ROP Auto Detection with Deep Neural Network

First Author
Institution1
Institution1 address
firstauthor@i1.org

Second Author
Institution2
First line of instituation2 address
http://www.author.org/~second

Abstract

ROP(Retinopathy of Prematurity) is a blinding disease, which primarily occurs on premature infants whose birth weights about 1250 grams or gestation less than 31 weeks. Premature infants in advance separate theirselves from the maternal environment to receive artificial oxygen, the development of the blood vessels are very sensitive to high oxygen concentrations, resulting in retinal vascular hyperplasia, even resulting in ROP. Nowadays, in many developing countries, it would take much time and energy to train an ophthalmologist, which means making ROP diagnosis on each premature infant is not realistic. In order to overcome this dilemma, we have developed an automation system to analyse premature infant retinal photographs using deep nerual network, judging the existence and the severity of ROP. The initial aim of our system is to help these premature infants who can not get the diagnosis of professsional ophtalmologist, make them enjoy equal treatment like those born in big cities.

1. Introduction

Retinopathy of prematurity (ROP) is a potentially blinding eye disorder that primarily affects premature infants weighing about 2 pounds (1250 grams) or less that are born before 31 weeks of gestation (A full-term pregnancy has a gestation of 38–42 weeks). The smaller a baby is at birth, the more likely that baby is to develop ROP. Nowadays, ROP has become a major cause of blindness in children in the developing and developed world despite current surgical treatment in the late-stage of the disease[?, ?]. ROP is a proliferative retinal vascular disease that occurs when abnormal blood vessels grow and spread throughout the retina. If a baby is born prematurely, before these blood vessels have reached the edges of the retina, normal vessel growth may stop. The edges of the retina may not get enough oxygen and nutrients. Scientists believe that the periphery of the retina then sends out signals to other areas of the retina for nourishment. As a result, new abnormal vessels begin to grow. These new blood vessels are fragile

and weak and can bleed, leading to retinal scarring. When these scars shrink, they pull on the retina, causing it to detach from the back of the eye. The available data suggest that blindness due ROP varies enormously from country to country, and that over 50,000 children are blind from ROP worldwide[?, ?, ?, ?]. Eighty percent of infants with birth weight less than 1500 g born in the United Kingdom survive and the incidence of stage 3 ROP of approximately 8% to 10% has been reported[?, ?, ?, ?]. Thus all babies having gestational age 31 wk or 1500 g are screened in United Kingdom[18]. American guidelines given by the American Academy of Pediatrics state that, infants with a birth weight 1500 g or gestational age of 30 wk and selected infants with a birth weight between 1500 and 2000 g or gestational age of more than 30 wk with an unstable clinical course, should be screened for ROP[?]. In China, National Ministry of Health promulgated the "guide for the prevention and treatment of premature infants with oxygen and retinopathy", birth weight of less than 2000 g, gestational age less than 37 wk of preterm infants should undergo fundus examination for ROP.

2. Related Work

Many auto or semi-auto dignose methods have been proposed to diagnose ROP plus disease. According to Aslam *et al*. [], it takes three phrases to dignose,(1) image segmentation.With image processing algorithms to construct vascular tree; (2) measurement of vessel diameter. It is a challenge that describe vessel's thickness and location exactly, cause that the process are subjective and reliable; (3) measurement of tortuosity. with varied algorithms to evaluate tortuosity of blood vessel result in different results.

Nowadays, cause of the excellent results Convolution Neural network(CNN) has achieved, more and more researchers start to focus on how to apply CNN on medical image process field, ROP images is no exception. Worrall *et al*. proposed a novel CNN architecture to diagnose ROP plus disease(waiting to complete). Brown *et al*. [] use two CNN to diagnose ROP plus disease, the first CNN is vessel segmentation network, which is used to segment retinal vessel through output a probability map whose size is same

as input image and the value is between 0 to 1. The second CNN is classification network whose architecture adopts Inception version1 [] architecture.

Above of all, Every method mentioned above is to to diagnose the existence of ROP plus disease, or calssify normal, pre-plus, plus disease. In this article, we develop a novel neural network, through enhance the ability of feature map represent feature on every layer in CNN, which is described on the next section. with this network to diagnose the existence and severity of ROP gain state-of-the-art result.

3. Data and Methodology

3.1. Data

Our data comes from Sichuang Provincial Peoples Hospital and Chengdu Women & Childern's Central Hospital, which contains xxx ROP examinations from 2014 to 2017. Every examination consists of 4 to 12 retinal photographs, which reflects each premature infant's fundus situation. which annotated by two professional ophtahalmologists.

First of all, we develop online label system to help ophtahalmologists annotate data, ophtahalmologist label on each examination, Figure 3.1 displays the detail of inspecting fundus photograph on label system, ophtahalmologist enable adjust brightness, saturation and contractness of retinal fundus photograph, and also we provide the draw line tools to aid ophtahalmologist caculate the distance between optic disc and "ridge/valley traversal", and the length of "ridge/valley traversal". Both of them assist ophtahalmologist diagnose the existence of ROP and severity of ROP certainly.

Second, to ensure the consistence of data labels and prevent from false annotate result from annotator's carelessness, such as annotator's mishandle and lack of attention. we ask for three ophtahalmologists to annotate every ROP examination. one ophtahalmologist has more than 10 years clinical experience on ROP, and the other ophtahalmologists have about 5 years clinical experience relatively. We adopt examinations to construct dataset that the examinations have consistent label ammong three ophtahalmologists. Based on the theory [?], retinal vessel tortuosity and the existence of "ridge/valley traversal" reflect the existence and severity of ROP. In order to construct a model mapping fundus image to ROP judgement, we need utilize the model to extract fundus image features. To better extract these features, we construct dataset with per-image rather than per-examination. (waiting to complete)

Last, we divide all images into three sets, training sets, validation sets and testing sets respectively. Table 3.1 displays data distribution on every sets.

3.2. Vessel Segmentation

Ophtahalmologist generally regard tortuosity of premature infants fundus image as one of most important standard to diagnose ROP, in fact, premature infants' retinal vessel haven't develop mature, when they separate themselves from maternal enviorment. And the infants unable to breath their own need artificial oxygen, which high oxygen concentrations causes tortuosity of retinal vessel, become ROP finally.

We try to utilize u-net [?] to segment vessel of fundus images. U-net is a symmetrical image segmentation framework, its left side consists of many modules serially, which every module has same construction, two 3×3 convolution operation used as feature extraction, one 2×2 max pooling operation used as downsampling. so does the right side, except for deconvolution operation is adopted to replace convolution operation. U-net uses skip connection fuse low level and high level features, which contributes to looking for dense and hierarchical image features.

However, Annotating these fundus images to train segmentation model costs most resources. Consider that premature infants fundus image has part of similiarity with adult fundus image,"transfer learning" could assist we pretrain model, and then we fine tune parameters θ on segmentation model to access better segmentation performace.

3.3. Classification Network

Since 2014, Google proposed Inception network architecture continuely [?, ?, ?, ?], which imporved the best published result on ImageNet again and again. The successful secret of Inception network is a module named "Inception module" is proposed. Different from traniditional convolutional network, Inception network used a "Inception module" as a layer. And the inventor of Inception network extend the width of "Inception module" with use different sizes of kernel to extract different spatial features, which superior than traniditional hierarchical convolutional network in performace.

Median Frequency Balancing: Consider that dataset is imbalanced, the number of normal samples is much more than ROP samples. We use median frequency balancing on loss function to deal with such problem. According to median frequency balancing, α_c denotes coordinate of class c while training, *e.g.* $totalloss = \sum_1^n \alpha_c \cdot loss(c)$, which is formulated as: $\alpha_c = \frac{median freq}{freq(c)}$

$freq(c)$ denotes the number of class c divided set number, and $median_f req$ is the median of all frequencies of classes. For binary classification tasks, we can formulate median frequency balancing as: $\alpha_p = \frac{n+p}{2 \cdot p} \alpha_n = \frac{n+p}{2 \cdot n}$

α_p denotes frequency of positive samples, α_n denotes frequency of negative samples, and p, n denotes the number of positive and negative samples.

3.4. Transfer Learning

The primary aim of "Transfer Learning" is proposed to save manual labeling costs with transferring model parameters from labeled dataset to unlabeled dataset. However, considering the difficulties to collect large scale data and cost to label data, more researches gradually to focus on transferring model parameters on two domains, one domain possesses labeled large scale avilable data, but the other don't.

Let's give definitions of task and domain firstly. Domain contains feature space χ and Edge probability distribution $p(X)$ where $X = x_1, x_2, ..., x_n \in \chi$, and task contains label space γ and target prediction function $f(\cdot)$, which is conditional probability distribution $p(Y|X)$. Accordingly, we have the definition of "Transfer Learning". Given a source domain D_s, a target domain D_t, a source task T_s and a target task T_t, transfer learning applies the knowledge which has been learned from D_s and T_s to help learn prediction function $f_t(\cdot)$ on domain D_t.

On the basis of the difference of domain and task, transfer learning can be partition as inductive transfer learning, transductive transfer learning and unsupervised transfer learning. In inductive transfer learning, we believe $T_s! = T_t$, and not care relationship between D_s and D_t. It utilizes labeled large scale data on domain D_s to train a objective model, and then transfer the parameters to D_t for improving performace on T_t; In transductive tranfer learning, T_t is same as T_s, but D_t is not same as D_s, mainly considering the difference of feature space or edge probability distribution; In unsupervised tranfer learning, it focus on unsupervised learning, *e.g.* clustering, dimensionality reduction, density estimation. ImageNet, cifar-10, CelebA, ... all these are good materials for transfer learning.

3.5. Draft and final copy

The LaTeX style defines a printed ruler which should be present in the version submitted for review. The ruler is provided in order that reviewers may comment on particular lines in the paper without circumlocution. The camera ready copy should not contain a ruler. (LaTeX users may uncomment the `\cvmfinalcopy` command in the document preamble.)

3.6. Miscellaneous

Compare the following:

| `$conf_a$` | $conf_a$ |
| `$\mathit{conf}_a$` | $conf_a$ |

See The TeXbook, p165.

The space after *e.g.*, meaning "for example", should not be a sentence-ending space. So *e.g.* is correct, *e.g.* is not. The provided `\eg` macro takes care of this.

When citing a multi-author paper, you may save space by using "et alia", shortened to "*et al.*" (not "*et. al.*" as
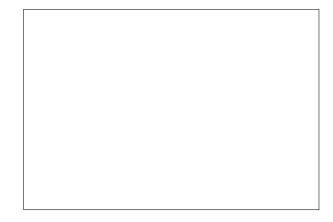

Figure 1. Example of caption.

Name	Performance
A	OK
B	Bad
Ours	Great

Table 1. An example for using tables.

"*et*" is a complete word.) However, use it only when there are three or more authors. Thus, the following is correct: " Frobnication has been trendy lately. It was introduced by Alpher [1], and subsequently developed by Alpher and Fotheringham-Smythe [2], and Alpher *et al.* [3]."

This is incorrect: "... subsequently developed by Alpher *et al.* [2] ..." because reference [2] has just two authors. If you use the `\etal` macro provided, then you need not worry about double periods when used at the end of a sentence as in Alpher *et al.*.

For this citation style, keep multiple citations in numerical (not chronological) order, so prefer [2, 1, 4] to [1, 2, 4].

3.7. References

List and number all bibliographical references in 9-point Times, single-spaced, at the end of your paper. When referenced in the text, enclose the citation number in square brackets, for example [4]. Where appropriate, include the name(s) of editors of referenced books.

3.8. Illustrations, graphs, and photographs

All graphics should be centered. Please ensure that any point you wish to make is resolvable in a printed copy of the paper. Resize fonts in figures to match the font in the body text, and choose line widths which render effectively in print. Many readers (and reviewers), even of an electronic copy, will choose to print your paper in order to read it. You cannot insist that they do otherwise, and therefore must not assume that they can zoom in to see tiny details on a graphic.

When placing figures in LaTeX, it's almost always best to use `\includegraphics`, and to specify the figure width as a multiple of the line width as in the example below

```
\usepackage[dvips]{graphicx} ...
\includegraphics[width=0.8\linewidth]
                {myfile.eps}
```

References

[1] A. Alpher. Frobnication. *Journal of Foo*, 12(1):234–778, 2002. 3

[2] A. Alpher and J. P. N. Fotheringham-Smythe. Frobnication revisited. *Journal of Foo*, 13(1):234–778, 2003. 3

[3] A. Alpher, J. P. N. Fotheringham-Smythe, and G. Gamow. Can a machine frobnicate? *Journal of Foo*, 14(1):234–778, 2004. 3

[4] Authors. The frobnicatable foo filter, 2012. Face and Gesture submission ID 324. Supplied as additional material `fg324.pdf`. 3