SKIN CANCER DETECTION USING CAPSULE NET

[1]MEENALOSINI VIMAL CRUZ, [2]SIBI CHAKKARAVARTHY, [3]ANUPAMA NAMBURU, [4]JARED BASILE

Keene State College, VIT India
E-mail: [1]Meenalosini.vimal.cruz@keene.edu, [2]sb.sibi@gmail.com, [3]namburianupama@gmail.com, [4]jared.basile@ksc.keene.edu

Abstract - In recent decades, there have been numerous investigations concentrated on early detection of skin cancer with non-invasive or negligibly intrusive strategies in lieu of conventional excisional biopsy. The early, fast and effective detection of skin cancer is very crucial as it makes treatment easier and increasingly successful. Researchers are proficient in pre-processing the skin images but fail in identifying efficient classifiers for classifying the skin cancer due to complexity with variety of lesion sizes, colors, and shapes. No single classifier is sufficient for classifying the skin cancer.
Recently, convolutional neural networks have played an important role in deep learning. Classification using CNN's have proven successful in various fields. However, CNNs suffer from not taking important spatial relations between features into consideration. Also, they lack rotational invariance. The CNNs classify only if certain features are present in the test data ignoring the relative spatial relation with each other, resulting in false positives. The lack of rotational invariance assigns an object to other classes, leading to false negatives.
The capsule network is designed to overcome the problems associated with CNNs. Capsule networks use modules or capsules other than pooling as an alternative to translational invariance. The capsule network uses layer-based squashing and dynamic routing. The Capsule network uses vector-output capsules and max-pooling with routing by agreement, unlike scale-output feature detectors of CNNs. These assist in avoiding the false positives and false negatives. The capsule network architecture is created with many convolution layers and one capsule layer as the final layer.
Hence, in the proposed work, skin cancer classification is performed based on CapsNet architecture which can work well with high dimensional hyperspectral images of skin.

Index Terms - Computer Aided Diagnosis, Skin Cancer Detection, Capsule Network, CNN.

I. INTRODUCTION

Skin cancer is one of the most common cancers all over the world. It is estimated that approximately 9,500 people in the U.S. are diagnosed with skin cancer everyday [1]. There are 5.4 million new cases of skin cancer in the United States every year [2], and according to the current estimation, one in five Americans will develop skin cancer in their lifetime. Like the other types of cancers, most skin cancers can be cured if it is diagnosed early enough. Skin cancer is primarily diagnosed visually, beginning with an initial clinical screening, followed by dermoscopic analysis, and ultimately a biopsy and histopathological examination. The biopsy method is done by removing or scraping off a skin sample that will undergo a series of laboratory testing, which is painful and time consuming. Dermoscopy is one of the golden techniques to exam skin lesions because it can capture high-resolution images of the skin regardless of interruption from surface reflections [2]. Experienced clinicians use this high-resolution imaging to evaluate the possibility of melanoma at the early stage with considerate accuracy. Unfortunately, there are not enough experienced dermatologists all over the world. In order to solve the problems specified above, computer-aided diagnosis is performed, to classify different skin associated diseases using dermoscopy images. However, within the last decade, computer-aided classification strategies were unfavorable due to various reasons. For instance, the number of cases in the database were inadequate for learning accuracy and for extracting useful features. Moreover, Dermoscopic images are acquired via a specialized instrument and histological images are acquired via invasive biopsy and microscopy. Although both modalities yield highly standardized images, hardware up until now was incapable of handling such data. The proposed approach combines image processing techniques and advanced deep learning techniques. The proposed approach achieves automated classification of skin lesions using images. This is a challenging task because of the close-grained variability in skin lesions. The main contribution of this paper is stated as follows:

- This research uses advanced deep learning architecture to classify the skin cancer from the given images.
- The capsule architecture used in this paper is modified in such a way that it overcomes the drawback of pooling layer. CapsNet has a drastic advantage that it is capable to distinguish the relationship between the features by making the translation-equivariant or viewpoint-equivariant. Hence, the CapsNet is capable of detecting the cancer spot anywhere in the image.
- Experimental validation is carried out using BioGPS and HAM 10000 – Skin cancer dataset.

II. RELATED WORKS

This section discusses the state of the art medical image classification techniques.

Esteva, A. et al [3] proposed a technique produced remarkable low-test errors on classification tasks over different types of images. Further his work exhibited the deep convolutional neural network with very high melanoma classification ability. In their research, the author utilized CNN architecture with well-trained weight on 1.28 million ImageNet, a fine-tuned dataset containing more than 13,000 images. The significance of the mentioned work is that the first-time scientists proved the diagnostic efficiency of computer aided method with deep learning techniques in skin cancer screening.

Alquran et al [4] proposed a technique for skin cancer detection. In their approach author used statistical feature measure such as Gray Level Co-occurrence Matrix (GLCM), Asymmetry, Border, Color, Diameter, (ABCD) and utilized PCA for feature selection.

Soumya et al [5] exhibited a method which utilizes techniques such as colour correlogram and texture analysis. Their method yields better results than the existing skin lesion analysis system. Their method outperforms the existing ABCDE parameters for feature selection.

Neha et al. [6] invented a hardware-based technique which used antenna for skin cancer detection. Their proposed antenna operating in the X-band allows sensing of varying differences in dielectric properties among skin tissues over a wide band of frequency and thus, can be used for skin cancer detection.

III. DRAWBACKS IN CNN

Deep learning in medical image classification is nothing new. Researchers around the world are utilizing the convolutional neural networks (CNNs) to solve common problems in the field of image classification. CNNs were built at first to classify images; they do so by using successive layers of convolutions and pooling. When training a traditional CNN, we only care about whether the model predicts the right classification or not. The pooling layer in a convolutional block is used to reduce the data dimension and achieve something called spatial invariance, which means regardless of where the object is placed in the image, it identifies the object and classifies it. While this is a powerful concept, it has some drawbacks.

The drawback of a CNN is that during pooling it tends to lose a lot of useful information while performing tasks such as image segmentation and object detection. The pooling layer loses the required spatial information due to the rotation, location, scale, and different positional attributes of the object. Hence, the process of object detection and segmentation becomes very difficult and the outputs are invariant to small changes in the inputs. This is a problem when detailed information must be preserved throughout the network. CNNs accumulate sets of features at each subsequent layer, starting with finding edges, shapes, and finally objects. However, little information about the spatial relationships between these features such as size and orientation, is lost. Therefore, the result in ambiguous conclusions is due to samples of the same images with different orientations being considered as different images. One way to overcome the problem is with excessive training for all of the possible angles, which ultimately takes a lot
more time and computational efficiency. This issue is addressed by building complex architectures around CNNs to recover some of the lost information. While modern CNN architecture has managed to reconstruct the positional information using a combination of advanced techniques, it still requires more enhancements to achieve 100% efficiency. Another remarkable shortcoming of the pooling layer is if the position of the object is slightly changed, the activation does not seem to change with its proportion. This leads to good accuracy in terms of image classification, but results in poor performance.

IV. CAPSNET IN CLASSIFICATION

A CapsNet is a collection of capsules. A capsule is a group of neurons representing an entity present in an image. A capsule provides two values corresponding to an entity it represents. The first is the probability that the entity represented by a capsule exists in the image. The second is the initial parameters for that entity.

The probability value of the capsule conveys the existence of the object in the image. These initial parameters represent certain characteristics of the entity represented by the capsule, which includes position, size, hue, position, saturation, etc. Additionally, the initial parameters of the capsule include its geometric properties and color statistics (color, length, breadth, and texture). As mentioned, a capsule is a collection of neurons. Each neuron can be considered a property of the entity and in this case the capsule encompasses 4 neurons for each property. This can be represented as a vector with parameters as x1, x2, x3, and x4. The probability of existence is the length of this vector. The probability of the presence of the entity can be derived from these parameters as

$$p = \sum_{i=1}^{n} \sqrt{x_i^2}$$

But this probability value must lie within the range [0,1]. The capsule output must be transformed to fit into this range. This non-linear transformation of the output is called the squashing function. This function

serves as an activation function for the capsule networks. The capsule network, like other neural networks, is structured with different layers. There are two basic layers in the capsule network known as Primary layer and Secondary layer. The capsules in the primary layer or the lowest layer are termed as primary capsules. The capsules in the secondary layer or higher layers are called the routing capsules. The primary capsules represent small regions of an image, biologically known as a receptive field. These capsules try to locate the existence of an entity or an object with a probability based on different geometric parameters. The routing capsules in the higher layers recognize more complex objects which are composed of many smaller entities. Routing by agreement is one of the important algorithms of CapsNet by which the lower layers and the higher layers work together to recognize the presence of an object in the image. This routing between the capsule leads to a coupling effect. When the capsules in the lower layer acknowledge the higher layer about the presence of a high-level entity, the capsules in the higher layer will send feedback to the lower layer capsules. The feedback helps improve the output manipulated by the lower layer capsules. The primary capsule layer in a CapsNet is usually a combination of regular convolutional layers which helps in extracting the primary features of the images through feature maps. The feature maps are reshaped accordingly and with the assistance of the squashing function, the output of the primary capsule layer will be a vector of length between 0 and 1.

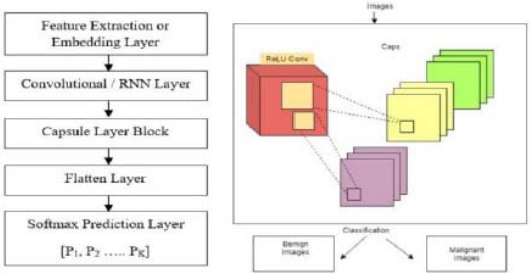


Fig. 1. - A Basic CapsNet Architecture and Proposed Process

The proposed convolution layer uses a traditional convolution network with the ReLU activation function to extract the basic features of the image. The input layer is of size 553*736*3. The kernel size used is 5x5, the stride is 2, and the kernel is 256.

The subsequent layers of the CapsNet are the ones that help in detecting the entities and their pose through a routing algorithm. The CapsNet architecture, as shown in Figure 1, is similar to a general neural network architecture except for the introduction of the capsule layer block consisting of a sequence of capsule layers. This layer is used to divide features into capsules. This part contains a convolution layer with a kernel size of 55 and a stride size of 2. The output size is 6x6x256(6x6x32x8), and the dimensionality of each capsule is 8.

The CapsNet architecture overcomes one of the important drawbacks of the Convolutional Neural Network. A CNN's architecture is composed of a sequence of convolutional, pooling, and fully connected layers. The Max Pooling layer is a type of routing mechanism. Only the active feature from that layer is passed to the further layers, but the higher layers do not provide any feedback on the routing from the pooling layer. This is introduced in the CapsNet with a dynamic routing mechanism. Additionally, the CapsNet has the capability to recognize things with few iterations, whereas the existing networks like CNN, need training on huge image sets. This clearly indicates the significance of the CapsNet in the image recognition field.

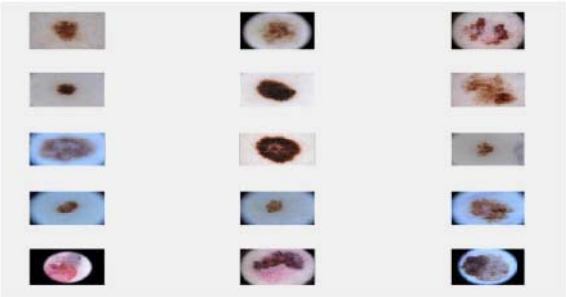


Fig. 2. - Sample BioGPS and HAM 10000 - Skin Cancer Dataset

The final layer in the proposed method is a fully connected layer, usually using the softmax activation function for classification. Here, the probabilities returned by the softmax activation function for each input are assigned to one of the mutually exclusive classes. Automatic classification of skin cancer based on the object skin lesion images helps to improve the dermoscopic analysis and examination. Since there can be a fine-grained difference in the appearance of skin lesions, CapsNet can be utilized to achieve a performance competent enough to the dermatologists.

V. RESULTS AND DISCUSSION

The proposed approach is validated using BioGPS and HAM 10000 – Skin cancer dataset. Figure 2 shows the sample dataset. Transfer learning is used to train the Capsule network used in this research to identify the skin cancer cell from the dermatological photo. Transfer learning is an efficient method used for training the CNN. In fact, the datasets used for experimental validation are small and training from them is not effective. To address this, the first few layers of the network are used to train a linear classifier on the output from these initial layers. The method works by reusing the already learned features of a benchmarked neural network. Here we used DNN which has more than 1000 categories with a thousand images in each category. The experimentation is carried out in a Windows 10 machine with NVIDIA Titan Xp GPU and Intel Xeon - Octo- core processor, 128 GB RAM. Figure 3

shows the accuracy rate for both training and testing. Figure 4 shows the loss rate for both training and testing.

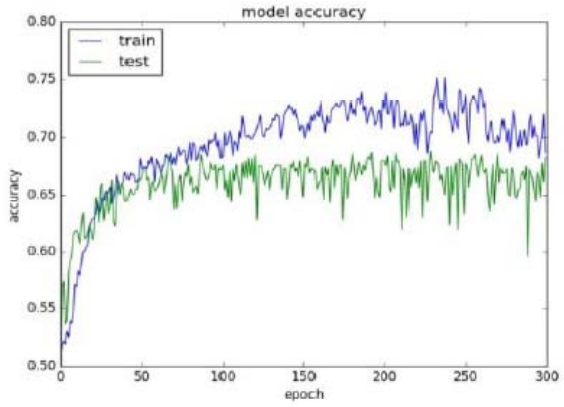


Fig. 3. – Accuracy vs. Epoch

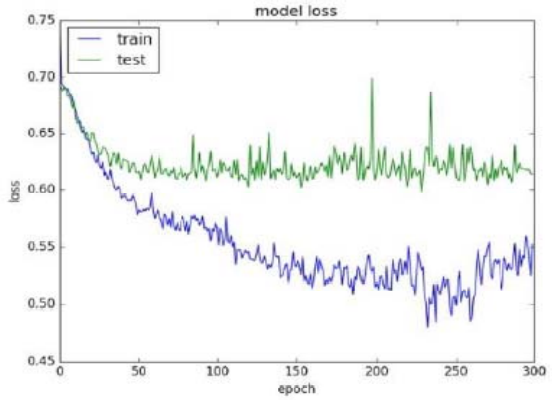


Fig. 4. – Loss vs. Epoch

CONCLUSION

In this paper, an approach to detect skin cancer that is robust enough to segment the cancer region is proposed. The training images are equally weighted so that the distribution among the samples are best approximated and further used for classification in the CapsNet model. From the experimental results, it is concluded that the proposed approach is better in terms of segmentation and classification than the state of the art classifiers available for skin cancer detection. In the experiments, various features were compared and determined the proposed approach is more sensitive than the state of the art classifiers.

ACKNOWLEDGEMENT

NH-INBRE New Hampshire INBRE, (IDeA Network of Biomedical Research Excellence) which is a program funded by the National Institutes of Health.

REFERENCES

[1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA CancerJ Clin. 2018; doi: 10.3322/caac.21442.

[2] Rogers, H. W. et al. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012. JAMA Dermatology 151.10, 1081–1086.

[3] Esteva, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115-118, 2017, doi:10.1038/nature21056.

[4] H. Alquran et al., "The melanoma skin cancer detection and clas- sification using support vector machine," 2017 IEEE Jordan Confer- ence on Applied Electrical Engineering and Computing Technologies (AEECT), Aqaba, 2017, pp. 1-5.

[5] R. S. Soumya, S. Neethu, T. S. Niju, A. Renjini and R. P. Aneesh, "Advanced earlier melanoma detection algorithm using colour correl- ogram," 2016 International Conference on Communication Systems and Networks (ComNet), Thiruvananthapuram, 2016, pp. 190-194.

[6] Neha and A. Kaur, "Wearable antenna for skin cancer detection," 2016 2nd International Conference on Next Generation Computing Tech- nologies (NGCT), Dehradun, 2016, pp. 197-201.

★★★