DEEP MULTIPLE INSTANCE LEARNING FOR AUTOMATIC DETECTION OF DIABETIC RETINOPATHY IN RETINAL IMAGES

Eshwar p.k[1]
Computer Science and Engineering
SRM Institute of Science and Technology
Chennai, India

Rohit Shaw[2]
Computer Science and Engineering
SRM Institute of Science and Technology
Chennai, India

Harish Rudru[3]
Computer Science and Engineering SRM
Institute of Science and Techno logy
Chennai, India

Mrs T. Malathi[4]
Computer Science and Engineering
SRM Institute of Science and Technology
Chennai, India

Abstract— In this paper, we propose an efficient approach for deep multiple instance learning for automatic detection of diabetic retinopathy in retinal images.Glaucoma is a disorder in which severe damage to the optic nerve leads to vision loss. It identification involves the measurement of shape and size of optic cup. Due to the interweavement of optic cup with blood vessels the optic cup segmentation is bit tedious task. Pre-processing followed by segmentation is used for optic cup segmentation which is further processed to find it's dimension. Based on the fact that the fractal dimension is used to find the dimension of irregular objects, a novel approach is proposed for glaucoma detection using perimeter method of fractal analysis. The glaucoma is detect by digital image processing. a weakly supervised learning technique, multiple instance learning (MIL) has shown an advantage over supervised learning methods for automatic detection of diabetic retinopathy (DR): only the image-level annotation is needed to achieve both detection of Diabetic Retinopathy images and Diabetic Retinopathy lesions, making more graded and de-identified retinal images available for learning.

Keywords—Image classification, learning (artificial intelligence), medical image processing

1. INTRODUCTION

Diabetes now has become one of the main challenges to our human health. According to the estimation of International Diabetes Federation, the number of diabetic patients will rise up to 592 million by 2035. As a common complication of diabetes, diabetic retinopathy (DR) can cause severe vision loss and even blindness in this population. It is not only a personal catastrophe to the individual but also a threat to the nation's economy. Over the last two decades, numerous methods have been proposed for automatic DR detection. Specifically, it treats retinal images as bags and their inside patches as instances. Then following the standard assumption that a DR image contains at least one DR patch and that a normal image only contains normal patches, various models have been explored to find out those discriminative DR patches. So both classifiers for DR patches and DR images can be learned using only image-level labels. However, so far these models still use handcrafted features to characterize image patches. They fail to benefit from the larger datasets brought by SVM, often resulting in inferior detection performance compared to previous supervised learning methods. Compared with the whole image classification achieved by CNNs, our method provides explicit locations of DR lesions so that the detected retinal images can be easily checked by clinicians. The main pipeline of our method is described as follows. First of all, image pre-processing is applied to all retinal images, normalizing factors such as image scale, illumination and Diabetes is a chronic disease caused by the increase in blood sugar, mainly either due to the less production or no production of insulin in body or due to the fact that cells do not respond to the produced insulin. In recent years, the number of diabetic patients has increased drastically. Moreover, diabetes is the major cause for heart stroke, kidney failure, lower-limb amputations and blindness. The development of the computer-based methods that would enable the high probability recognition of pre-diabetic or diabetic condition can be an efficient support to the decision making in healthcare.

2. LITERATUREREVIEW

1.Manoj Kumar, Anubha Sharma and Sonali Agarwal, **"Clinical decision support system for diabetes disease diagnosis using optimized neural network" ,** Institute of Electrical and Electronic Engineers 2014.

The research paper is organized as follows- the section two of the paper is based on related work in health care using data mining. In section three the fundamentals of health data mining is discussed. Section four elaborates feature selection method used to select the most appropriate feature form the dataset to perform effective data mining. Section five highlights basic of ant colony optimization and neural network. A hybrid model based on ACO and Neural network with feature selection has been proposed in section six. Section seven and eight includes performance analysis of proposed model with meaningful concluding remarks. Accuracy of an algorithm with and without feature selection is found with ACO (Ant Colony Optimization) neural network, this is found to be the findings of the paper. The cons in this paper is that the theoretical analysis is difficult and sequences of random decisions (not independent).

2. Vaishali R, Dr R.sasikala,S. Ramasubbareddy, S.Remya and SravaniNalluri,**"Genetic algorithm based feature selection and MOE fuzzy classification algorithm on Pima Indians Diabetes dataset"**, Institute of Electrical and Electronic Engineers 2017.

Multi Objective Evolutionary Fuzzy Classifier works on the principle of maximum classifier rate and minimum rules. As a result of feature selection with GA, the number of features is reduced to 4 from 8 and the classifier rate is improved to 83.0435 %, this is found to be the findings of the paper .Due to feature selection, the size of the dataset gets reduced and thus rules based classifiers may end up with the problem of model overfitting is found to be the cons. These two algorithms are combined in order to achieve good accuracy in lesser time. The interpretability of the machine learning model on critical medical problems is improved. Hence, we have chosen MOE fuzzy classifier.

3.Ayush Anand and DivyaShakti,**" Prediction of Diabetes based on personal lifestyle indicators"**, 2015 1st International Conference on Next Generation Computing Technologies(NGCT-2015),Dehradun,India, 4-5 September 2015.

This research paper is a discussion on establishing a relationship between diabetes risk likely to be developed from a person's daily lifestyle activities such as his/her eating habits, sleeping habits, physical activity along with other indicators like BMI (Body Mass Index), waist circumference etc. Chi-Squared Test of independence was performed followed by application of the CART(Classification And Regression Trees) machine learning algorithm on the data and finally using Cross-Validation, the bias in the results was removed ,this is found to be the findings of the paper. Results of CART came to be a little biased because of a smaller dataset and it took only one or two factors into account, this is found to be the cons. CART prediction model has been applied with 75% accuracy, for a highly categorical dataset. Blood Pressure has been identified as a significant factor in causing diabetes, along with the others such as roadside eating, late sleeping, diabetic family history or heredity, intake of rice and the level of physical activities performed

3. PROPOSED METHOD

In this project Glaucoma is a silent thief of sight. A detail of eye vision used for image processing uses preprocessing, feature extraction, feature selection, CNN techniques and data sets used for testing and training purpose. Glaucoma identification involves the measurement of shape and size of optic cup. Due to the interweavement of optic cup with blood vessels the optic cup segmentation is bit tedious task. Pre-processing followed by segmentation is used for optic cup segmentation which is further processed to find it's dimension. Based on the fact that the fractal dimension is used to find the dimension of irregular objects, a novel approach is proposed for glaucoma detection using perimeter method of fractal analysis. The glaucoma is detected by means of CNN algorithms. The disease is detected, then the detected information sends to the microcontroller. If disease detected then the data will be send to microcontroller and then the controller will update patient information in webpage by using IOT. If disease detected microcontroller will display the data in LCD and produce alert sound by using buzzer.

4. MODULES

- Preprocessing (Median filter)
- Edge detection(canny)
- Segmentation(otsu's)
- Feature extraction(GLCM)
- Classification(CNN)

Preprocessing (Median filter)

The Median Filter is a non-linear digital filtering technique, often used to remove noise from an image or signal. Such noise reduction is a typical pre-processing step to improve the results of later processing (for example, edge detection on an image). Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise (but see discussion below), also having applications in signal processing. The main idea of the median filter is to run through the signal entry by entry, replacing each entry with the median of neighboring entries. The pattern of neighbors is called the "window", which slides, entry by entry, over the entire signal. For 1D signals, the most obvious window is just the first few preceding and following entries, whereas for 2D (or higher-dimensional) signals such as images, more complex window patterns are possible such as "box" or "cross". Typically, by far the majority of the computational effort and time is spent on calculating the median of each window. Because the filter must process every entry in the signal, for large signals such as images, the efficiency of this median calculation is a critical factor in determining how fast the algorithm can run.

The naïve implementation described above sorts every entry in the window to find the median; however, since only the middle value in a list of numbers is required, selection algorithms can be much more efficient. Furthermore, some types of signals (very often the case for images) use whole number representations: in these cases, histogram medians can be far more efficient because it is simple to update the histogram from window to window, and finding the median of a histogram is not particularly onerous.

Edge detection (canny)

Median filtering is one kind of smoothing technique, as is linear Gaussian filtering. All smoothing techniques are effective at removing noise in smooth patches or smooth regions of a signal, but adversely affect edges. Often though, at the same time as reducing the noise in a signal, it is important to preserve the edges. Edges are of critical importance to the visual appearance of images, for example. For small to moderate levels of Gaussian noise, the median filter is demonstrably better than Gaussian blur at removing noise whilst preserving edges for a given, fixed window size. However, its performance is not that much better than Gaussian blur for high levels of noise, whereas, for speckle noise and salt-and-pepper noise (impulsive noise), it is particularly effective. Because of this, median filtering is very widely used in digital image processing.

Non-linear filters have many applications, especially in the removal of certain types of noise that are not additive. For example, the median filter is widely used to remove spike noise — that affects only a small percentage of the samples, possibly by very large amounts. Indeed, all radio receivers use non-linear filters to convert kilo- to gigahertz signals to the audio frequency range; and all digital signal processing depends on non-linear filters (analog-to-digital converters) to transform analog signals to binary numbers.

Segmentation (OTSU)

In computer vision and image processing, Otsu's method, named after Nobuyuki Otsu, is used to automatically perform clustering-based image thresholding, or, the reduction of a gray level image to a binary image. The algorithm assumes that the image contains two classes of pixels following bi-modal histogram (foreground pixels and background pixels), it then calculates the optimum threshold separating the two classes so that their combined spread (intra-class variance) is minimal, or equivalently (because the sum of pairwise squared distances is constant), so that their inter-class variance is maximal. Consequently, Otsu's method is roughly a one-dimensional, discrete analog of Fisher's Discriminant Analysis. Otsu's method is also directly related to the Jenks optimization method.

Otsu's method exhibits the relatively good performance if the histogram can be assumed to have bimodal distribution and assumed to possess a deep and sharp valley between two peaks. But if the object area is small compared with the background area, the histogram no longer exhibits bimodality. And if the variances of the object and the background intensities are large compared to the mean difference, or the image is severely corrupted by additive noise, the sharp valley of the gray level histogram is degraded. Then the possibly incorrect threshold determined by Otsu's method results in the segmentation error. (Here we define the object size to be the ratio of the object area to the entire image area and the mean difference to be the difference of the average intensities of the object and the background)

From the experimental results, the performance of global thresholding techniques including Otsu's method is shown to be limited by the small object size, the small mean difference, the large variances of the object and the background intensities, the large amount of noise added, and so on.

Feature extraction (GLCM)

A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also known as the gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. (The texture filter functions, described in Texture Analysis cannot provide information about shape, that is, the spatial relationships of pixels in an image.) A co-occurrence matrix or co-occurrence distribution is a matrix that is defined over an image to be the distribution of co-occurring pixel values (grayscale values, or colors) at a given offset.

Whether considering the intensity or grayscale values of the image or various dimensions of color, the co-occurrence matrix can measure the texture of the image. Because co-occurrence matrices are typically large and sparse, various metrics of the matrix are often taken to get a more useful set of features. Features generated using this technique are usually called Haralick features, after Robert Haralick.

Texture analysis is often concerned with detecting aspects of an image that are rotationally invariant. To approximate this, the co-occurrence matrices corresponding to the same relation, but rotated at various regular angles (e.g. 0, 45, 90, and 135 degrees), are often calculated and summed.

Texture measures like the co-occurrence matrix, wavelet transforms, and model fitting have found application in medical image analysis in particular.

The texture filter functions provide a statistical view of texture based on the image histogram. These functions can provide useful information about the texture of an image but cannot provide information about shape, i.e., the spatial relationships of pixels in an image.

Another statistical method that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also known as the gray-level spatial dependence matrix. The toolbox provides functions to create a GLCM and derive statistical measurements from it.

Classification (CNN)

In machine learning, a convolutional neural network (CNN, or ConvNet) is a class of deep, feed-forward artificial neural networks that has successfully been applied to analyzing visual imagery.

CNNs use a variation of multilayer perceptrons designed to require minimal preprocessing. They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on their shared-weights architecture and translation in variance characteristics.

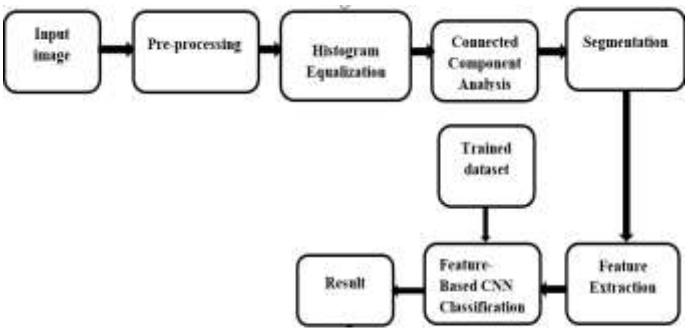
Convolutional networks were inspired by biological processes in that the connectivity pattern between neurons resembles the organization of the animal visual cortex. Individual cortical neurons respond to stimuli only in a restricted region of the visual field known as the receptive field. The receptive fields of different neurons partially overlap such that they cover the entire visual field.

CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were hand-engineered. This independence from prior knowledge and human effort in feature design is a major advantage.

A CNN consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers and normalization layers

Description of the process as a convolution in neural networks is by convention. Mathematically it is a cross-correlation rather than a convolution. This only has significance for the indices in the matrix, and thus which weights are placed at which index.

5. ARCHITECTURE DIAGRAM



6. CONCLUSION

In this project, we propose a CNN method for DR detection by taking the complementary advantages from MIL and deep learning: only the image-level annotation is needed to achieve both detection of DR images and DR lesions, meanwhile, features and classifiers are jointly learned from data. The pre-trained Alex Net is adapted and deeply fine-tuned in our framework to achieve the patch-level DR estimation. An end-to-end multi-scale framework is applied to help better handle the irregular DR lesions. Compared to existing methods for DR detection, our method significantly improves the detection performance. In the future work, we are going to incorporate techniques such as semi-supervised learning and active learning into our method to further achieve the accurate segmentation of DR lesions.

7. ACKNOWLEDGEMENT

We would like to express our gratitude to our guide Mrs. T.Malathi, who guided us throughout this project and offered a deep insight into the study. We wish to acknowledge the help provided by the technical and support staff in computer science department of SRM Institute of Science and Technology.

8. REFERENCES

[1] N. Cho, J. Shaw, S. Karuranga, Y. Huang, J. da Rocha Fernandes, A. Ohlrogge et al., “Idf diabetes atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045,” Diabetes research and clinical practice, vol. 138, pp. 271–281, 2018.

[2] S. M. S. Islam, M. M. Hasan, and S. Abdullah, “Deep learning based early detection and grading of diabetic retinopathy using retinal fundus images,” International Conference on Machine Learning, Image Processing, Network Security and Data Sciences, 2018.

[3] K. Zhou, Z. Gu, W. Liu, W. Luo, J. Cheng, S. Gao et al., “Multi-cell multi-tas convolutional neural networks for diabetic retinopathy grading,” in International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2018, pp. 2724–2727.

[4] J. Krause, V. Gulshan, E. Rahimy, P. Karth, K. Widner, G. S. Corrado et al., “Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy,” Ophthalmology, vol. 125, no. 8, pp. 1264– 1272, 2018.

[5] F. Ren, P. Cao, D. Zhao, and C. Wan, “Diabetic macular edema grading in retinal images using vector quantization and semisupervised learning,” Technology and Health Care, no. Preprint, pp. 1–9, 2018.

[6] Q. Chen, Y. Peng, T. Keenan, S. Dharssi, E. Agro, W. T. Wong et al., “A multi-task deep learning model for the classification of age-related macular degeneration,” AMIA Summits on Translational Science Proceedings, vol. 2019, p. 505, 2019.