

Classification methods on different brain imaging modalities for Alzheimer disease studies

A. Khvostikov^{1,2}, J. Benois-Pineau¹, A. Krylov², G. Catheline³

xubiker@gmail.com|jenny.benois-pineau@u-bordeaux.fr

kryl@cs.msu.ru|gwenaelle.catheline@u-bordeaux2.fr

¹LaBRI UMR 5800, University of Bordeaux, Bordeaux, France;

²CMC Depart., Lomonosov Moscow State University, Moscow, Russia;

³INICIA UMR 5287, University Victor Segalen Bordeaux 2, Bordeaux, France

Computer-aided early diagnosis of Alzheimer's Disease (AD) and its prodromal form, Mild Cognitive Impairment (MCI), has been the subject of extensive research in recent years. Some recent studies have shown promising results in the AD and MCI determination using structural and functional Magnetic Resonance Imaging (sMRI, fMRI), Positron Emission Tomography (PET) and Diffusion Tensor Imaging (DTI) modalities.

This paper reviews the major trends in automatic classification methods such as feature extraction based methods as well as deep learning approaches in medical image analysis applied to the field of Alzheimer's Disease diagnostics. Different fusion methodologies to combine heterogeneous image modalities to improve classification scores are also considered.

Keywords: *Medical imaging, Alzheimer's Disease, Mild cognitive impairment, Machine Learning, Deep learning, Convolutional neural networks, Fusion, Review*

1. Introduction

Alzheimer's Disease (AD) is the most common type of dementia. Dementia refers to diseases that are characterized by a loss of memory or other cognitive impairments, and is caused by nerve cells degeneration in the brain.

Structural Magnetic Resonance Imaging (sMRI) and amyloid-Positron Emission Tomography (amyloid-PET) have been added in the newly proposed criteria for the predementia phase of the disease [1]. Other modalities as diffusion MRI (DTI) and functional MRI measured at rest (fMRI) are not yet used in these definitions. However, AD subjects also present modification on these two modalities as well [2, 3].

Since AD-related neurodegeneration is associated with gray matter atrophy, most previous works were focused on volumetric approaches that are based on comparison of anatomical brain structures assuming one-to-one correspondence between subjects. The wide-spread voxel-based morphometry (VBM) [4] is an automatic volumetric method for studying the differences in local concentrations of white and gray matter and comparison of brain structures of the subjects to test with reference normal control (NC) brains. Tensor-based morphometry (TBM) [5] was proposed to identify local structural changes from the gradients of deformations fields when matching tested brain and the reference healthy NC. Object-based morphometry (OBM) [6] was introduced for shape analysis of anatomical structures.

In general, the automatic classification on brain images of different modalities can be applied to the whole brain [7–10], or performed using the domain knowledge on specific regions of interest (ROIs). Concerning AD, hippocampal structural changes are strongly correlated to the severity of disease [11]. The

changes in such regions are considered as AD biomarkers.

Advances in computer vision and content-based image retrieval research make penetrate the so-called feature-based methods into classification approaches for (AD) detection [12–14]. The reason for this is in inter-subject variability, which is difficult to handle in VBM. On the contrary, the quantity of local features which can be extracted from the brain scans together with captured particularities of the image signal allowed an efficient classification with lower computational workload [14]. The obtained feature vectors are classified using machine learning algorithms.

Lately with the development of neural networks the feature-based approach became less popular and is gradually replaced with convolutional neural networks of different architectures.

The goal of the present paper is to give a substantial overview of these recent trends in classification of different brain imaging modalities in the problem of computer-aided diagnostics of Alzheimer disease and its prodromal stage, i.e. mild cognition impairment (MCI).

The paper is organized as follows. In Section 2 we will overview the main feature-based approaches and in Section 3 we will compare different approaches based on using neural networks. Particular attention will be paid in each case to fusion of modalities. All reviewed approaches are compared in Table 1. Section 4 concludes our review.

2. Feature-based classification

Feature-based classification can be performed on images of different modalities. Here we compare and discuss the usage of sMRI, DTI and sMRI fusion with other modalities.

sMRI In previous joint work [13], Ahmed et al. computed local features on sMRI scans in hippocampus and posterior cingulate cortex (PCC) structures of the brain. The originality of the work consisted in the usage of Gauss-Laguerre Circular Harmonic Functions (GL-CHFs) instead of traditional SIFT and SURF descriptors [15, 16]. CHFs perform image decomposition on the orthonormal functional basis, which allows capturing local directions of the image signal and intermediate frequencies. It is similar to Fourier decomposition, but is more appropriate in case of smooth contrasts of MRI modality. For each projection of each ROI a signature vector was calculated using a bag-of-visual-words model (BoVWM) with a low-dimensional dictionary with 300 clusters. This led to the total signature length of 1800 per image. Principal component analysis was then applied to reduce the signature length to 278. The signatures then were classified using SVM with RBF kernel and 10-fold cross-validation and reached the accuracy level of 0.838, 0.695, 0.621 for AD/NC, NC/MCI and AD/MCI binary classification problems accordingly.

DTI This modality is probably the most recent to be used for AD classification tasks. Both Mean Diffusivity (MD) and Fractional Anisotropy (FA) maps are being explored for this purpose. In [17] the authors acquired DTI images of 15 AD patients, 15 MCI patients, and 15 healthy volunteers (NC). After the preprocessing steps the FA map, which is an indicator of brain connectivity, was calculated. The authors considered 41 Brodmann areas, calculated the connectivity matrices for this areas and generated a connectivity graph with corresponding 41 nodes. Two nodes corresponding to Brodmann areas are marked with an edge if there is at least one fiber connecting them. Then the graph is described with the vector of features, calculated for each node and characterizing the connectivity of the node neighborhood. Totally each patient is characterized by 451 feature. The vectors were reduced to the size of 430 and 110 using ANOVA-based feature selection approach. All vectors were classified with the ensemble of classifiers (Logistic regression, Random Forest, Gaussian native Bayes, 1-nearest neighbor, SVM) using 5-fold cross-validation. The authors have achieved the 0.8, 0.833, 0.7 accuracy levels for AD/NC, AD/MCI and MCI/NC accordingly on their custom database.

Another methodology is described in [18]. The authors use the fractional anisotropy (FA) and mode of anisotropy (MO) values of DTI scans of 50 patients from the LONI Image Data Archive (<https://ida.loni.usc.edu>). After non-linear registration to the standard FA map, the authors calculate the skeleton of the mean FA image as well as MO and perform the second step of registration. After that a Relief feature algorithm is performed on all voxels of the image, relevant ones are used for 10-cross valida-

tion training the SVM classifier with RBF kernel. The declared accuracy is 0.986 and 0.977 for classification AD/MCI, AD/NC accordingly.

Data fusion In [19] authors use a fusion of sMRI and PET images together with canonical correlation analysis (CCA). After preprocessing and aligning images of 2 modalities given the covariance data of sMRI and PET image they find the projection matrices by maximizing the correlation between projected features. Here

$$X_1 \in R^{d \times n}, X_2 \in R^{d \times n}$$

are the d -dimensional sMRI and PET features of n samples,

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

is a covariance matrix,

$$(B_1, B_2) = \arg \max_{(B_1, B_2)} \frac{B_1^T \Sigma_{12} B_2}{\sqrt{B_1^T \Sigma_{11} B_1} \sqrt{B_2^T \Sigma_{22} B_2}}$$

are the projection matrices and

$$Z_1 = B_1^T X_1, Z_2 = B_2^T X_2$$

are the resulting projections. The authors construct the united data representation for each patient:

$$F = [X_1; X_2; Z_1; Z_2] \in R^{4d \times n}$$

and calculate SIFT descriptors. This descriptors are used to form the BoVW model, the classification is performed using SVM. The achieved accuracy is 0.969, 0.866 and for classifying AD/NC and MCI/NC accordingly.

Ahmed et al. in [14] demonstrated the efficiency of using the amount of cerebrospinal fluid (CSF) in the hippocampal area calculated by an adaptive Otsu's thresholding method as an additional feature for AD diagnostics. In [12] they further improved the result of [13] by combining visual features derived from sMRI and DTI MD maps with a multiple kernel learning scheme (MKL). Similar to [13] they selected hippocampus ROIs on the axial, saggital and coronal projections and described them using Gauss-Laguerre Harmonic Functions (GL-CHFs). These features are clustered into 250 and 150 clusters for sMRI and MD DTI modalities and encoded using the BoVW model. Thus they got three sets of features: BoVW histogram for sMRI, BoVW histogram for MD DTI and CSF features. The obtained vectors are classified using MKL approach based on SVM. The achieved accuracy is 0.902, 0.794, 0.766 for AD/NC, MCI/NC and AD/MCI classification.

3. Classification with neural networks

Deep neural networks (DNN) and specifically convolutional NN have become popular now due to their good generalization capacity and available GPU Hardware needed for parameter optimization. Their main drawback for AD classification is the small amount of available training data and also a low resolution of input images when the ROIs are considered. This problem can be eliminated in several ways: i) by using

shallow networks with relatively small number of neurons, ii) applying transfer learning from an existing trained network or iii) pretraining some of the layers of the network.

Forming shallow networks kills the idea of deep learning to recognize structures at different scales and reduces the generalization ability of the network, so this methodology has not often been used since recently, despite it has shown decent results [20].

One way to enlarge the dataset is to use domain-dependent data augmentation. In case of medical images this often comes down to mirror flipping, small-magnitude translations and weak gaussian blurring [20].

Autoencoders The idea of pretraining some of the layers in the network is easily implemented with autoencoders (AE) or in image processing tasks more often with convolutional autoencoders (CAE). Autoencoder consists of an input layer, hidden layer and an output layer, where the input and output layers have the same number of units (Fig.1). Given the input vector $x \in \mathbb{R}^n$ autoencoder maps it to the hidden representation h :

$$h = f(Wx + b),$$

where $W \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^p$, n is the number of input units, p is the number of hidden units, f is an encoder function e.g. sigmoid. After that the hidden representation h is mapped back to $\hat{x} \in \mathbb{R}^n$:

$$\hat{x} = g(\widehat{W}h + \widehat{b}),$$

where $\widehat{W} \in \mathbb{R}^{n \times p}$, $\widehat{b} \in \mathbb{R}^n$, g is the identity function. The weights and biases are found by gradient methods to minimize the cost function:

$$J(W, b) = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} \|\hat{x}^{(i)} - x^{(i)}\|^2,$$

where N is the number of inputs.

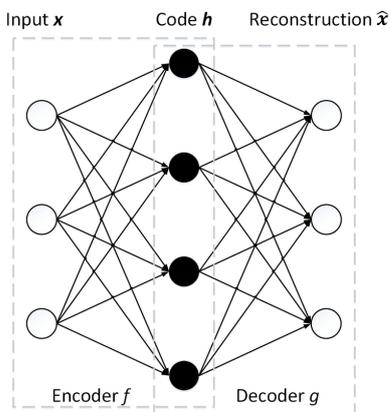


Figure 1. Architecture of autoencoder

The overcompleted hidden layer is used to make the autoencoder extracting features.

Introducing spatial constraints with convolutions easily aligns the model of autoencoder to the convolutional autoencoder (CAE) and 3D convolutional autoencoder (3D-CAE).

In [8] authors added a sparsity constraint to prevent hidden layers of autoencoder from learning the identity function. They use 3D convolutions on the both sMRI and PET modalities and train the autoencoder on random $5 \times 5 \times 5$ image patches. Maxpooling, fully-connected and softmax layers were applied after autoencoding. Mixing data of sMRI and PET modalities is performed at FC layer. The use of autoencoders allowed the authors to increase the classification accuracy by 4-6% and leads to the level of 0.91 for AD/NC classification.

Nearly the same approach with sparse 3D autoencoder was used in [9] to classify sMRI images into 3 categories (AD/MCI/NC). The proposed network architecture is shown in Fig.2. Larger obtained dataset and more accurate network parameters configuration allowed the authors to reach the accuracy of 0.954, 0.868 and 0.921 in AD/NC, AD/MCI and NC/MCI determination accordingly.

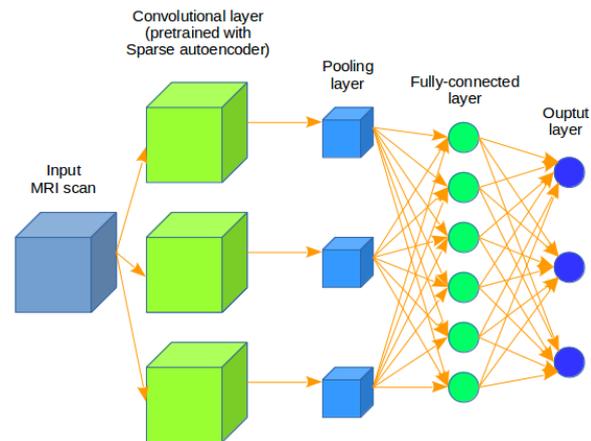


Figure 2. Typical CNN architecture with CAE pretraining.

The authors of [7] extended the idea of applying autoencoders. They proposed using three stacked 3D convolutional autoencoders instead of only one. Two fully-connected layers before the softmax were used for a progressive dimension reduction. The usage of stacked 3D CAE allowed the authors to achieve one of the best accuracy level: 0.993, 1, 0.942 for AD/NC, AD/MCI and MCI/NC classification on sMRI images only.

Transfer Learning Transfer learning is considered as the transfer of knowledge from one learned task to a new task in machine learning. In the context of neural networks, it is transferring learned features of a pre-trained network to a new problem. Glozman and Liba in [21] used the widely known AlexNet [22], pretrained

on the ImageNet benchmark and fine-tuned the last 3 fully-connected layers (Fig.3). The main problem of transfer learning is the necessity to transform the available data so that it corresponds to the network input. In [21] the authors created several 3-channel 2D images from the 3D input of sMRI and PET images by choosing central and nearby slices from axial, coronal and sagittal projections. They then interpolated the slices to the size 227×227 compatible with AlexNet. Naturally one network was used for each projection. To augment the source data only mirror flipping was applied. This transfer learning based approach allowed the authors to reach 0.665 and 0.488 accuracy on 2-way (AD/NC) and 3-way (AD/MCI/NC) classifications accordingly on a subset of ADNI database.

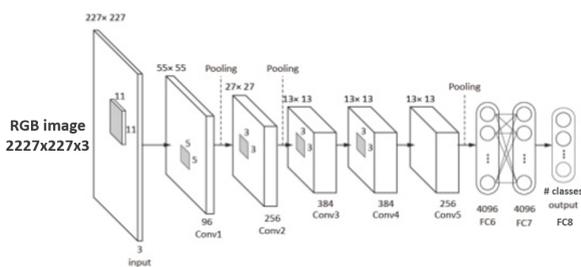


Figure 3. AlexNet architecture. Includes 5 convolutional layers and 3 fully-connected layers.

2D convolutional neural networks In [23], [24], [25] the authors compared the classification of structural and functional MRI images using the LeNet-5 architecture by transforming the source 3D and 4D (in case of MRI) data to a batch of 2D images. LeNet-5 consists of two convolutional and two fully-connected layers. The reached level of accuracy for 2-class classification (AD/NC) was 0.988 for sMRI and 0.999 for fMRI images.

Billones et al. proposed in [10] to use a modified 16-layered VGG network [26] to classify sMRI images. The key feature of this paper was to use 2D convolutional network to classify each slice of source data separately. The authors selected 20 central slices for each image and the final score was calculated as the output of the last softmax layer of the network. The accuracy of each slice among all images was also studied, 17 slices were selected as representative, 3 slices (the first and two last slices in the image sequence) demonstrated lower level of accuracy. All in all authors reached a very good accuracy level: 0.983, 0.939, 0.917 for AD/NC, AD/MCI and MCI/NC classification.

In [20] Aderghal et al. used 3 central slices in each projection of a hippocampal ROI. The network architecture represented three 2D convolutional networks (one network per projection) that were joined in the last fully-connected layer. The reached accuracy

for AD/NC, AD/MCI and MCI/NC classification is 0.914, 0.695 and 0.656 accordingly.

Other networks A new approach was proposed in [27]. Shi et al. used a deep polynomial network to analyze sMRI and PET images. It differs from classical CNNs by non-linearity of operations. The building block of the architecture is shown in Fig.4. Here, n^i represents a layer of nodes, (+) means a layer of nodes that calculate the weighted sum $n(z) = \sum_i w_i z_i$, all other nodes compute $n(z_1, z_2) = \sum_i w_i (z_1)_i (z_2)_i$. These blocks were combined into a deep network, the input layers were fed with the average intensity of the 93 ROIs selected on sMRI and PET brain images.

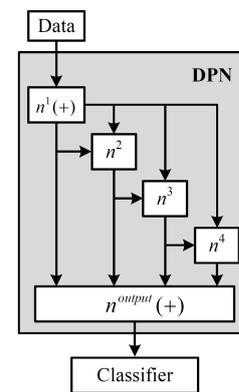


Figure 4. An example of a DPN module.

This architecture allowed the authors to reach very good level of accuracy: 0.971, 0.872 for AD/NC, MCI/NC classification. The used algorithm also demonstrated a good level of accuracy (0.789) for MCI-C/MCI-NC determination, where MCI-C stands for MCI patients that lately converted to AD and MCI-NC stands for MCI patient that were not converted.

In [28] the authors compared the residual and plain 3D convolutional neural networks for sMRI image classification. Here the authors examined the four binary classification tasks AD/LMCI/EMCI/NC, where LMCI and EMCI stands for the late and early MCI stages accordingly. Both networks demonstrated nearly the same performance level, the best figures being obtained for AD/NC classification 0.79-0.8.

In [29] Suk et al. try to combine two different methods: sparse regression and convolutional neural networks. The authors got different sparse representations of the 93 ROIs of the sMRI data by varying the sparse control parameter, which allowed them to produce different sets of selected features. Each representation is a vector, so the result of generating multiple representations can be treated as a matrix. This matrix is then fed to the convolutional neural network with 2 convolutional layers and 2 fully-connected layers. This approach led to the classification accuracy

| Algorithm | Methodology | Modalities | Content | Data (size) | Accuracy | | |
|-----------------------|---------------|------------|------------|-------------|--------------|--------|--------|
| | | | | | AD/NC | AD/MCI | MCI/NC |
| Magnin et al. [6] | Volumetric | sMRI | Full brain | custom (38) | 0.945 | - | - |
| Ahmed et al. [13] | Feature-based | sMRI | 2 ROIs | ADNI (509) | 0.838 | 0.695 | 0.621 |
| Ebadi et al. [17] | Feature-based | DTI | Full brain | custom (34) | 0.8 | 0.833 | 0.7 |
| Lee et al. [18] | Feature-based | DTI | Full brain | LONI (141) | 0.977 | 0.977 | - |
| Lei et al. [19] | Feature-based | sMRI + PET | Full brain | ADNI (398) | 0.969 | - | 0.866 |
| Ahmed et al. [14] | Feature-based | sMRI + DTI | 1 ROI | ADNI (203) | 0.902 | 0.766 | 0.794 |
| Vu et al. [8] | NN-based | sMRI + PET | Full brain | ADNI (203) | 0.91 | - | - |
| Payan and Montana [9] | NN-based | sMRI | Full brain | ADNI (2265) | 0.993 | 1 | 0.942 |
| Glozman and Liba [21] | NN-based | sMRI + PET | Full brain | ADNI (1370) | 0.665 | - | - |
| Sarraf et al. [23] | NN-based | sMRI, fMRI | Full brain | ADNI (302) | 0.988, 0.999 | - | - |
| Billones et al. [10] | NN-based | sMRI | Full brain | ADNI (900) | 0.983 | 0.939 | 0.917 |
| Aderghal et al. [20] | NN-based | sMRI | 1 ROI | ADNI (815) | 0.914 | 0.695 | 0.656 |
| Shi et al. [27] | NN-based | sMRI + PET | Full brain | ADNI (202) | 0.971 | - | 0.872 |
| Korolev et al. [28] | NN-based | sMRI | Full brain | ADNI (231) | 0.79-0.8 | - | - |
| Suk et al. [29] | NN-based | sMRI | 93 ROIs | ADNI (805) | 0.903 | - | 0.742 |

Table 1. Comparison of different classification methods.

level of 0.903 and 0.742 for AD/NC and MCI/NC classification.

4. Discussion and Conclusion

As it can be seen from the Table 1 relatively new feature-based and neural network-based methods demonstrate very good level of performance compared to the classical volumetric methods that are performed manually by medical experts.

It should be mentioned, that the direct comparison of the reviewed algorithms for Alzheimer's disease diagnostics is impossible. The proposed results were obtained using images from several databases and in different quantities (see Table 1). Moreover different classification problems were challenged: although most papers focus on the 3-class AD/MCI/NC binary classification, some of them consider only 2-class AD/NC classification [8, 23–25] and even 4-class AD/eMCI/lMCI/NC classification [28]. Also [10, 29] deserve special attention as the authors try to solve a problem in demand of prediction of Alzheimer converters.

This review allows the community of researchers working on AD classification problems to position their approach and design more efficient classification schemes.

5. Acknowledgements

This research was supported by Ostrogradsky scholarship grant established by French Embassy in Russia. We thank Pierrick Coupé from LABRI UMR 5800 University of Bordeaux/CNRS/Bordeaux-IPN who provided insight and expertise that greatly assisted the research.

6. References

- [1] Jeffrey L Cummings, Bruno Dubois, et al. International work group criteria for the diagnosis of alzheimer disease. *Medical Clinics of North America*, 97(3):363–368, 2013.
- [2] Massimo Filippi and F Agosta. Imaging biomarkers in multiple sclerosis. *Journal of Magnetic Resonance Imaging*, 31(4):770–788, 2010.
- [3] Julio Acosta-Cabronero and Peter J Nestor. Diffusion tensor imaging in alzheimer's disease: insights into the limbic-diencephalic network and methodological considerations. *Frontiers in aging neuroscience*, 6, 2014.
- [4] J. Ashburner and K. Friston. Voxel-based morphometry—the methods. *Neuroimage*, 11(6):805–821, 2000.
- [5] C. Studholme et al. Deformation-based mapping of volume change from serial brain mri in the presence of local tissue contrast change. *IEEE transactions on Medical Imaging*, 25(5):626–639, 2006.
- [6] B. Magnin et al. Support vector machine-based classification of alzheimer's disease from whole-brain anatomical mri. *Neuroradiology*, 51(2):73–83, 2009.
- [7] E. Hosseini-Asl et al. Alzheimer's disease diagnostics by a deeply supervised adaptable 3d convolutional network. *arXiv preprint arXiv:1607.00556*, 2016.
- [8] Tien Duong Vu et al. Multimodal learning using convolution neural network and sparse autoencoder. In *Big Data and Smart Computing (Big-Comp), 2017 IEEE International Conference on*, pages 309–312. IEEE, 2017.
- [9] A. Payan and G. Montana. Predicting alzheimer's disease: a neuroimaging study with 3d convolutional neural networks. *arXiv preprint arXiv:1502.02506*, 2015.
- [10] C. Billones et al. Demnet: A convolutional neural network for the detection of alzheimer's disease and mild cognitive impairment. In *Region 10 Conference (TENCON), 2016 IEEE*, pages 3724–3727. IEEE, 2016.
- [11] V. Planche et al. Hippocampal microstructural damage correlates with memory impairment in clinically isolated syndrome suggestive of mul-

- tiple sclerosis. *Multiple Sclerosis Journal*, page 1352458516675750, 2016.
- [12] Olfa Ben Ahmed et al. Recognition of alzheimer's disease and mild cognitive impairment with multimodal image-derived biomarkers and multiple kernel learning. *Neurocomputing*, 220:98–110, 2017.
- [13] Olfa Ben Ahmed et al. Alzheimer's disease diagnosis on structural mr images using circular harmonic functions descriptors on hippocampus and posterior cingulate cortex. *Computerized Medical Imaging and Graphics*, 44:13–25, 2015.
- [14] Olfa Ben Ahmed et al. Classification of alzheimer's disease subjects from mri using hippocampal visual features. *Multimedia Tools and Applications*, 74(4):1249–1266, 2015.
- [15] David G Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pages 1150–1157. Ieee, 1999.
- [16] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. *Computer vision—ECCV 2006*, pages 404–417, 2006.
- [17] A. Ebadi et al. Ensemble classification of alzheimer's disease and mild cognitive impairment based on complex graph measures from diffusion tensor images. *Frontiers in Neuroscience*, 11, 2017.
- [18] W. Lee, B. Park, and K. Han. Svm-based classification of diffusion tensor imaging data for diagnosing alzheimer's disease and mild cognitive impairment. In *International Conference on Intelligent Computing*, pages 489–499. Springer, 2015.
- [19] B. Lei et al. Discriminative learning for alzheimer's disease diagnosis via canonical correlation analysis and multimodal fusion. *Frontiers in aging neuroscience*, 8, 2016.
- [20] K. Aderghal et al. Fuseme: Classification of smri images by fusion of deep cnns in 2d+eps projections. In *International Workshop on Content-Based Multimedia Indexing*, pages 1–7, 2017.
- [21] T. Glozman and O. Liba. Hidden cues: Deep learning for alzheimer's disease classification cs331b project final report.
- [22] A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [23] S. Sarraf and G. Tofghi. Classification of alzheimer's disease structural mri data by deep learning convolutional neural networks. *arXiv preprint arXiv:1607.06583*, 2016.
- [24] S. Sarraf and G. Tofghi. Deep learning-based pipeline to recognize alzheimer's disease using fmri data. *bioRxiv*, page 066910, 2016.
- [25] S. Sarraf et al. Deepad: Alzheimer's disease classification via deep convolutional neural networks using mri and fmri. *bioRxiv*, page 070441, 2016.
- [26] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [27] Jun Shi et al. Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of alzheimer's disease. *IEEE journal of biomedical and health informatics*, 2017.
- [28] S. Korolev et al. Residual and plain convolutional neural networks for 3d brain mri classification. *arXiv preprint: 1701.06643*, 2017.
- [29] Heung-Il Suk et al. Deep ensemble learning of sparse regression models for brain disease diagnosis. *Medical image analysis*, 37:101–113, 2017.

About the authors

Alexander Khvostikov is a PhD student, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, Russia. His contact e-mail is xubiker@gmail.com.

Jenny Benois-Pineau is a full professor of Computer Science at the University Bordeaux, Bordeaux, France. She is a head of video and indexing research group at LABRI UMR 5800. Her contact email is jenny.benois-pineau@u-bordeaux.fr.

Andrey Krylov is a professor, head of the Laboratory of Mathematical Methods of Image Processing, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, Russia. His contact e-mail is kryl@cs.msu.ru.

Gwenaelle Catheline is an assistant professor at the University Victor Segalen Bordeaux 2, Bordeaux, France. She is a researcher at INCIA UMR 5287. Her contact email is gwenaelle.catheline@u-bordeaux.fr.