ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2019, i6doc.com publ., ISBN 978-287-587-065-0. Available from http://www.i6doc.com/en/.

Blind-Spot Network for Image Anomaly Detection: A New Approach to Diabetic Retinopathy Screening

Shaon Sutradhar[1,2], José Rouco[1,2], Marcos Ortega[1,2] *

1- CITIC - Research Center of Information and Communication Technology, University of A Coruña, Spain

2- Department of Computer Science, University of A Coruña, Spain

Abstract. The development of computer-aided screening (CAS) systems is motivated by the high prevalence and severity of the target disease along with the time taken to manually assess each case. This is the case with diabetic retinopathy screening, that is based on the manual grading of retinography images. The development of CAS systems, however, usually involves data-driven approaches that require extensive and usually scarce manually labeled datasets. With this in mind, we propose the use of unsupervised anomaly detection methods for screening that can take advantage of the large amount of healthy cases available. Concretely, we focus on reconstruction-based anomaly detection methods, which are usually approached with autoencoders. We propose a new network architecture, the Blind-Spot Network, that, according to the presented experiments, improves the performance of autoencoders in this setting.

1 Introduction

The retina is composed of specialized neurons and cells that are connected to the visual cortex making it an extension of the brain and the unique organ in human body allowing the analysis of the central nervous system with non-invasive methods [1]. The retina contains metabolically active tissues irrigated with blood vessels. Thus several chronic diseases, such as diabetes, stroke, hypertension, along with neurodegenerative and cardiovascular diseases [1], leave significant biomarkers on the retina. This allows these diseases to be prognosed using the retinal information, and potentially in far advance to their development.

Diabetes has a high risk of developing into diabetic eye diseases like diabetic retinopathy (DR) or diabetic macular edema (DME). These are the most common causes of blindness among the working-age people in the developed countries [2]. A systematic screening program can, however, significantly minimize the risk of vision loss through the early detection of DR and DME symptoms, which would allow an early intervention to prevent visual impairment [3].

In a typical DR screening program, highly skilled experts manually analyze the retinal fundus images to grade the progression of the disease. Such manual

*This research is supported by the Instituto de Salud Carlos III, Government of Spain and the ERDF of the EU through the DTS15/00153 research project.

ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2019, i6doc.com publ., ISBN 978-287-587-065-0. Available from http://www.i6doc.com/en/.

grading is a subjective and time-consuming process [3]. Also, with the increasing prevalence of DR, this manual approach may not meet the demand.

Computer-aided screening (CAS) systems can automate the retinal image assessment to provide a fast and repeatable process. The development of such systems for DR screening has been a developing research area in the last three decades [4]. Many of the existing approaches employ machine learning algorithms for the automatic grading of DR. These data-driven procedures usually require manually annotated retinal image dataset for training [4]. However, there is scarcity of large datasets with annotated lesions, because manual labeling is a time consuming and tedious task that requires the intervention of expert clinicians. This motivates the development of unsupervised learning techniques that can take advantage of the existent unlabeled data. Furthermore, some rare pathologies may lack sufficient data to represent their whole variability in a representative way, even using the unlabeled data. This further motivates the use of unsupervised anomaly detection approaches that allow the recognition of any pattern that diverges from the healthy cases.

Anomaly detection methods have been applied to video surveillance, cybersecurity, disease recognition in medical imaging and among others [5, 6]. In image data, it usually refers to the detection of abnormal local regions or objects. Machine learning techniques, with both supervised and unsupervised methods, are extensively applied in image anomaly detection [5]. Unsupervised anomaly detection methods typically involve learning the appearance of normal structures in order to identify the anomalous data which do not fit to the learned model. This approach is practical when a large dataset of normal cases is available, while there is insufficient representation of abnormal cases. Some frequently used unsupervised anomaly detection methods include one-class support vector machine (OC-SVM) [7], autoencoders (AEs) [8] and their variants [9] or generative adversarial networks (GAN) [10].

In the work herein described we employ the concept of reconstruction-based unsupervised anomaly detection for screening. This is usually approached with AEs [11] and, to the best of our knowledge, it has not been used before for DR screening. Additionally, we propose a new network architecture, the Blind-Spot Network, as an alternative to AEs, in this setting. The provided experimental results demonstrate the advantage of the BSNet against AEs.

2 Methods

A scheme of the proposed screening system is shown in Fig 1. The system is a reconstructor-based anomaly detection method which consists of an image reconstructor model that is trained to minimize the reconstruction error of normal data. It is assumed that the model will not reconstruct the anomalous data samples as accurately as normal data and thus will produce a higher reconstruction error. Therefore, the error could be used to decide if the input sample is normal or not. This screening method should be general enough to cover a wide variety of diseases, although, in this paper, we only experiment with DR.

ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2019, i6doc.com publ., ISBN 978-287-587-065-0. Available from http://www.i6doc.com/en/.

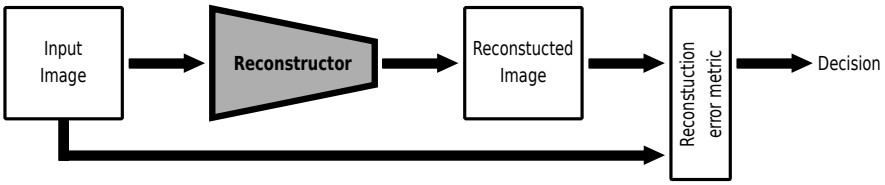


Fig. 1: Proposed image anomaly detection method for screening. The reconstructor is trained on healthy images obtaining lower reconstruction error than pathological ones.

Autoencoders are usually used for reconstruction based anomaly detection [8], and we tested them as a first approach. An autoencoder [12] transforms the high dimensional input data into a lower dimensional representation which is then transformed back to the original input dimension. Our experiments, however, evidenced that autoencoders do not efficiently learn meaningful semantic features of the input distribution and tend to add a blurring to the reconstructed output. This means that the features learned from normal images are general enough to encode and reconstruct the pathological patterns with equally low error. This motivates the proposal of alternative approaches that could add more pressure to learn higher level features. To that end, we propose a new architecture, the Blind-Spot Network (BSNet), that instead of having equal input and output, the target output is not available at the input. Thus the network is 'blind' to the target part of its input region, and forced to interpolate this 'blind spot' from its surrounding context only.

2.1 Blind-Spot Network Architecture

Given patch of 96×96 pixels, the BSNet first divides it into 9 equal 32×32 subpatches, as illustrated Fig. 2(a). The BSNet takes the 8 boundary segments as inputs and the middle segment as the target output. Each input segment is precessed separately with independent convolutional neural networks, whose outputs are concatenated and fed into a common fully-connected network that generates the target center patch. The architecture is depicted in Fig. 2(b). This network is similar to context encoders [13], however, in this case, instead of masking an area of the input image, the network architecture is designed to take the patch surroundings as input and to predict the unseen middle region.

The activation function of the output layer is linear, while the rest of the layers use ReLU. The loss function is mean absolute error (MAE). The network is initialized following the Glorot [14] method and the optimization algorithm is Adam [15] with learning rate $1e$-5. Early stopping with minimum change in validation loss of $1e$-7 and 500 epochs of patience is used.

3 Results and Discussion

In the experiments we use the public Messidor database [16]. The database consists of 1200 color fundus images with three different image sizes. In the pre-

ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2019, i6doc.com publ., ISBN 978-287-587-065-0.
Available from http://www.i6doc.com/en/.

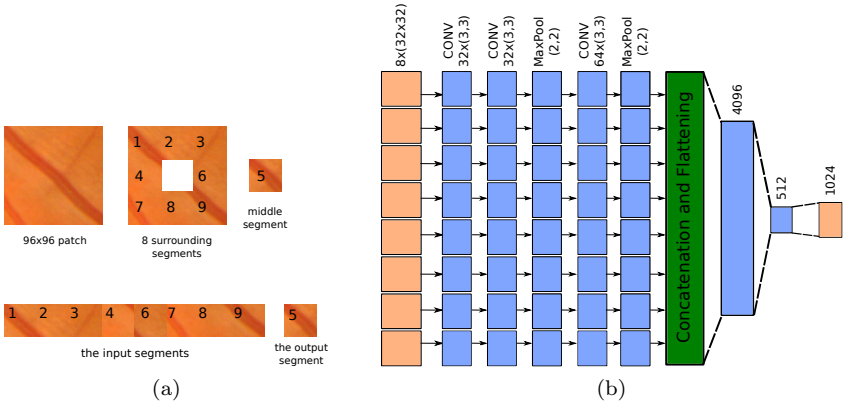


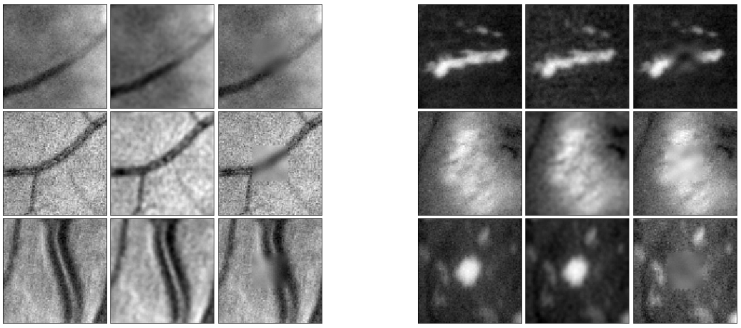
Fig. 2: The Blind-Spot Network architecture. (a) Illustration of the patch segmentation process. (b) The neural structure.

liminary experiments herein described we only use the 1488×2240 sized images. For each image, the DR grade (0 to 3) is provided. A total of 546 images are classified as normal and 654 as having different grades of DR. 27 normal images are used to prepare the training dataset. Manually selected 13900 target points, covering the macula, optic disk and blood vessel regions, are used to extract patches centered on them. We extract patches of three different sizes (96×96, 144×144 and 48×48 pixels) from each point. For augmentation, we rotate the 96×96 patches with 90°, 180° and 270° angles and scale the 144×144 and 48×48 patches with 0.5 and 1.5 factors respectively. A total of 83400 patches are generated with the above mentioned augmentation. An additional validation dataset of 3135 normal patches is generated with a similar extraction procedure from 13 randomly selected normal images. Finally, a test dataset of pathological patches is collected with additional care so that each patch contains some portion of a lesion. We randomly select 26 pathological (abnormal) images and extract 398 patches. The patches of the three sets are converted into gray-scale, and normalized between 0 and 1.

We compare the proposed BSNet model with a plain fully-connected AE. Several AE architectures have been tested achieving similar reults. The best performance was obtained with a fully-connected AE with three hidden layers of 4096, 2048, 4096 ReLU neurons, and linear output. The used initialization and training methodology is the same as with the BSNet.

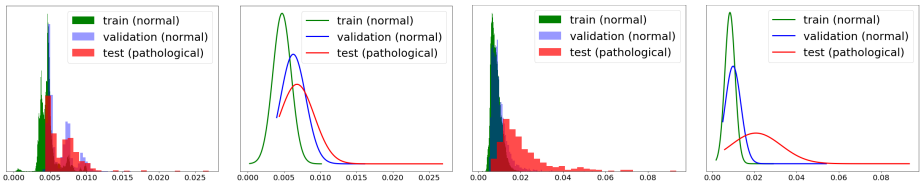
Fig. 3 depicts examples of reconstructed patches using both the BSNet and AE approaches. It can be observed that the AE model is able to similarly reconstruct both the healthy and pathological patches, while the BSNet demonstrates different performance on both classes. To quantitatively evaluate the quality of reconstructed images we use MAE. In Fig. 4 we plot the histogram of error distributions for the train, validation and test data. Additionally, we fit the three error distributions to Gaussians by computing the means and standard

ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2019, i6doc.com publ., ISBN 978-287-587-065-0.
Available from http://www.i6doc.com/en/.



(a) (b)

Fig. 3: Reconstruction examples of unseen healthy (a) and pathological (b) patches. Each line contains the input patch, and the AE and BSNet outputs.



(a) (b) (c) (d)

Fig. 4: Comparison of reconstruction error histograms and Gaussian distributions. (a,b) AE reconstructor. (c,d) BSNet reconstructor.

deviations. It is observed that, for the AE, the error distributions of the validation set, i.e. unseen healthy patches, and the test set, i.e. pathological patches, are highly overlapped. These distributions, however, are clearly separated for the BSNet. To quantitatively asses the separability of the distributions we use Mahalanobis distance, defined as:

$$M(A,B) = \sqrt{\frac{(\mu_A - \mu_B)^2}{\sigma_A^2 + \sigma_B^2}} \tag{1}$$

where A and B denote two different distribution with averages μ_A and μ_B and standard deviations of σ_A and σ_B, respectively. Table 1 shows the obtained values, confirming that the reconstruction error distributions of unseen healthy and pathological cases are more separable for the BSNet than for the AE.

4 Conclusions

In this paper we have introduced a DR screening system employing an image reconstruction based anomaly detection approach. Using this concept, we tested

ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2019, i6doc.com publ., ISBN 978-287-587-065-0. Available from http://www.i6doc.com/en/.

Table 1: Mahalanobis distance between error distributions.

Reconstructors	Test-Train	Test-Validation
AE	0.7614	0.1736
BSNet	1.0116	0.8664

AEs and compared the results with a newly proposed architecture, the Blind-Spot Network (BSNet). Our preliminary experiments show that the BSNet performs significantly better than AEs.

References

[1] T.J. MacGillivray, E. Trucco, J.R. Cameron, et al. Retinal imaging as a source of biomarkers for diagnosis, characterization and prognosis of chronic illness or long-term conditions. *The British journal of radiology*, 87(1040):20130832, 2014.

[2] World Health Organization. Global report on diabetes, 2016.

[3] J.H.K. Goh, C.Y. Cheung, S.S. Sim, et al. Retinal imaging techniques for diabetic retinopathy screening. *Journal of diabetes science and technology*, 10(2):282–294, 2016.

[4] M.F. Nørgaard and J. Grauslund. Automated screening for diabetic retinopathy–A systematic review. *Ophthalmic Research*, 60(1):9–17, 2018.

[5] V. Chandola, A. Banerjee, and V. Kumar. Anomaly detection: A survey. *ACM Comput. Surv.*, 41(3):15:1–15:58, 2009.

[6] M.A.F. Pimentel, D.A. Clifton, et al. A review of novelty detection. *Signal Processing*, 99:215–249, 2014.

[7] B. Schölkopf, R.C. Williamson, et al. Support vector method for novelty detection. In *Advances in neural information processing systems*, pages 582–588, 2000.

[8] V. Ciesielski and V.P. Ha. Texture detection using neural networks trained on examples of one class. In *Australasian Joint Conference on Artificial Intelligence*, pages 140–149. Springer, 2009.

[9] C. Baur, B. Wiestler, S. Albarqouni, et al. Deep autoencoding models for unsupervised anomaly segmentation in brain MR images. *CoRR*, abs/1804.04488, 2018.

[10] T. Schlegl, P. Seeböck, S.M. Waldstein, et al. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *International Conference on Information Processing in Medical Imaging*, pages 146–157. Springer, 2017.

[11] S. Hawkins, H. He, G. Williams, et al. Outlier detection using replicator neural networks. In *Data Warehousing and Knowledge Discovery, 4th International Conference, DaWaK 2002, Aix-en-Provence, France, September 4-6, 2002, Proceedings*, pages 170–180, 2002.

[12] G.E. Hinton. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, jul 2006.

[13] D. Pathak, P. Krahenbuhl, J. Donahue, et al. Context encoders: Feature learning by inpainting. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 2536–2544, 2016.

[14] X. Glorot and Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256, 2010.

[15] D.P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015.

[16] E. Decenciére, X. Zhang, G. Cazuguel, et al. Feedback on a publicly distributed database: the messidor database. *Image Analysis & Stereology*, 33(3):231–234, August 2014.