Delft University of Technology

Automated Lesion Detection by Regressing Intensity-Based Distance with a Neural Network

van Wijnen, Kimberlin M.H.; Dubost, Florian; Yilmaz, Pinar; Ikram, Mohammad Arfan; Niessen, Wiro; Adams, Hieab; Vernooij, Meike W.; de Bruijne, Marleen

Citation (APA)
van Wijnen, K. M. H., Dubost, F., Yilmaz, P., Ikram, M. A., Niessen, W., Adams, H., Vernooij, M. W., & de Bruijne, M. (2019). Automated Lesion Detection by Regressing Intensity-Based Distance with a Neural Network. In D. Shen (Ed.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019* (pp. 234-242). (Lecture Notes in Computer Science; Vol. 11767). Springer. https://doi.org/10.1007/978-3-030-32251-9-26

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Automated Lesion Detection by Regressing Intensity-Based Distance with a Neural Network

Kimberlin M. H. van Wijnen[1(✉)], Florian Dubost[1], Pinar Yilmaz[2],
M. Arfan Ikram[2,3], Wiro J. Niessen[1,4], Hieab Adams[2], Meike W. Vernooij[2],
and Marleen de Bruijne[1,5]

[1] Biomedical Imaging Group Rotterdam, Department of Radiology and Nuclear
Medicine, Erasmus MC, Rotterdam, The Netherlands
{k.vanwijnen,f.dubost,marleen.debruijne}@erasmusmc.nl
[2] Departments of Radiology and Nuclear Medicine and of Epidemiology,
Erasmus MC, Rotterdam, The Netherlands
[3] Department of Neurology, Erasmus MC, Rotterdam, The Netherlands
[4] Imaging Physics, Faculty of Applied Sciences, TU Delft, Delft, The Netherlands
[5] Department of Computer Science, University of Copenhagen,
Copenhagen, Denmark

Abstract. Localization of focal vascular lesions on brain MRI is an important component of research on the etiology of neurological disorders. However, manual annotation of lesions can be challenging, time-consuming and subject to observer bias. Automated detection methods often need voxel-wise annotations for training. We propose a novel approach for automated lesion detection that can be trained on scans only annotated with a dot per lesion instead of a full segmentation. From the dot annotations and their corresponding intensity images we compute various distance maps (DMs), indicating the distance to a lesion based on spatial distance, intensity distance, or both. We train a fully convolutional neural network (FCN) to predict these DMs for unseen intensity images. The local optima in the predicted DMs are expected to correspond to lesion locations. We show the potential of this approach to detect enlarged perivascular spaces in white matter on a large brain MRI dataset with an independent test set of 1000 scans. Our method matches the intra-rater performance of the expert rater that was computed on an independent set. We compare the different types of distance maps, showing that incorporating intensity information in the distance maps used to train an FCN greatly improves performance.

Keywords: Lesion detection · Geodesic distance · Fully convolutional neural network · Dot annotations · Perivascular spaces

---

1   Introduction

Obtaining the location of focal vascular lesions on brain scans, such as white matter hyperintensities, lacunes, enlarged perivascular spaces or microbleeds is extremely useful for studying the association of these lesions with neurological disorders. However the manual annotation of these lesions can be challenging, time-consuming and subject to observer bias due to the difficulty of distinguishing a specific type of lesion from other similarly appearing structures. An automated method for detecting lesions could improve reliability, generalization and speed of lesion detection, which could greatly advance neuropathology research.

Various promising automated methods have been proposed to detect lesions. Deep learning methods often provide the best accuracy, but depend on expensive manual annotations for training like voxel-wise segmentations [5,11] or bounding boxes [8] marking the lesions. This hinders applicability of these techniques in practice.

Annotating by placing a single dot per lesion instead is considerably more time-efficient, allowing to collect larger annotated datasets for training and evaluation. In this paper we therefore propose a novel method for lesion detection that requires only dot annotations. Dot annotations have been effectively used to train convolutional neural networks (CNNs) for other applications, such as cell detection in histology images [18], lacune detection in placental ultrasound [15] and landmark detection in retinal images [14]. An approach that has shown great promise is regression of a distance map (DM) that is computed from these dot annotations [14,15,18]. Contrary to many other deep learning detection methods that use a two-stage approach [8], this approach directly outputs predicted detections and is optimized in an end-to-end fashion.

We use a similar approach for detecting lesions based on dot annotations. Previous distance regression approaches for detection [14,18] have used Euclidean distance. This is especially suited for the detection of circular objects such as cells. Brain lesions on the other hand often have a morphology that is complex and discriminative [4].

In this paper we investigate the effect of including intensity information in DMs for lesion detection. Intensity distance incorporates local image context enabling the DM to capture complicated morphologies. Voxels surrounding dot annotations which have similar intensity values (inside the lesions) will have a lower value in the DM than dissimilar voxels (outside the lesions). This could encourage the CNN to learn the characteristic morphology of the lesions and propose more accurate detections than when trained on a Euclidean distance map (EDM) that does not make this distinction. We compare Euclidean distance, intensity distance, and geodesic distance that combines both Euclidean and intensity distances. For geodesic distance the image is seen as a curved surface defined by the spatial coordinates and one intensity coordinate, where the shortest path on the surface is the geodesic distance [17].

In this paper we show that including image intensity information in the DM improves optimization of a CNN for detecting lesions in brain MRI. We compute DMs from the dot annotations and their corresponding intensity images.

Subsequently we train a fully convolutional neural network (FCN) to predict these DMs for unseen intensity images. The local minimal distances in the predicted DMs correspond to the proposed detection candidates.

We show the potential of regressing intensity-based DMs for the detection of enlarged perivascular spaces (PVS). PVS burden has been associated with cerebral small vessel disease [6]. As PVS follow the course of the vessel they surround, they appear as elongated structures on 3D brain MRI scans. Several methods have been proposed to detect PVS. The majority of the proposed algorithms is however evaluated on a relatively small sets (less than 30 images) due to the need for voxel-wise annotations for testing (and training) [4,13]. We train and validate on a set of 1202 MRI scans and test on a separate set of 1000 images. As the centrum semiovale (CSO) is seen as the most difficult brain region for PVS detection and most clinically relevant, we focused on this brain region [3].

2   Method

We train an FCN to regress a DM for a given intensity image. Our approach requires MRI scans with dot annotations for training. The local optima in the predicted DMs are expected to correspond to lesion locations. We compare geodesic distance maps (GDMs), EDMs and intensity distance maps (IDMs).

2.1   Distance Transform

To compute DMs we use a distance transform, that requires a definition of the foreground – in our case the set of dot annotations Φ – and a gray-scale image $G(x)$ in the case of intensity and geodesic distances, with x the position in the image. The distance map $DM(x)$ is defined by

$$DM(x) = min(\Lambda(\gamma), \gamma \in \Psi(x, \Phi)) \tag{1}$$

with $\Psi(x, \Phi)$ the set of possible paths γ between a position x in the image and the set of dot annotations Φ. The length $\Lambda(\gamma)$ of the path γ is

$$\Lambda(\gamma) = \sum_{i=1}^{n-1} d(x_i, x_{i+1}) \tag{2}$$

with n the number of voxels in the path γ between a position x and a dot annotation $x_{dot} \in \Phi$ and d the distance measure. The geodesic distance d_G in a 2D gray-scale image between voxel x_i and the next voxel in the path x_{i+1}, with intensities $G(x_i)$ and $G(x_{i+1})$ respectively, is defined by [17] as

$$d_G(x_i, x_{i+1}) = \sqrt{d_I\big(x_i, x_{i+1}\big)^2 + d_E\big(x_i, x_{i+1}\big)^2} \tag{3}$$

with the intensity distance $d_I(x_i, x_{i+1}) = G(x_i) - G(x_{i+1})$ and the Euclidean distance $d_E(x_i, x_{i+1})$ which is 1 for $x_{i+1} \in N_4(x_i)$ (voxels connected horizontally

and vertically) and $\sqrt{2}$ for $x_{i+1} \in N_8(x_i) \setminus N_4(x_i)$ (voxels connected diagonally). EDMs are consequently computed by setting $d_I = 0$ in Eq. 3, while IDMs are computed by setting $d_E = 0$. We approximate these DMs using the optimization algorithm *iterative raster scan* described in [17]. This approach is for computing DMs in 2D, though it can easily be extended to 3D.[1]

The resulting $DM(x)$ is normalized by dividing by the maximum distance in the $DM(x)$ and inverted as this is convenient for implementation. Furthermore, we add a parameter p to influence how steeply the distance decays. The final map $M_p(x)$ is calculated using

$$M_p(x) = \left(1 - \frac{DM(x)}{max\big(DM(x)\big)}\right)^p \tag{4}$$

Fig. 1. Network architecture, on the left the input preprocessed brain scan is shown and the output predicted distance map is shown on the right

2.2 Fully Convolutional Neural Network

We use an architecture similar to a shallow U-Net for our FCN shown in Fig. 1, which was shown to work well for regressing the number of perivascular spaces in the basal ganglia [9,16]. For optimization we use mean square error loss $MSE = \frac{1}{N} \sum_x \left(\widehat{M_p}(x) - M_p(x)\right)^2$, with $\widehat{M}(x)$ the predicted map and N the number of voxels in $M_p(x)$.

Non-maximum suppression is applied to the predicted distance map to detect local optima. We use a 5×5 maximum filter with a connectivity of 8. By thresholding the local optima the proposed detections are acquired.

---

[1] Our code for computing 2D as well as 3D distance maps is available at https://github.com/kimvwijnen/geodesic_distance_transform.

3 Experiments

3.1 Data

Our data set consists of 2202 T2-weighted MRI scans from the Rotterdam Scan Study. All scans were from different individuals and were acquired on a 1.5 T MRI scanner. The images have a size of $512 \times 512 \times 192$ with a voxel resolution of $0.49 \times 0.49 \times 0.8$ mm^3. Further details on the image acquisition of this data are discussed by Ikram et al. [12].

The number of PVS in the axial slice 1 cm above the lateral ventricles is highly correlated with the total number of PVS in the CSO [2]. The rater selected this specific slice and annotated it with dots indicating PVS between 1–3 mm in diameter in line with the guidelines described by Adams et al. [1]. The intra-rater performance was evaluated on a separate set of 40 MRI scans (see Table 1 and Fig. 3).

3.2 Preprocessing

Images are preprocessed as proposed by [10]. We segment the CSO with the FreeSurfer multi-atlas segmentation algorithm [7] producing a binary mask that we smooth with a Gaussian kernel. The image are multiplied with the smoothed mask and cropped to a fixed size containing only the slices close to the annotated slice. The resulting images are normalized to the range [0, 1] by dividing by the maximum intensity in the image.

Annotated dots were not always inside the PVS. To solve this problem, we shift the dots to the highest intensity value within the same connected component and within 3 voxels distance. The shifted dots were only used to compute the distance maps for the training and validation set. For evaluation of the detection performance, the original annotated dots were used.

3.3 Experimental Setup

Random sampling was used to split the 2202 scans into a set of 1202 for development of the method (1000 for training and 202 for validation) and a separate set of 1000 for testing. As only one slice per scan was annotated, DMs were computed in 2D and the loss was only evaluated for this slice. Non-maximum suppression and evaluation of detection performance was also only done on the slice that was annotated.

Weights for the convolutional layers were initialized by random sampling from a truncated normal distribution with zero mean and unit variance. For optimization we use Adadelta and a batch of one due to memory limitations. We use on-the-fly augmentation for the training set. For every image a random rotation around the depth direction with a maximum of 20° in both directions is applied combined with random flipping in horizontal and in vertical direction. Methods were implemented in Python and Keras with Tensorflow as backend.

3.4 Detection Performance

The candidate detections of each method are compared to the expert annotations using the hungarian algorithm to find a one-to-one mapping between these sets. Only detections within a 6 voxel radius of the annotations were counted as true positive. We use 6 voxels as this is the maximum PVS diameter (corresponds to 3 mm [1]).

The detection performance is mainly evaluated with the Free-Response Operating Characteristic (FROC) curve and its area under the curve (FAUC) until 10 FP_{avg}, which is approximately twice the FP_{avg} of the rater. The FAUC is calculated as the percentage of the highest possible area. We used bootstrapping to quantify the uncertainty, resulting in a mean FAUC and confidence interval based on 1000 sampled sets. Bootstrapping was performed by random sampling with replacement from the test set.

3.5 Evaluation Approach

We ran experiments varying the decay parameter p (see Fig. 2). For higher values of p the FCN did not train, we expect because of label imbalance. Based on the FAUC on the validation set we set p to 5 for geodesic distance, to 6 for intensity distance and 9 for euclidean distance. During training, the model parameters were chosen as the ones minimizing the FAUC computed on the validation set. Only the best model per distance type ($GDM_5(x)$, $EDM_9(x)$, $IDM_6(x)$) was tested on the test set of 1000 scans.

The operating point on the FROC was chosen per model as the threshold with a sensitivity on the validation set closest to the average intra-rater sensitivity. For $GDM_5(x)$ the threshold was chosen at 0.525, for $EDM_9(x)$ at 0.500 and for $IDM_6(x)$) at 0.495. This threshold was used as the detection threshold during evaluation on the test set.
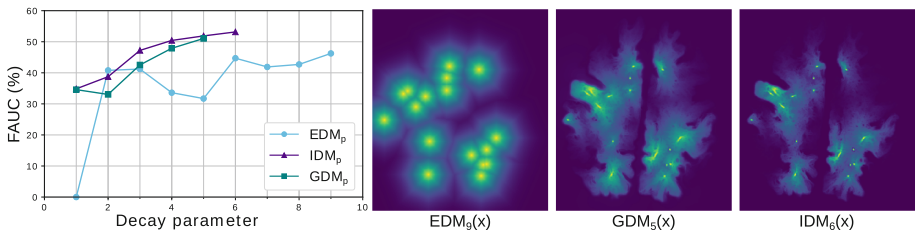

Fig. 2. Influence of decay parameter p on detection performance on the validation set and the chosen distance maps
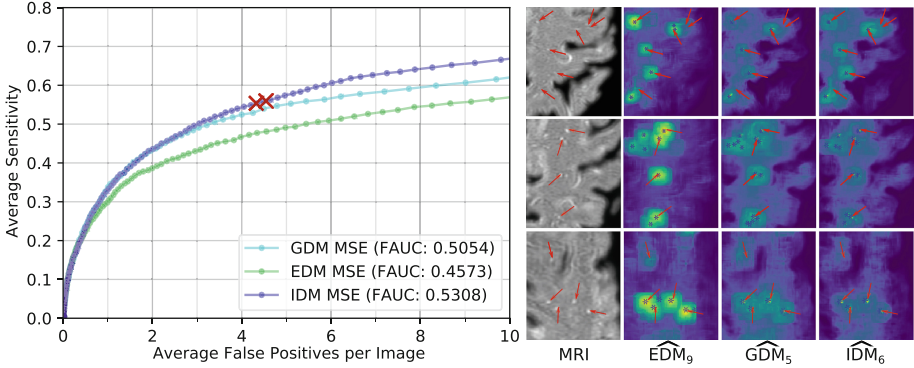
Fig. 3. FROC curves and crops of the output of the FCNs and their proposed detections (intra-rater performance indicated with red crosses, annotations with red arrows and predictions with blue stars) (Color figure online)

3.6   Results

Figure 3 shows the FROC curves computed on the test set and examples of the output of the FCNs. Table 1 shows the corresponding FAUCs, the sensitivity and FP_{avg} of the methods on the test set at the chosen thresholds (based on the validation set) and the average intra-rater performance.

Table 1. PVS detection performance on the test set for the detection methods and the average intra-rater performance on a smaller independent set

	FAUC	FP_{avg}	Sensitivity
$EDM_9(x)$	45.761 ($\pm$ 0.052)	7.49	53.63
$GDM_5(x)$	50.575 ($\pm$ 0.050)	5.10	55.26
$IDM_6(x)$	53.078 ($\pm$ 0.051)	4.35	55.35
Average intra-rater	-	4.43	55.66

4   Discussion and Conclusion

Our experiments indicate that incorporating image intensity information in a distance map used to train an FCN substantially improves performance of PVS detection. Results show that using GDMs and IDMs both result in a similar detection performance, with IDMs sometimes reaching higher performance than GDMs. This indicates that intensity difference is the most discriminative information, and that Euclidean distance could even be ignored. Using higher values of the decay parameter also increases the PVS detection performance, and stabilizes the optimization.

The FCN trained using IDMs reaches a sensitivity and FP_{avg} similar to the intra-rater performance computed on a smaller independent set (Fig. 3).

We expect our method could perform well for detecting other types of focal vascular lesions in the brain. Using intensity information in the computation of DMs could help the detection lesions that either have a complex morphology, or can have substantial variation in their size, such as microbleeds, white matter hyperintensities or lacunes. Additionally, in this work we evaluate the intensity-based distance maps only for their performance in detecting PVS. However, we observe that the PVS detections in the output maps of the FCNs trained on intensity-based distance maps (Fig. 3) seem to approximate the PVS shape quite well. We therefore expect our approach might also work well for segmentation.

References

1. Adams, H.H.H., et al.: Rating method for dilated Virchow-Robin spaces on magnetic resonance imaging. Stroke **44**, 1732–1735 (2013)
2. Adams, H.H., et al.: A priori collaboration in population imaging: the uniform neuro-imaging of Virchow-Robin spaces enlargement consortium. Alzheimer's Dement. Diagn. Assess. Dis. Monit. **1**(4), 513–520 (2015)
3. Ballerini, L., et al.: Perivascular spaces segmentation in brain MRI using optimal 3D filtering. Sci. Rep. **8**(1), 2132 (2018)
4. Boespflug, E.L., Schwartz, D.L., Lahna, D., Pollock, J., Iliff, J.J., Kaye, J.A., et al.: MR imagingbased multimodal autoidentification of perivascular spaces (mMAPS): automated morphologic segmentation of enlarged perivascular spaces at clinical field strength. Radiology **286**(2), 632–642 (2018)
5. Brosch, T., Tang, L.Y.W., Yoo, Y., Li, D.K.B., Traboulsee, A., Tam, R.: Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation. IEEE T-MI **35**(5), 1229–1239 (2016)
6. Charidimou, A., et al.: Enlarged perivascular spaces as a marker of underlying arteriopathy in intracerebral haemorrhage: a multicentre MRI cohort study. J. Neurol. Neurosurg. Psychiatry **84**(6), 624–629 (2013)
7. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., et al.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage **31**(3), 968–980 (2006)
8. Dou, Q., Chen, H., Yu, L., Zhao, L., Qin, J., Wang, D., et al.: Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. TMI **35**(5), 1182–1195 (2016)

9. Dubost, F., et al.: GP-Unet: lesion detection from weak labels with a 3D regression network. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 214–221. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66179-7_25

10. Dubost, F., et al.: Enlarged perivascular spaces in brain MRI: automated quantification in four regions. NeuroImage **185**, 534–544 (2019)

11. Ghafoorian, M., et al.: Deep multi-scale location-aware 3D convolutional neural networks for automated detection of lacunes of presumed vascular origin. NeuroImage: Clin. **14**, 391–399 (2017)

12. Ikram, M.A., et al.: The rotterdam scan study: design update 2016 and main findings. Eur. J. Epidemiol. **30**(12), 1299–1315 (2015)

13. Lian, C., et al.: Multi-channel multi-scale fully convolutional network for 3D perivascular spaces segmentation in 7T MR images. Med. Image Anal. **46**, 106–117 (2018)

14. Meyer, M.I., Galdran, A., Mendonça, A.M., Campilho, A.: A pixel-wise distance regression approach for joint retinal optical disc and fovea detection. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11071, pp. 39–47. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00934-2_5

15. Qi, H., Collins, S., Noble, J.A.: Automatic lacunae localization in placental ultrasound images via layer aggregation. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11071, pp. 921–929. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00934-2_102

16. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

17. Toivanen, P.J.: New geodosic distance transforms for gray-scale images. Pattern Recogn. Lett. **17**(5), 437–450 (1996)

18. Xie, Y., Xing, F., Shi, X., Kong, X., Su, H., Yang, L.: Efficient and robust cell detection: a structured regression approach. Med. Image Anal. **44**, 245–254 (2018)