q-Space Deep Learning for Alzheimer's Disease Diagnosis: *Global Prediction and Weakly-Supervised Localization*

Vladimir Golkov[1] Phillip Swazinna[1] Marcel M. Schmitt[1] Qadeer A. Khan[1] Chantal M.W. Tax[2] Marat Serahlazau[1] Francesco Pasa[1,3] Franz Pfeiffer[3] Geert Jan Biessels[4] Alexander Leemans[5] Daniel Cremers[1]

[1] Computer Vision Group, Technical University of Munich, Germany [2] CUBRIC, Cardiff University, UK [3] Physics Department, Technical University of Munich, Germany

[4] Department of Neurology, University Medical Center Utrecht, Netherlands [5] Image Sciences Institute, University Medical Center Utrecht, Netherlands

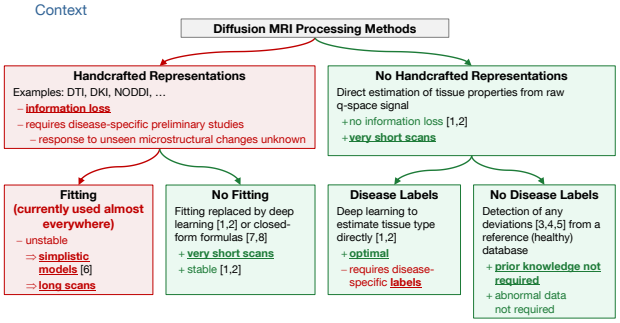
Goals

- q-Space deep learning [1,2]:
  - Prediction of tissue properties directly from q-space data
  - Usually voxel-wise
- Here:
  - AD/healthy label for entire scan (not voxel-wise)
  - ConvNet with global prediction
  - Weakly-supervised learning: voxel-wise reasons for global decision

Overfitting of Intensities → Data Augmentation 1

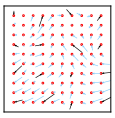
- Prevent overfitting of image intensity values:
  - Divide each scan by its mean intensity
  - Then multiply each scan by a random number between 0.5 and 1.5 (different in each training iteration)

Context

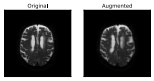


Overfitting of Patches → Data Augmentation 2

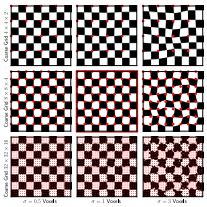
- Few scans, many voxels → danger of overfitting irrelevant image patterns
- Random elastic spatial deformations → more variablility of irrelevant patterns, less overfitting



Random and interpolated deformation vectors



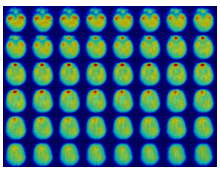
Before and after random 3D deformation



Different deformation grid densities and vector standard deviations

Overfitting of Distortion Artifacts → Cropping

- Reliable classification, but network attends to distorted regions (probably each scan has unique "overfittable" features there)



➢ Solution: Crop training images

➢ This reveals the next "overfitting problem"

Weakly-Supervised Learning Methods

- One label for every **scan**, one prediction for every **voxel**

Global supervised learning:

- One label for every **scan**, one prediction for every **scan**

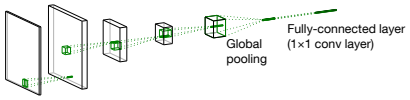
Local supervised learning:

- One label for every **voxel**, one prediction for every **voxel**

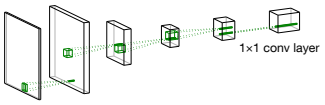
- Global label: Alzheimer's disease
- Local prediction: which areas of given image influence classifier the most

Class Activation Mapping (CAM) [12,13]

- ConvNets (illustration: 2D) often have a global pooling layer
- Train with global pooling (→ global prediction)



- Test without global pooling (→ local prediction)



- Fully-connected layers (if any) can be considered as 1×1 convolutional layers
- Result: local prediction of a network that was trained for global prediction
- Fewer fully-connected layers → more well-founded

Guided Backpropagation [14]

- Which voxel intensities should be changed how to strongly influence the prediction

Data [9,10]

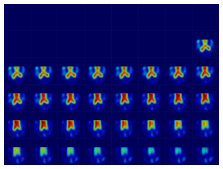
- 47 AD patients, 58 healthy controls, 5-fold cross-validation
- One b=0 image (averaged over 3 repetitions), 45 diffusion directions (b=1200s/mm²)
- Single-shot SE-EPI, TR=6638ms, TE=73ms, voxels 1.72mm×1.72mm×2.5mm, matrix 128×128, 48 axial slices, motion/distortion-corrected using ExploreDTI [11]

Network and Training

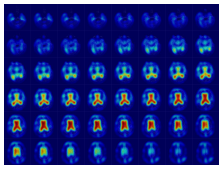
- 3D ConvNet: C128-P-C256-P-C512-GP-FC2000-FC1
  - C*n* – 3D convolutional layer with *n* 3×3×3 filters
  - P – 3D 2×2×2 max-pooling
  - GP – 3D global pooling
  - FC*n* – fully-connected layer with *n* neurons
- Hidden layers: ReLU nonlinearity, output: sigmoid
- Binary cross-entropy loss (target: AD/healthy label), Adam optimizer [15], learning rate $2 \cdot 10^{-5}$

"Over"fitting of Macrostructure & Open Questions

- Network attends to ventricles (enlarged in AD, easy macrostructural features) instead of q-space (microstructural features)

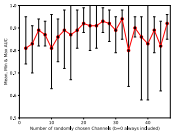


Trained and tested on cropped image



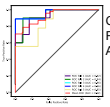
Trained on cropped, tested on full image

- This is further supported by the fact that just the *b*=0 image yields good results



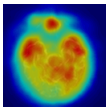
- Open question: smaller receptive fields to use only q-space information, but for global prediction

Good Classification, Relevant Brain Regions



Cross-validation ROC, AUC=0.92

CAM attends to AD-relevant brain regions



Discussion & Conclusions

- q-Space deep learning for global (scan-wise) prediction yields good results
- No handcrafted representations (such as DTI or NODDI)
- Diagnosis directly from raw q-space measurements, data-driven
- ConvNet with large receptive field uses macrostructural features (image space, ventricles) more than microstructural features (q-space in brain voxels) for Alzheimer's disease classification
- Open question: study q-space information – small (e.g. minimal [1]) receptive field combined with scan-wise labels

References

[1] V. Golkov et al. *q-Space Deep Learning for Twelve-Fold Shorter and Model-Free Diffusion MRI Scans.* MICCAI 2015

[2] V. Golkov et al. *q-Space Deep Learning: Twelve-Fold Shorter and Model-Free Diffusion MRI Scans.* IEEE TMI 2016

[3] V. Golkov et al. *Model-Free Novelty-Based Diffusion MRI.* ISBI 2016

[4] V. Golkov et al. *q-Space Novelty Detection in Short Diffusion MRI Scans of Multiple Sclerosis.* ISMRM 2018 (#5378)

[5] A. Vasilev et al. *q-Space Novelty Detection with Variational Autoencoders.* arXiv:1806.02997 2018

[6] U. Ferizi et al. *White matter compartment models for in vivo diffusion MRI at 300 mT/m.* NeuroImage 2015

[7] B. Hansen et al. *Experimentally and computationally fast method for estimation of a mean kurtosis.* MRM 2013

[8] B. Hansen et al. *Fast imaging of mean, axial and radial diffusion kurtosis.* NeuroImage 2016

[9] Y.D. Reijmer, A. Leemans, K. Caeyenberghs, et al. *Disruption of cerebral networks and cognitive impairment in Alzheimer disease.* Neurology 2013

[10] M. Bach, F.B. Laun, A. Leemans, et al. *Methodological considerations on tract-based spatial statistics (TBSS).* NeuroImage 2014

[11] A. Leemans, B. Jeurissen, J. Sijbers, D.K. Jones. *ExploreDTI: a graphical toolbox for processing, analyzing, and visualizing diffusion MR data.* ISMRM 2009

[12] B. Zhou et al. *Learning Deep Features for Discriminative Localization.* CVPR 2016

[13] R.R. Selvaraju et al. *Grad-CAM: visual explanations from deep networks via gradient-based localization.* ICCV 2017

[14] J.T. Springenberg, A. Dosovitskiy, T. Brox, M. Riedmiller. *Striving for simplicity: the all convolutional net.* ICLR Workshops 2015

[15] D. Kingma, J. Ba. *Adam: a method for stochastic optimization.* ICLR 2015

Acknowledgments: This study was partly funded by the European Research Council, Deutsche Telekom Foundation, German Research Council (DFG, LA 2804/1-1), and the contribution of Chantal Tax is supported by an FC-EW grant (No. 612.001.104) from the Dutch Scientific Foundation (NWO). The authors would like to thank the members of the Utrecht Vascular Cognitive Impairment Study Group for providing the diffusion MRI data (in alphabetical order by department): University Medical Center Utrecht, the Netherlands, Department of Neurology: E. van den Berg, G.J. Biessels, M. Brundel, W.H. Bouvy, S.M. Heringa, L.J. Kappelle, Y.D. Reijmer; Department of Radiology/Image Sciences Institute: J. de Bresser, H.J.Kuijf, A. Leemans, P.R. Luijten, W.P.Th.M. Mali, M.A. Viergever, K.L. Vincken, J.J.M. Zwanenburg; Department of Geriatrics: H.L. Koek, J.E. de Wit; Hospital Diakonessenhuis Zeist, the Netherlands: M. Hamaker, R. Faaij, M. Pleizier, E. Vriens.