Yang, L. (2017, July). Mdnet: A semantically and visually interpretable medical image diagnosis network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 6428-6436).

[6] Simonyan, K., Vedaldi, A., & Zisserman, A. (2013). Deep inside convolutional networks: Visualising image classification models and saliency maps. arXiv preprint arXiv:1312.6034.

[7] H.-C. Shin, L. Lu, L. Kim, A. Seff, J. Yao, and R. M. Summers. Interleaved text/image deep mining on a very large-scale radiology database. In CVPR, 2015.

[8] H.-C. Shin, K. Roberts, L. Lu, D. Demner-Fushman, J. Yao, R.M. Summers, learning to read chest X-rays: recurrent neural cascade model for automated image annotation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[9] Shin, H. C., Roberts, K., Lu, L., Demner-Fushman, D., Yao, J., & Summers, R. M. (2016). Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2497-2506).

[10]Mariam A.Sheha,MaiS.Mabrouk," Automatic Detection of Melanoma Skin Cancer",[2012],International Journalof Computer Applications (0975 -8887) March,2012.

[11] Alexandros Karargyris, Orestis Karargyris, AlexandrosPantelopoulos, "An Advanced image-Processing Mobile Application for Monitoring Skin Cancer",2012,Conferene peper of IEEE 24th International Conference.

[12]CharalamposDoukas, Paris Stagkopoulos, Chris T. Kiranoudis, and IliasMaglogiannis"Automated Skin Lesion Assessment", 2012,34th Annual International Conference of the IEEE EMBS San Diego, California USA.

[13]E. Barati, M.Saraee, A.Mohammadi, N. Adibi and M. R. Ahamadzadeh,"A Survey on Utilization of Data Mining Approaches for Dermatological (Skin) Diseases Prediction",2011,Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Health Informatics (JSHI).

[14]Suhail M. Odeh,"Automatic Diagnosis of Skin Cancer", 2010,Journal of Communication and Computer 8 (2011) 751-755 at David Publishing.

[15]Radu Dobrescu, Mateidobrescu, Stefan Mocnu, Danpopescu,"Medical images classification for skin cancer", 2010,WseasTransactionson Biologyand Biomedicine;ISSN-1109-9518

[16]Gabriella Fabbrocini, Giovanni Betta, Giuseppe Di Leo, Consolatina Liguori, Alfredo Paolillo,AntonioPietrosanto, Paolo Sommella, Orsola Rescigno, Sara Cacciapuoti, Francesco Pastore,Valerio De Vita, Ines Mordente and Fabio Ayala,"Epiluminescence Image Processing for Melanocytic Skin Lesion Diagnosis Based on 7-Point Check-List", 2010,The Open Dermatology Journal, 2010,Vol 4.