Convolutional Neural Networks for the Identification of Regions of Interest in PET Scans: A Study of Representation Learning for Diagnosing Alzheimer's Disease

Andreas Karwath, Markus Hubrich, Stefan Kramer(✉),
and the Alzheimer's Disease Neuroimaging Initiative

Institut für Informatik, Johannes Gutenberg-Universität, Mainz, Germany
kramer@informatik.uni-mainz.de

Abstract. When diagnosing patients suffering from dementia based on imaging data like PET scans, the identification of suitable predictive regions of interest (ROIs) is of great importance. We present a case study of 3-D Convolutional Neural Networks (CNNs) for the detection of ROIs in this context, just using voxel data, without any knowledge given a priori. Our results on data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) suggest that the predictive performance of the method is on par with that of state-of-the-art methods, with the additional benefit of potential insights into affected brain regions.

Keywords: PET Scans · Alzheimer's disease · Dementia · Deep learning · Convolutional Neural Networks · Regions of interest · Representation learning

1 Introduction

Alzheimer's disease is a progressive, degenerative and incurable disease of the brain and a common cause of dementia. One of the imaging modalities often used in the Computer-Aided Diagnosis (CAD) of Alzheimer's disease is positron emission tomography (PET). Recent years have seen a number of machine learning based approaches applied to PET scans to distinguish between Alzheimer's disease (AD), mild cognitive impairment (MCI) and normal control (NC), amongst others [1,2]. Deep learning approaches have attracted particular interest in attempts to transfer the progress made in computer vision to medical image analysis.[1] Besides classification, the identification of regions of interest (ROIs) is of major concern in medical research, i.e. regions in 2-D or 3-D medical images that are of particular importance for assessing the functioning of organs, for diagnosis and for understanding disease progression. In this paper, we continue

[1] For a good overview see the recent paper by Vieira *et al.* [2]. Apparently many more papers can be found in online archives than papers that have appeared already.

A. ten Teije et al. (Eds.): AIME 2017, LNAI 10259, pp. 316–321, 2017.
DOI: 10.1007/978-3-319-59758-4_36

our previous work [1] and report first results towards the identification of ROIs in PET scans, in the context of AD diagnosis, using deep learning methods. While deep learning on PET scans is currently an active research topic and ROIs have been investigated in other contexts (e.g. investigating different tracers), the combination of deep learning and ROIs has not received much attention yet. The approach employs a binary convolutional neural network (CNN) based classifier to generate importance maps of PET voxel data.

2 Methodology

The main aim is to establish ROIs with respect to AD without medical background knowledge in an automated manner. In the following, we describe how to apply a binary classifier to generate importance maps for extracting ROIs.

2.1 Data Acquisition and Preprocessing

To train the binary classifier used to extract importance maps (or ROIs), we employed the Alzheimer's Disease Neuroimaging Initiative[2] (ADNI) data set. The initiative has a large pool of PET images (co-registered, averaged), which have been acquired on various scanners using different imaging parameters. In the work presented here, we chose the $\left[^{18}\mathrm{F}\right]$ fluorodeoxyglucose PET scans from enrolled normal control (NC), mild cognitive impairment (MCI) and Alzheimer's disease (AD) subjects. Not all scans from a single patient possess the same class, as the disease might occur at a later stage of a patient's life time. To avoid possible information leakage during the validation of our approach, we always employed a k-fold validation based on patients and not on scans. Our aim is to use this data to compare the predictive performance of the identified ROIs to a smaller data set used in a previous study by Li *et al.* [1]. However, as this data also originates from the same source, we excluded 84 subjects from the ADNI data set (see Table 1). Overall, we used 1258 subjects and 2630 scans.

The ADNI data set contains globally normalized scans [3], which are not suitable for our purposes as this kind of normalization does not easily allow the

Table 1. The two different data sets used in this work originating from the ADNI project, $\mathcal{D}_{\mathcal{C}}$ the ADNI data set without the patients used in Li *et al.* ($\mathcal{D}_{\mathcal{ROI}}$).

Data Set	Name	Class Label	No.Patients (male:female)	Age	No.Scans
$\mathcal{D}_{\mathcal{C}}$	ADNI \ Li *et al.*	NC	396 (186:210)	74 (±6)	749
		MCI	662 (376:286)	73 (±7)	1292
		AD	330 (199:131)	76 (±8)	589
$\mathcal{D}_{\mathcal{ROI}}$	Li *et al.*	NC	30 (21:9)	74 (±5)	30
		MCI	29 (23:6)	74 (±6)	29
		AD	25 (15:10)	72 (±6)	25

[2] URL: http://adni.loni.usc.edu/ (visited on Jan. 12, 2016).

comparison of PET images to each other. Therefore, and to allow comparison to the works of Li *et al.*, we used a different type of normalization. We employed the same kind of spatial normalization as Li *et al.* based on the Automated Anatomical Labeling (AAL) digital brain atlas: The normalization yields scans in a $91 \times 109 \times 91$ with $2\,\text{mm}^3$ voxel format. Next, each scan is intensity normalized with respect to its cerebral global mean (also called grand mean normalization). Finally, all scans are smoothed by an 8 mm FWHM Gaussian kernel.

2.2 Generation of Binary Classifiers

As mentioned above, the approach taken is to train a binary classifier in order to extract informative regions of interest (ROIs). To train the binary classifier a 3-D CNN is employed, expected to be able to incorporate local voxel relations as well as to extract complex spatial patterns. Training a 3-D CNN on input volumes with a size of $91 \times 109 \times 91$ can be computationally expensive. Therefore, we down-sampled all scans reducing the number of operations. First, all non-brain voxels are removed to eliminate noise caused by the PET scanners. As a consequence, we trim all scans to their minimum bounding box, which yields a new size of $88 \times 108 \times 88$ voxels. The first and last ten sagittal slices contain very little brain volume, but a considerably large ratio of the skull. Hence, we discard these slices and obtain cubic volumes of dimension $88 \times 88 \times 88$. Next, the scans are down-sampled using "average pooling", i.e. the input is divided into cubic pooling regions of a particular size and the output consists of the average of each region. We chose a pooling size of 4, resulting in scans of size $22 \times 22 \times 22$.

For the binary classifier, we chose the following deep architecture: Overall seven convolutional layers are followed by three fully connected layers, where all of them use rectified linear units as activation function. The input layer corresponds to the 22^3 down-sampled scan size and the output layer consists of a single sigmoid output node (see Fig. 1). All feature maps use a stride of 1. We employed 10-fold cross validation to estimate the optimal parameters using $\mathcal{D}_\mathcal{C}$. To cater for variations, we augmented the data set by mirroring the original scans on the center coronal plane. Furthermore, we used a mini-batch size of 128 and trained the final classifier $\mathcal{C}$ on $\mathcal{D}_\mathcal{C}$ with the parameters found in the 10-fold cross validation.

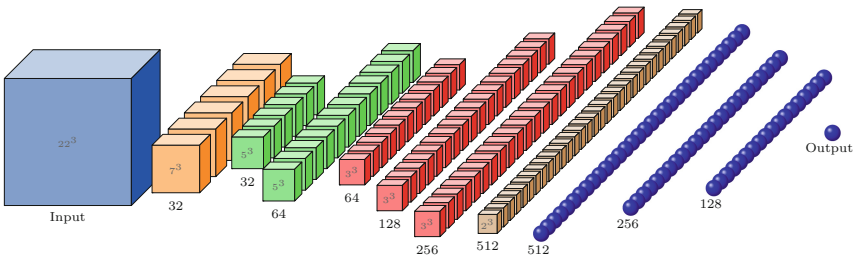


Fig. 1. The architecture for the CNN classifier.

2.3 ROI Discovery and Ranking

Given the binary classifier $\mathcal{C}$ from the previous section and a scan X, one can assign each voxel of X a value of its importance towards the classification of X. We assume $X : \mathbb{N}^{n\times n\times n} \rightarrow \mathbb{R}_{\geq 0}$. For $1 \leq x, y, z \leq n$, we denote with v_{xyz} the voxel at position (x, y, z). We can exclude the information of v_{xyz} from X by setting $X \setminus v_{xyz}(x, y, z) = X(x, y, z) = \gamma$[3]. We denote $Conf_{\mathcal{C}}(X) \in [0, 1]$ to be the confidence for correctly classifying the example X by $\mathcal{C}$. Using this information, the influence of $v_{xyz} \in X$ can be assessed as $\Delta Conf(\mathcal{C}, X, v_{xyz}) = Conf_{\mathcal{C}}(X) - Conf_{\mathcal{C}}(X \setminus v_{xyz})$. If $\Delta Conf(\mathcal{C}, X, v_{xyz}) = 0$, v_{xyz} can be regarded as redundant and can be omitted. However, if $\Delta Conf(\mathcal{C}, X, v_{xyz}) > 0$, the voxel v_{xyz} can be considered as important. Finally, $\Delta Conf(\mathcal{C}, X, v_{xyz}) < 0$ indicates a decrease in the classification performance. Let $b^k_{\tilde{x}\tilde{y}\tilde{z}}$ be a cubic block of voxels with center $(\tilde{x}, \tilde{y}, \tilde{z})$ and size $2k-1$. Analogously to the exclusion of single voxels, we can exclude the information contained in a block from X as follows:

$$X \setminus b^k_{\tilde{x}\tilde{y}\tilde{z}}(x, y, z) = \begin{cases} \gamma, & \text{if } v_{xyz} \in b^k_{\tilde{x}\tilde{y}\tilde{z}} \\ X(x, y, z), & \text{otherwise.} \end{cases} \tag{1}$$

Using this, we can can generate a map of voxel importance for the exclusion of blocks with length $2k - 1$ as $I^k_{\mathcal{C},X}(x, y, z) = \frac{1}{|b^k_{xyz}|}\Delta Conf(\mathcal{C}, X, b^k_{xyz})$. To accommodate for blocks in the border areas, we normalize $I^k_{\mathcal{C},X}$ using the total amount of voxels in a particular block, i.e. $|b^k_{xyz}|$. To investigate the varying size of potential patterns, we employed different block sizes k from $\mathcal{K} \subset \{1, 2, \ldots, \frac{n+1}{2}\}$.

Workflow 1: General workflow for the extraction of ROIs

Input: $\mathcal{D}_{\mathcal{C}}$: a data set to train a binary classifier, $\mathcal{D}_{ROI}$: a data set to extract ROIs, δ: a threshold value, $\mathcal{K}$: the set of block sizes

Output: an importance map of voxels ($\geq \delta$)

$\mathcal{C} = Train_{\mathcal{C}}(\mathcal{D}_{\mathcal{C}})$ (train binary classifier $\mathcal{C}$ using $\mathcal{D}_{\mathcal{C}}$)

$\forall(x, y, z) \in \mathbb{N}^{n\times n\times n}, \forall X \in \mathcal{D}_{\mathcal{ROI}}$ calculate $\mathcal{I}^{\mathcal{K}}_{\mathcal{C},X}(x, y, z)$

with $\mathcal{I}^{\mathcal{K}}_{\mathcal{C},X}(x, y, z) = \frac{1}{|\mathcal{K}|}\sum_{k\in\mathcal{K}} I^k_{\mathcal{C},X}(x, y, z), \quad \mathcal{K} \subset \{1, 2, \ldots, \frac{n+1}{2}\}$

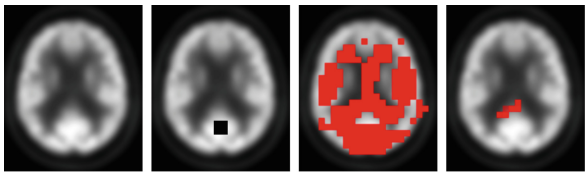
return $sort(filter(\forall(x, y, z) \in \mathbb{N}^{n\times n\times n}, \forall X \in \mathcal{D}_{\mathcal{ROI}}\ \mathcal{I}^{\mathcal{K}}_{\mathcal{C},X}(x, y, z) \geq \delta))$

Parts of the process of generating $\mathcal{I}^{\mathcal{K}}_{\mathcal{C},X}$ (definition in Workflow 1) can exemplarily seen in Fig. 2.

3 Experiments and Results

First, we trained the initial binary classifier $\mathcal{C}$ as described above on $\mathcal{D}_{\mathcal{C}}$ using all scans with label NC and AD. The accuracy of a 10-fold cross validation was 89%, with $TPr = 0.85$, $TNr = 0.91$ and $AUC = 0.95$. To evaluate the benefit of using importance maps, we used two different settings on $\mathcal{D}_{\mathcal{ROI}}$ using a 10×10-fold

[3] Different γs are possible. Here, we used $\gamma_1 = \bar{v}_{xyz}$ (the average voxel of all scans in the data set) as well as $\gamma_2 = round\left(\frac{v_{xyz}}{max(X)}\right)$.



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

Fig. 2. The original scan X is shown in (a). In (b), a block of voxels is excluded from X. Subfigure (c) shows all voxels of $\mathcal{I}^{\mathcal{K}}_{\mathcal{C},X}$ for a threshold of $\delta \geq 10^{-6}$. Lastly, (d) shows some of the 50 most informative voxels obtained by $\mathcal{I}^{\mathcal{K}}_{\mathcal{C},X}$.

Table 2. Performance measure obtained from different settings.

Performance Measure	Li *et al.*	S_{all}	$S_{\mathcal{ROI}}$			
			$\delta = 0.001$ ($\#v = 3251$)	$\delta = 0.05$ ($\#v = 198$)	$\delta = 0.10$ ($\#v = 42$)	$\delta = 0.15$ ($\#v = 16$)
Acc (%)	88.1	$87.1_{\pm 2.5}$	$87.0_{\pm 1.8}$	$91.6_{\pm 2.9}$	$85.6_{\pm 3.8}$	$79.3_{\pm 2.7}$
TPr	0.91	$0.84_{\pm 0.06}$	$0.81_{\pm 0.05}$	$0.88_{\pm 0.05}$	$0.82_{\pm 0.06}$	$0.73_{\pm 0.06}$
TNr	0.83	$0.90_{\pm 0.02}$	$0.92_{\pm 0.02}$	$0.95_{\pm 0.02}$	$0.89_{\pm 0.05}$	$0.85_{\pm 0.04}$
AUC	0.97	$0.94_{\pm 0.02}$	$0.95_{\pm 0.02}$	$0.99_{\pm 0.01}$	$0.94_{\pm 0.02}$	$0.92_{\pm 0.03}$

cross validation. Setting S_{all} uses all available voxels, while $S_{\mathcal{ROI}}$ only employs voxels considered important according to varying δs. To only evaluate the use of individual voxels and not their spatial neighbours, we train linear support vector machines (SVM) based on individual voxels. A simple grid search for C was performed to find optimal parameters. We used eight folds for training, one as validation set, and the final one to evaluate the performance using the optimal C. Furthermore, we employed block sizes $\mathcal{K} = \{1, 3, 5\}$ and γ_2. The averaged performance of S_{all} and $S_{\mathcal{ROI}}$ for different δs is given in Table 2. The comparison to the works by Li *et al.* in the table is given as indication of state-of-the-art methods using the same data set $\mathcal{D}_{\mathcal{ROI}}$.

4 Conclusions and Future Work

We have presented an approach to extract regions of interest from PET scans based on binary classifiers. The extracted regions, or importance maps, can be employed as a voxel subset to differentiate between scans labeled NC and AD. The binary classifier employed here is based on a deep CNN architecture for the extraction of complex spatial patterns. In contrast to many deep learning approaches just aiming for the classification, our approach explicitly aims for ROIs, but is, at this point, evaluated only quantitatively. As a next step, we are going to evaluate the ROIs qualitatively together with our collaboration partners.

Acknowledgments. This work was partially supported by the Carl-Zeiss-Foundation Competence Center for High Performance Computing.

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

References

1. Li, R., Perneczky, R., Drzezga, A., Kramer, S.: Gaussian mixture models and model selection for [18F] fluorodeoxyglucose positron emission tomography classification in Alzheimer's disease. PLoS ONE **10**(4), e0122731 (2015)
2. Vieira, S., Pinaya, W.H., Mechelli, A.: Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. Neurosci. Biobehav. Rev. **74**, (Part A), 58–75 (2017)
3. Jagust, W.J., Landau, S.M., Koeppe, R.A., Reiman, E.M., Chen, K., Mathis, C.A., Price, J.C., Foster, N.L., Wang, A.Y.: The ADNI PET core: 2015. Alzheimers Dement. **11**(7), 757–771 (2015)