Full Paper Acceptance Notification

Vancouver, Canada, Aug. 23-25, 2019
www.iscai.org

Dear Snehil Kumar,

Congratulations! On behalf of the ISCAI 2019 committee, we are pleased to inform you that your paper entitled "Diabetic retinopathy diagnosis with ensemble deep learning" (ID: CA022) has been accepted for oral presentation and publication at the 2019 International Symposium on Computing and Artificial Intelligence (ISCAI 2019). We would like to formally invite you to attend ISCAI 2019 that is going to be held in Vancouver, Canada, during Aug. 23-25, 2019.

You have 15 minutes time slot to present your work during the conference in English. And we'd like to advice you to prepare the presentation file in both PDF and PPT format in advance.

Accepted papers will be published in the digital conference proceedings which will send to be indexed by all major citation databases such as Ei Compendex, Scopus, Google Scholar, Cambridge Scientific Abstracts (CSA), Inspec, SCImago Journal & Country Rank (SJR), EBSCO, CrossRef, Thomson Reuters (WoS), etc.
A selection of papers will be recommended to be published in journals.

Since the presentation and publication of your paper are subject to your successful registration, you are encouraged to finish your registration before Aug 15, 2019. Kindly notice that you can enjoy the regular registration fee if you could finish your registration before the above deadline. To finish your registration process, please refer to the 2nd page of this notification.

Congratulations again and we look forward to meeting you.

Yours sincerely,

ISCAI 2019

Organizing Committee Aug 13, 2019

Diabetic Retinopathy Diagnosis With Ensemble Deep-Learning

Snehil Kumar
Jadavpur University
Kolkata, WB, India
snehil03july@gmail.com

ABSTRACT

Diabetic retinopathy is an eye disease, is a condition in which retina is damaged due to diabetes mellitus. It is a major cause of blindness. Although many artificial intelligence methods has been applied to diabetic retinopathy diagnosis. This method is a new approach in this problem domain. It's early detection can help in tackling the future damages due to the disease. Here, model is ensemble of pre-existing GoogleNet, AlexNet and ResNet50 architectures. The best machine learning models that were used was GoogLeNet achieving highest accuracy for this job. Here, The results are standing out with the GoogLeNet's accuracy.

CCS Concepts

Artificial Intelligence • Computer Vision • AI in Healthcare

Keywords

Ensemble learning; Deep learning; Transfer Learning; Computer vision; AI in healthcare.

1. INTRODUCTION

1.1 Diabetic Retinopathy

Diabetes can even lead to blindness but that can be prevented, diagnosed before the severe conditions comes. The blindness due to diabetes is mostly seen in the people of the age group 30years to 69 years. Approximately, all the patients of the type 1 diabetes and nearly 60% of the people with type 2 diabetes, do have some degree of retinopathy after 20 years from the day of the onset of diabetes. About 25% of cases of the type 2 diabetes even sometimes at the time of diagnosis have a background of this disease. In the current scenario, treatment of this disease is possible and should be treated before the problem in vision occurs. Diabetes retinopathy (DR) is a major reason for the vision loss in the people of varied age groups from middle-aged to older-aged people.

This, disease can be prevented through early detection as well as proper treatment and can lead to the prevention of diabetic related visual impairments. Patients who are suffering diabetes need to have a periodic follow-up with primary care physicians to reduce their glycaemic, blood pressure and lipid control to reduce the risk of development and progression of DR and other diabetes-related problems. Diabetic retinopathy (DR) can also take place because of higher body mass index, puberty and pregnancy, and cataract surgery [8]. To minimize the effects of the visual loss due to this problem, it is vital that all expertise look forward for innovative methods of overseeing and avoiding diabetes, and ideate cost-worth camouflaging schemes in the community. Here, in this experiment the method used is a new approach in this problem domain. It's early detection can help in tackling the future damages due to the disease. Here, model is ensemble of pre-existing GoogleNet, AlexNet and ResNet50. Here, Figure. 1 shows the image of retina obtained from device.

ISCSIC '18, September 21–23, 2018, Stockholm, Sweden

ACM ISBN 978-1-4503-6628-1/18/09…$15.00
https://doi.org/10.1145/3284557.3284727

Figure 1. Images of retina obtained from various devices.

1.2 Ensemble-Learning

Ensemble learning [20] can be termed as the method in which multiple methods such as (classifiers or experts) are generated strategically and combined to solve a particular computational intelligence problem. Ensemble learning is basically used to improve the performance of a model such as (classification, prediction, function approximation, etc.), or nullify the similarity of an selection of a poor one. There are much more applications of ensemble learning for example assigning a confidence to the decision made by the model, selecting optimal (or near optimal) features, data fusion, incremental learning, non-stationary learning and error-correcting. Here, in this experiment we have ensemble of pre-trained models for computer-vision that is GoogLeNet, AlexNet and ResNet50. The combination was arranged in different arrangements in the ensembles and the study of accuracies are made. Ensemble learning has made the remarkable accuracies which are standing near to highest accuracies attained.

1.3 Transfer-Learning

The classical machine learning algorithms make predictions on the new or the future data using mathematical and statistical models which are trained on the experience or the data which are either labeled or unlabeled. Transfer learning [15], basically, welcomes the domains, tasks, and distributions which are utilized in training and testing to be different. There are many examples of transfer learning present in the current scenario. For an instance, the model used to recognize a tiger may, help in recognizing the image of an cheetah and other animals of the similar animal family. In the same manner learning to play guitar may, help in learning violin also. The intuition of Transfer learning came with the fact that people can easily and wisely apply knowledge learned from the past experience to solve a new problems faster or with better solutions. And the same mathematical and statistical model can be developed for computers to act the same way. Here we are using pre-trained computer vision models GoogLeNet, AlexNet and ResNet50 utilizing transfer learning in this way.

1.4 GoogleNet

GoogLeNet [4], is basically, a 22 layered deep learning convolutional neural network model basically designed and trained for object detection. GoogLeNet [4] was previously already trained with the processed data. And the accuracies for 2-arry [5], 3-arry [5] and 4-arry [5] classification were 0.7275, 0.6425 and 0.5525 respectively.

1.5 AlexNet

AlexNet [2] came much later after the advent of various CNNs [11] (e.g. Yann LeCun's LeNet paper in 1998). It basically consists of 60 million parameters and 650,000 neurons. Today there are we have various more complex CNNs [11] and that can run efficiently on GPUs. This was also trained on the processed data and the results are shown below.

1.6 ResNet

Resnet [9] basically stands for Deep residual networks. It is basically a a 50 layer deep convolutional neural network which is trained over more than a million images. This folder contains an implementation of ResNet for the ImageNet dataset written in TensorFlow. This was also trained using the processed data.

2. The DATASET

The dataset contains images of 35,126 high-resolution colour fundus retinal images which are labelled to five classes .

0 - No DR
1 – Mild
2 – Moderate
3 – Severe
4 - Proliferative DR

which correspond to the five stages of the diabetic retinopathy. The data were collected from the famous kaggle.com. These all 5 classes were numbered from 0 t0 4 with different extent of the effect of the disease from No DR to Severe and Proliferative DR. The images in the dataset were extracted from varied devices, which can have different affects on the visual appearance of left and right retinas. Others are shown as one would see through a microscope condensing lens (i.e. inverted, as one sees in a typical live eye exam). There is also noise in both the images and labels. Images may contain antiques, be out of focus, underexposed [6], or overexposed [17] and are of diverse resolutions. The data is public.

3. DATA PRE-PROCESSING

Here, a previously used image processing method is used. At first, all the images that were collected were converted to a stratified data configuration for the further steps of data pre-preprocessing, data augmentation, and afterwards model-training. There were multiple steps in pre-processing: The images were sliced using Otsu's method [13] to remove the round-colored images of the retina. Images were tabulated by removing the minimum pixel intensity [14] from each channel and dividing by the mean pixel intensity [15] to represent pixels in the range 0 to 14. The contrast reconcilation was done with the contrast limited adaptive histogram equalization (CLAHE) [16] filtering algorithm.

4. EXPERIMENT- SETUP

4.1 Hardware-Used

Here, there was access to the powerful hardware on Intel AI dev cloud. Information is in Table.1.

Table 1. The infrastructure for the training.

Components	Details
Architecture	x86_64
CPU op-modes	32bit, 64bit
CPUs	24
Model Name	Intel® Xeon® Gold 6128 processor @ 3.40 GHz
RAM	92GB
Model	85
CPU Family	Six

4.2 Software-Used

Here, three great softwares that were used were tensorflow [12], opencv [3] and python was used in the implementation.

4.3 Image-processing

Here, the image dataset consists of images of different range of patients with heavily diverse levels of lighting in the fundus photography [1]. The pixel intensity values of the images is effected by the lighting and creates a redundant contrast unconnected to classification levels. A contrast limited adaptive histogram equalization filtering algorithm (CLATHE) , with the OpenCV python module was used to label this antique. The results of this pre-processing step is shown in Figure. 2. Here, digital image preprocessing technique helped improving detection of fleck ultra-fine features and convolutional neural networks helped in the further pre-processing and made the data ready for training. We anticipate this transformation is traceable to the channel wise contrast inflating effect of histogram equalization [19].

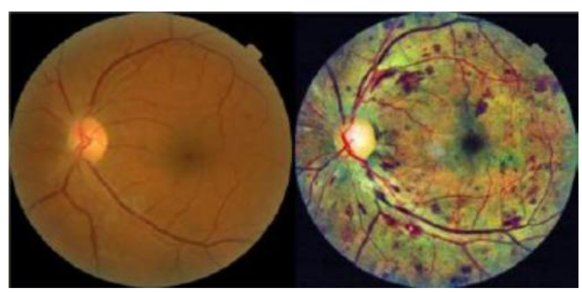
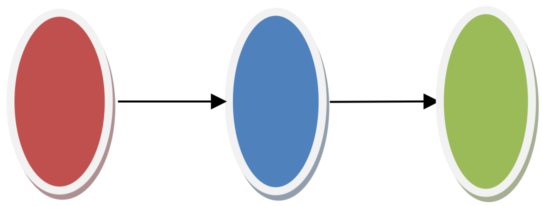


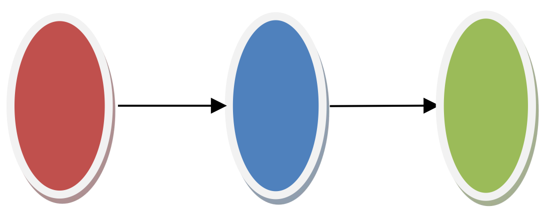
Figure 2. Two images before and after CLATHE application. The left one is before and the right image is after the application of clathe algorithm.

4.4 Models for Training

Here, basically 3 types of ensembles were made combining 3 existing computer-vision models that are GoogLeNet, AlexNet and ResNet50. Here, models are named as follows with the different orientation of the 3 existing models. Model-1 (GoogLeNet->AlexNet->Resnet), Model-2 (GooLeNet->ResNet->AlexNet), Model-3 (AlexNet->GoogLeNet->Resnet50), Model-4(AlexNet->ResNet50->GoogLeNet),Model-5(ResNet50->GoogLeNet->AlexNet), Model-6 (ResNet50->AlexNet->GoogLeNet).



Model 1. Alignment of different models in the first ensemble model (GoogLeNet->AlexNet->Resnet).



Model 2. Alignment of different models in the last ensemble model(AlexNet->GoogLeNet->Resnet50).

In this way, all the 6 ensembles configurations are made by interchanging the position of different models obtained through transfer learning. These all 6 models were trained and results were interesting and different for all the models. Here, all the models were trained with stochastic gradient descent [10] optimization algorithms with 30 epochs [7] and a learning rate [18] of 0.001.

5. RESULTS

Here, firstly the accuracies for the processed data was calculated with only AlexNet and Resnet. And the results were like this in Table. 2.

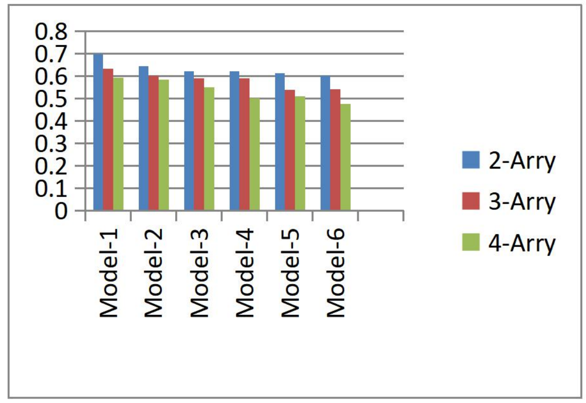
Table 2. Accuracies of different classifier.

	AlexNet	ResNet
2-Arry[5]	0.601	0.557
3-Arry[5]	0.572	0.519
4-Arry[5]	0.502	0.438

After, this the results were calculated for the different ensemble models and their accuracies are shown in Table. 2. And the Graph-1 also depicts the accuracy plot of different ensemble models. The results are showing that the accuracies attained by these are achieving quite similar heights as that was attained by the individual ones. The highest accuracy attained is by Model-1 with 2-arry classification that is 0.6994 and the lowest is 0.4746 which is attained by Model-6 with 4-Arry classification.

Table 3. This table depicts the different accuracy values.

	2-Arry	3- Arry	4-Arry
Model-1	0.6994	0.6345	0.593
Model-2	0.6452	0.6012	0.5834
Model-3	0.6620	0.5971	0.5489
Model-4	0.6213	0.5897	0.5027
Model-5	0.6142	0.5388	0.5109
Model-6	0.6018	0.5430	0.4746



Graph 1. Accuracies of 2-arry, 3-arry and 4-arry classifiers.

6. CONCLUSION

Here, It can be concluded that these models are touching the accuracies that were attained by the single model. In few cases they have overtaken the individual model and a few more modification can lead these methods to new heights. All the six models show the same trend with the 2-arry having highest accuracy and then 3-arry and 4-arry consecutively decrease respectively.

One Important conclusion that can be drawn is the occurrence of model with highest accuracy at the first position and decreasing rest 2 accuracy at 2nd and 3rd position respectively leads to a model with highest accuracy. So, we can say that *"In an ensemble if we put a model with high accuracy at previous positions in a model that ensemble will lead to a higher accuracy model".*

7. FUTURE-SCOPE

This is a quite emerging research field. Here, with a bit of some modifications in optimization methods and epochs and other aspect some remarkable heights can be accomplished. This particular experiment still has a lot of potential to reach the highest accuracy with a bit of modifications in the methods which are discussed above.

8. ACKNOWLEDGEMENTS

I am thankful to Intel for providing the infrastructure for conducting the research. There was a support of Intel AI Devcloud where the research was conducted. Here, these models were a new deep learning approach to this problem case and this will have applications in other similar problem cases also.

REFERENCES

[1] Abràmoff, Michael D.; Garvin, Mona K.; Sonka, Milan (2010-01-01). Retinal Imaging and Image Analysis. IEEE Transactions on Medical Imaging. 3: 169–208.

doi:10.1109/RBME.2010.2084567. ISSN 0278-0062. PMC 3131209. PMID 21743764

[2] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. *In Proceedings of the 25th International Conference on Neural Information Processing Systems ,*Volume 1 (NIPS'12), F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Eds.), Vol. 1. Curran Associates Inc., USA, 1097-1105.

[3] Bradski G. The OpenCV Library. Dr Dobb&#39, Journal of Software Tools, 2000

[4] C. Szegedy et al., Going deeper with convolutions, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 1-9.doi: 10.1109/CVPR.2015.7298594

[5] Carson Lam, MD, Darvin Yi, Margaret Guo, and Tony Lindsey, PhD. Automated Detection of Diabetic Retinopathy using Deep Learning.

[6] Chen-Jui Chung, Wei-Yao Chou and Chia-Wen Lin, Under-exposed image enhancement using exposure compensation," 2013 13th International Conference on ITS Telecommunications (ITST), Tampere, 2013, pp. 204-209. doi: 10.1109/ITST.2013.6685546

[7] Devarakonda, Aditya, Maxim Naumov and Michael Garland. AdaBatch: Adaptive Batch Sizes for Training Deep Neural Networks. ArXiv abs/1712.02029 (2017): n. pag.

[8] Frampton G, Harris P, Cooper K, et al. The clinical effectiveness and cost-effectiveness of second-eye cataract surgery: a systematic review and economic evaluation. Southampton (UK): NIHR Journals Library, 2014 Nov. (Health Technology Assessment, No. 18.68.) References. Available from: https://www.ncbi.nlm.nih.gov/books/NBK263913/

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep Residual Learning for Image Recognition, CoRR, abs/1512.03385, 2015, http://arxiv.org/abs/1512.03385, arXiv, 1512.03385.

[10] Kiefer, J.; Wolfowitz, J. *Stochastic Estimation of the Maximum of a Regression Function*. Ann. Math. Statist. 23 (1952), no. 3, 462--466. doi:10.1214/aoms/1177729392. https://projecteuclid.org/euclid.aoms/1177729392

[11] LeCun, Y., Bengio, Y. and Hinton, G., 2015. *Deep learning. Nature,* 521(7553), pp.436-444

[12] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Rafal Jozefowicz, Yangqing Jia, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Mike Schuster, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: *Large-scale machine learning on heterogeneous systems*, 2015. Software available from tensorflow.org.

[13] N. Otsu, "*A Threshold Selection Method from Gray-Level Histograms*," in IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62-66, Jan. 1979. doi: 10.1109/TSMC.1979.431007

[14] PANG, Z. , LAPLANTE, N. and FILKINS, R. J. (2012), Dark pixel intensity determination and its applications in normalizing different exposure time and autofluorescence removal. Journal of Microscopy, 246: 1-10. doi:10.1111/j.1365-2818.2011.03581.x

[15] Pratt, L. Y. (1993). "*Discriminability-based Transfer between Neural Networks" (PDF*). NIPS Conference: Advances in Neural Information Processing Systems 5. Morgan Kaufmann Publishers,pages 204–211.

[16] Reza, A.M. The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology (2004) 38: 35. https://doi.org/10.1023/B:VLSI.0000028532.53893.82

[17] S. Arora, M. Hanmandlu, G. Gupta and L. Singh, "*Enhancement of Overexposed Color Images*," 2015 3rd International Conference on Information and Communication Technology (ICoICT), Nusa Dua, 2015, pages 207-211. doi: 10.1109/ICoICT.2015.7231423

[18] Senior, Andrew W., Georg Heigold, Marc'Aurelio Ranzato and Ke Yang. "*An empirical study of learning rates in deep neural networks for speech recognition*." 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (2013): 6724-6728.

[19] Stephen M. Pizer, E. Philip Amburn, John D. Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart Ter Haar Romeny, and John B. Zimmerman. 1987. Adaptive histogram equalization and its variations. Comput. Vision Graph. Image Process. 39, 3 (September 1987), 355-368. DOI=10.1016/S0734-189X(87)80186-X http://dx.doi.org/10.1016/S0734-189X(87)80186-X

[20] Thomas G. Dietterich. 2000. Ensemble Methods in Machine Learning. *In Proceedings of the First International Workshop on Multiple Classifier Systems (MCS '00)*, Josef Kittler and Fabio Roli (Eds.). Springer-Verlag, London, UK, UK, 1-15.