Machine Learning on Cataracts Classification Using SqueezeNet

Xingzhi Qian
Dulwich International High School,
132 Zhenyuan, Xinbuhuating,
Changshu Suzhou, Zip Code 215500,
Jiangsu Province, P.R. China
felixqianxz@gmail.com

* Evan W. Patton
Department of Computer Science and Engineering, CSAIL
Massachussetts Instittue of Technology
Cambridge, MA 02139, U.S.A.
ewpatton@csail.mit.edu

*Justin Swaney
Department of Chemical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139, U.S.A.
jswaney@mit.edu

*Qian Xing
Department of Ophthalmology
Changshu 2nd People's Hospital,
Changshu, Zip Code 215500
Jiangsu Province, P.R. China
*corresponding authors

Tingying(Helen) Zeng
Division of Career Education
Academy for Advanced Research and Development (AARD)
Cambridge, MA 02142, U.S.A.
helen.zeng@us-aard.com

***Abstract*—Cataracts is a serious eye disease, affecting over 20 million people worldwide. It is the clouding of the lens, which blocks the light to go through the lens and project on the retina [1]. As a result, the nerve cannot transfer the whole image to the brain, leading to blindness. A vast majority of cataracts patients are people who are over 50 years old. To classify different areas of cataracts in lens, we use supervised training of convolutional neural network to train 420 images of cataracts on the lens taken from slit-lamps. The experiment can make the future of classifying cataracts more easily and ophthalmologists can apply operations to different categories of cataracts within a shorter time to cure patients with cataracts. For those people in the countryside, even not so experienced doctors can take the photo of lens and use the program to classify cataracts correctly.**

Keywords—cataracts classification, convolutional neural network, SqueezeNet, Lens Opacities Classification System, version II (LOCS II)

I. INTRODUCTION

Classifying cataracts is time-consuming for ophthalmologists, and it requires experts to classify correctly. Usually it is hard for old people in the countryside to receive the treatment on cataracts as soon as possible because of lack of well-experienced doctors and poor equipment, resulting in blindness of most patients. Therefore, using computer programs to classify the cataracts automatically from the images taken through slit-lamps can help solve the problem. The Lens Opacities Classification System, version II (LOCS II), uses a set of colored slit-lamp to classify and grade

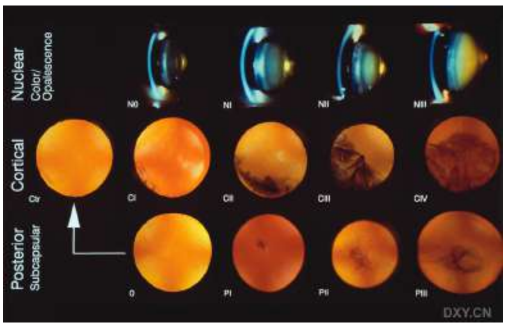

Fig. 1

different levels of nuclear, cortical, and posterior subcapsular cataract [2]. It is used specifically to classify age-related cataracts, not congenital cataracts or traumatic cataracts. Figure 1 shows the classification of LOCS II. The notation of a cataract after classifying with LOCS II is like N*C*P* (where * is different numbers according to the degree of cataracts in three different areas, e.g. N0C2P1, N3C0P0). Larger numbers correspond to more severe cataracts.

A nuclear cataract starts with a gradual hardening and yellowing of the central zone of the lens, also known as the nucleus. [3] The color of lens will turn yellow because of the nuclear cataract. The darker the yellow color is, the more serious the cataract will be. Figure 2(a) shows an example of a nuclear cataract.

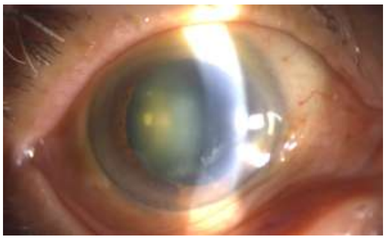

Fig. 2(a)

Cortical cataracts are associated with the local disruption of the structure of mature fiber cells. [4] The spoke shaped white opacities appears around the edge of lens. When the proportional area of the white opacities increase, cortical cataract becomes more severe. Figure 2(b) is the image of

Fig. 2(b)

cortical cataract.

Posterior subcapsular cataracts form beneath the lens capsule, which is a small "sac," or membrane, that encloses the lens and holds it in place [5]. There will be yellow clouding on the posterior subcapsular of lens. As the area of the clouding increases, the posterior subcapsular becomes worse. Figure 2(c) provides an example of a posterior subcapsular cataract.

Since the time limitation and lack of data for each individual levels of cataracts, we only did the classification of the three types of cataracts which are nuclear cataracts,

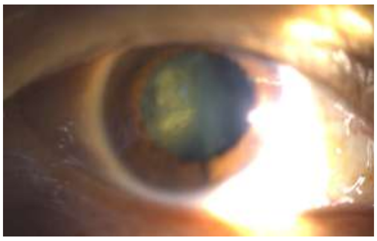

Fig. 2(c)

cortical cataracts and posterior subcapsular cataracts instead of getting specific results like notations of LOCS II.

II. Materials and Method

We generated a classifying neural network model using the following procedure: First we got our dataset with labels for training and did some image preprocessing such as cropping and image generating methods. We then trained new neural network layer on top of SqueezeNet, which is a pre-trained model that already achieved high accuracy on classifying images in ImageNet.[11] Finally, we tested the model using some images to get the output which is the classification of images in four categories.

A. Datasets of cataracts

The images of lens were collected from patients in both genders who had age-related cataracts in No.2 Hospital, Changshu, Jiangsu, China. There are 420 images in total in each of the four categories which are Nuclear Cataract, Cortical Cataract, Posterior Subcapsular Cataract and Transparent Lens. All of the images are in size of 1200*1920*3 in TIFF format. Therefore, it is feasible to use supervised learning in training data. Eighty percent of the images (336 images) are separated into training set, and twenty percent of the images (84 images) are separated into testing set.

B. Image Preprocessing

In order to classify the cataracts accurately, we only need to extract the lens from the images and get rid of other areas outside lens. There is one algorithm called HoughCircle Detection which is used to detect the circle part of one image [6]. Since there are spotlights on the lens when taking the photos from slit-lamps, which are also circles, it is hard to detect the lens. After trying HoughCircle Detection, some images will be cropped into spotlights part instead of whole lens part. Also, the lens have different sizes and are in different positions in images, so it is not possible to crop the images with specific coordinates. Therefore, cropping out the images manually was used and it did not cost a lot of time since the input dataset is limited in size.

After cropping all of the images, they had different sizes as lens in each individual image is in different size compared with other lenses. As a result, all the images were resized into 224*224*3 which made the training more easily.

C. Balancing Training Data

The numbers of images in Cortical Cataracts, Nuclear Cataracts and Transparent Lens are much greater than the number of images in Posterior Subcapsular. This leads to the imbalanced training data. When doing the machine learning, the imbalanced distribution of different classes exists [7]. Therefore, it is important to increase the number of images in the minority class which is the Posterior Subcapsular Cataracts. By duplicating some images in Posterior Subcapsular Cataracts and removing some images in classes with larger number of images, all four categories reached the same number of images finally that every category has 100 images [8].

D. Increasing Data

Since there is a limited number of images in four categories, the model may suffer from overfitting. To mitigate this concern, we used Keras' ImageDataGenerator to augment the input dataset. ImageDataGenerator is implemented in order to solve the problem of small dataset. It generates images with real-time data augmentation, which means it runs at the same time when the computer is training the model. It can rotate, flip, or shift one image in order to get more images.

E. Training Model

At first, we wanted to build our own CNN (Convolutional Neural Network). CNN is made up of neurons with learnable weights and biases [9]. However, it required a long time to change the parameters, layers and activation function in order to get better results. The accuracy after we trained the model was about fifty percent, which was not so well-trained.

To save the time on training and changing the model, we used transfer learning method. Transfer learning is a machine

layer name/type	output size	filter size / stride (if not a fire layer)	depth	s_{1x1} (#1x1 squeeze)	e_{1x1} (#1x1 expand)	e_{3x3} (#3x3 expand)
input image	224x224x3					
conv1	111x111x96	7x7/2 (x96)	1			
maxpool1	55x55x96	3x3/2	0			
fire2	55x55x128		2	16	64	64
fire3	55x55x128		2	16	64	64
fire4	55x55x256		2	32	128	128
maxpool4	27x27x256	3x3/2	0			
fire5	27x27x256		2	32	128	128
fire6	27x27x384		2	48	192	192
fire7	27x27x384		2	48	192	192
fire8	27x27x512		2	64	256	256
maxpool8	13x12x512	3x3/2	0			
fire9	13x13x512		2	64	256	256
conv10	13x13x1000	1x1/1 (x1000)	1			
avgpool10	1x1x1000	13x13/1	0			

activations (input/output data between layers)

parameters

Fig. 3

learning method where a model developed for a task is reused as the starting point for a model on a second task [10].

SqueezeNet is one type of convolutional neural network(CNN), which achieves the same accuracy on classifying images from ImageNet as AlexNet but needs 50 times fewer parameters than AlexNet. It can be compressed to 510 times smaller than AlexNet. This provided a quicker and more accurate result of the training model.

What is different in SqueezeNet from other CNN model is that there is one building block called fire module. As figure 3 shows, it contains one expand layer and one squeeze layer. There are 8 fire modules in total and 24 dimensional hyperparameters [11]. Figure 3 shows all layers in one SqueezeNet model.

III. RESULTS

In order to see the result of the training model, we used TensorBoard to visualize the graph.

We used Softmax as activation function in final output which is a probability distribution over K different results(in this case is 4 outcomes)[12]. The outcome with greater possibility is the final output of the classification.

Before cropping the images to extract the lens part, the accuracy was about 30 percent because the areas outside the lens affect the training a lot. After cropping the images, the accuracy increases in a greater degree.

Figure 4(a) is the validation graph and Figure 4(b) shows the validation loss graph after running twenty epochs of training model. The validation accuracy reaches 96.1% and validation loss reduces to 0.193, which means in 84 testing images, around 80 images are classified correctly.

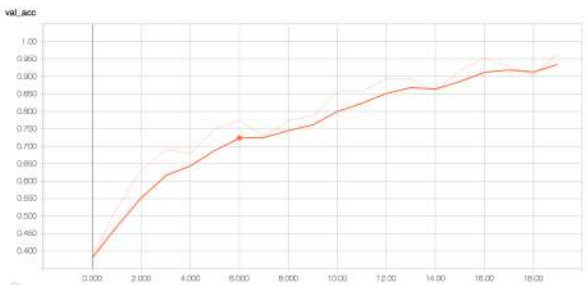

Fig. 4(a)

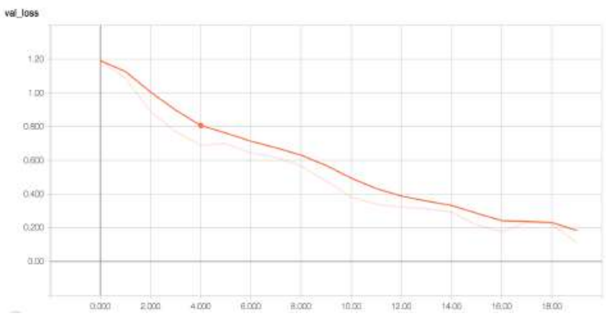

Fig. 4(b)

IV. CONCLUSION

We applied SqueezeNet model in order to classify and detect cataracts from nuclear cataracts, cortical cataracts, posterior cataracts and transparent lens, and implemented some techniques such as cropping manually and data generation using ImageDataGenerator in order to increase the level of accuracy of the model. In our training model, we successfully classified different cataracts with the accuracy around 96%. This can be further improved by training more data and more epochs. The spotlight from slit-lamps on lens of every image may affect the results since we didn’t remove them when we trained images. With more data in different levels of cataracts from three different areas of lens, it is possible to classify the cataracts the same as LOCS II and the output should be the notation as N*C*P*. Using this model, doctors can easily input images taken from slit-lamps and they will get the correct classification of cataracts, which helps to increase the accuracy and efficiency of classification. In the future, we expect that doctors in large hospitals will not need to spend a lot of time driving to countryside to diagnose every person’s eye, and that this will be scalable with commodity hardware.

V. ACKNOWLEDGMENT

We are so grateful for the sponsorship of Academy for Advanced Research and Development (AARD) (www.us-aard.com) for this internship research support.

REFERENCES

[1] “Facts About Cataract.” National Eye Institute, U.S. Department of Health and Human Services, 1 Sept. 2015, www.nei.nih.gov/health/cataract/cataract_facts.

[2] Jr, Leo T. Chylack. “Lens Opacities Classification System II (LOCS II).” JAMA Internal Medicine, American Medical Association, 1 July 1989, jamanetwork.com/journals/jamaophthalmology/article-abstract/637914.

[3] “Are There Different Types of Cataracts?” Continue Painting with Vision Loss - VisionAware, www.visionaware.org/info/your-eye-condition/cataracts/different-types-of-cataracts/125.

[4] “Cortical Cataracts.” American Academy of Ophthalmology, www.aao.org/bcscsnippetdetail.aspx?id=43e7b453-fe11-4eb9-bb0b-a59ba93ba944.

[5] “Are There Different Types of Cataracts?” Continue Painting with Vision Loss - VisionAware, www.visionaware.org/info/your-eye-condition/cataracts/different-types-of-cataracts/125.

[6] “An Improved Algorithm of Hough Circle Detection--《Applied Science and Technology》2006.06.” The Advantages and Disadvantages of DIV CSS Distribution Mode Applied in Making Static Web Pages--《Journal of Shanxi College for Youth Administrators》2011.04, en.cnki.com.cn/Article_en/CJFDTOTAL-YYKJ200606023.htm.

[7] Glander, Shirin. “Dealing with Unbalanced Data in Machine Learning.” Sitewide ATOM, shiring.github.io/machine_learning/2017/04/02/unbalanced.

[8] Dellana, Ryan, and Kaushik Roy. “Performance of Feature-Based Texture Classifiers versus Convolutional Neural Network under Sparse Training Data.” SoutheastCon 2017, 2017, doi:10.1109/secon.2017.7925325.

[9] CS231n Convolutional Neural Networks for Visual Recognition, cs231n.github.io/convolutional-networks/.

[10] “A Gentle Introduction to Transfer Learning for Deep Learning.” Machine Learning Mastery, 21 Dec. 2017, machinelearningmastery.com/transfer-learning-for-deep-learning/.

[11] Gao, Hao. “Notes on SqueezeNet – Hao Gao – Medium.” Medium, Augmenting Humanity, 19 Oct. 2017, medium.com/@smallfishbigsea/notes-of-squeezenet-4137d51feef4.

[12] “Comp.ai.neural-Nets FAQ, Part 2 of 7: LearningSection - What Is a Softmax Activation Function?” Scientific Child Rearing - Encyclopedia of Children and Childhood in History and Society, www.faqs.org/faqs/ai-faq/neural-nets/part2/section-12.htm