

5378

## q-Space Novelty Detection in Short Diffusion MRI Scans of Multiple Sclerosis

Vladimir Golkov<sup>1</sup>, Aleksei Vasilev<sup>1</sup>, Francesco Pasa<sup>1,2</sup>, Ilona Lipp<sup>3</sup>, Wassim Boubaker<sup>1</sup>, Eleonora Sgarlata<sup>3,4</sup>, Franz Pfeiffer<sup>2</sup>, Valentina Tomassini<sup>3,5</sup>, Derek K. Jones<sup>3</sup>, and Daniel Cremers<sup>1</sup>

<sup>1</sup>Department of Informatics, Technical University of Munich, Munich, Germany, <sup>2</sup>Physics Department, Technical University of Munich, Munich, Germany, <sup>3</sup>CUBRIC, Cardiff University, Cardiff, United Kingdom, <sup>4</sup>Department of Neurology and Psychiatry, Sapienza University of Rome, Rome, Italy, <sup>5</sup>Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, United Kingdom

### Synopsis

**Diffusion MRI can capture disease-related microstructural changes, but most methods use handcrafted data transformations that discard parts of the information and require quite long scan times. In contrast, q-space novelty detection (q-ND) circumvents these drawbacks, and does not require any knowledge whatsoever about the effect of disease on q-space measurements. Instead, q-ND highlights voxels that look unlike anything seen in a database of healthy scans. Here we show that novelty scores from q-ND largely coincide with multiple sclerosis lesions, and that q-ND also works at reduced scan times.**

### Introduction

So far, establishing the relationship between q-space measurements and microstructure has required mathematical/physical representations which are suboptimally simplistic (discarding parts of the information) and rely on unstable fitting procedures requiring quite long scan times. Moreover, some kind of prior knowledge about disease-related effects is required. In contrast, handcrafted-representations-free methods<sup>1,2</sup> circumvent most of these drawbacks. In particular, q-space novelty detection<sup>2</sup> (q-ND) does not even require any knowledge whatsoever about the effect of disease on q-space measurements. Instead, it highlights voxels that look unlike anything seen in a database of healthy scans. Here we apply q-ND to scans of multiple sclerosis (MS) patients. We compare voxel-wise novelty scores from q-ND with lesion labels obtained without diffusion MRI, and for the first time study the effect of scan time reduction on q-ND results.

### Methods

The data were as follows:<sup>3</sup> As the the healthy database to compare the patient datasets with, we used 26 diffusion MRI scans of healthy volunteers: six  $b=0$  images, 40 uniformly-distributed diffusion directions ( $b_{\max}=1200\text{s/mm}^2$ , SE-EPI, voxel size  $1.8\text{mm}\times 1.8\text{mm}\times 2.4\text{mm}$ , matrix  $128\times 128$ , 57 slices,  $TE=94.5\text{ms}$ ,  $TR=16\text{s}$ , motion/distortion-corrected,<sup>4</sup> upsampled to  $256\times 256\times 172$ ). Using machine-learning nomenclature, we refer to these 46 volumes (six  $b=0$ , 40 diffusion-weighted) as “channels”. As test data, we used scans of five MS patients (scan parameters as above). Lesions were marked by human raters in additional structural  $T_2$ -weighted scans for comparison with q-ND results. In order to avoid potential discrepancies in image intensity across scans, we divided each scan by its mean intensity. Additionally, in order to prevent some channels from dominating over others in novelty detection, we divided each channel by the channel-wise mean taken across all scans (a technique known as “feature scaling”).

We consider every voxel as a data sample. In the standard setting, the feature vector consists of the 46 q-space measurements. In additional evaluations, we used small random subsets of the channels in order to study the effect of reducing the scan time. This results in an  $n$ -dimensional feature space, where  $n\leq 46$  is the size of the chosen set of channels. Selected channels are the same for both reference and test data. For every test data sample, we define the novelty score in arbitrary units as the distance between its feature vector and the closest, in terms of Euclidean distance in feature space, feature vector corresponding to a voxel from the healthy database. Consequently, the more dissimilar the test point is to healthy voxels, the higher its novelty score. This translates to the following formula for the novelty score:  $r(x)=\min_{y\in Y}\|x-y\|$ , where  $Y$  is set of feature vectors  $y$  corresponding to healthy voxels in the  $n$ -dimensional feature space, and  $\|\cdot\|$  is the Euclidean norm. We took only brain voxels into account in the healthy database and in the tested images. To speed up q-ND runtime, we reduced the reference dataset size by randomly sampling 1/6 of brain voxels from each scan.

### Results

Fig. 1 shows the receiver operating characteristic (ROC) for MS lesion segmentation across all test voxels using all 46 channels. The area under the curve (AUC) is 0.888, indicating a promising performance. The per-patient AUC ranged from 0.599 (for a patient with particularly small extent of lesions) to 0.929. Fig. 2 shows the normalized histogram of novelty score outputs on the test set for 46 channels. Figures. 3 and 4 show how reducing the number of sampled channels (and thus reducing the scan time) affects the quality of the results.

### Discussion

These results show that high novelty scores largely coincide with MS lesions even with a small number of channels and thus short scan duration. Moreover, certain random subsets of channels (containing useful and complementary information) yield better results than others. Some imperfections in the predictions may be due to imperfections of the method, or imperfections of image registration; on the other hand, some *apparent* imperfections of predictions might be rather due to imperfections of the lesion labels used for reference.

### Conclusions

Our experiments show that by using q-space novelty detection it is possible to directly segment multiple sclerosis lesions from raw diffusion MRI data without fitting handcrafted representations and without requiring hand-marked labels. (Note that here lesion labels were used here merely for validation.) The approach that we use shows promising segmentation even at short scan times (Figs. 3-4). Moreover, we do not need to acquire labels which can be imperfect. To summarize: the method marks any q-space signal as a deviation from healthy data if this signal is affected by the disease in a measurable way.<sup>2</sup> And this happens without any representations-based assumptions or prior knowledge of the disease effect on the q-space signal. Other novelty detection methods<sup>5</sup> could be tried.

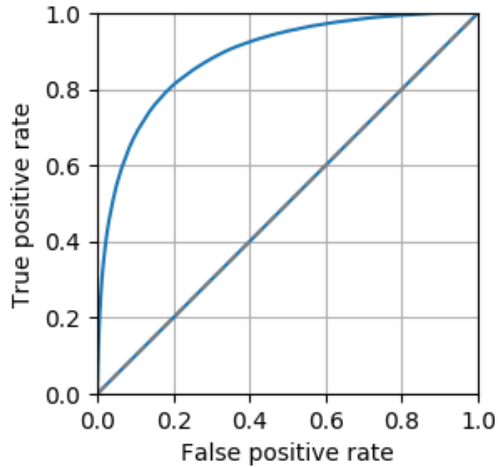
## Acknowledgements

This work was supported by the ERC Consolidator Grant "3DReloaded".

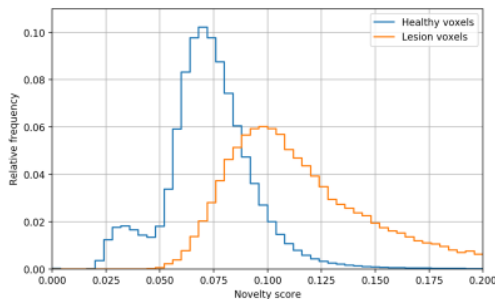
## References

1. V. Golkov et al. q-Space deep learning: twelve-fold shorter and model-free diffusion MRI scans. *IEEE Trans Med Imag.* 2016;35(5):1344–1351.
2. V. Golkov et al. Model-free novelty-based diffusion MRI. *Proc IEEE Int Symp Biomed Imag (ISBI).* 2016:1233–1236.9
3. I. Lipp et al. Comparing MRI measures of myelin changes in multiple sclerosis. *Org Human Brain Mapping (OHBM) Annual Meeting.* 2017:3060.
4. S. Klein et al. elastix: A toolbox for intensity-based medical image registration. *IEEE Trans Med Imag.* 2010;29(1):196–205.
5. M.A. Pimentel, et al. A review of novelty detection. *Signal Processing.* 2014;99:215–249.

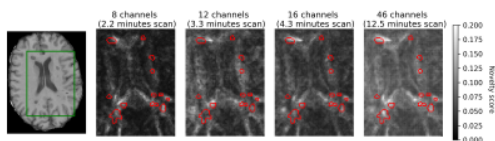
## Figures



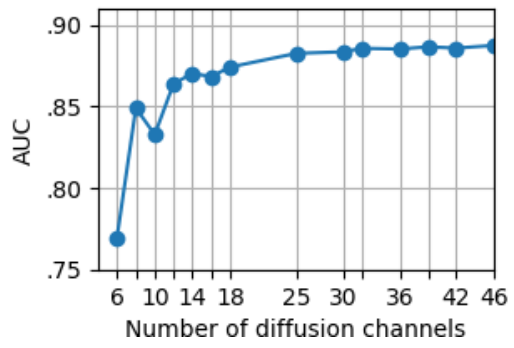
ROC measuring how much q-ND novelty scores (which use only knowledge about healthy q-space data, but no knowledge whatsoever about lesions) coincide with hand-marked lesions, AUC=0.888.



Normalized histograms of the novelty score for lesion voxels, and the novelty score for non-lesion voxels. The novelty score allows quite good separation of lesions from non-lesions.



Novelty scores from q-DL obtained using different numbers of channels (different scan times). Independently obtained labels are outlined in red. High novelty scores strongly coincide with lesion labels. Besides, novelty scores seem to indicate additional q-space abnormalities beyond the labelled lesions.



AUC for lesion segmentation with different numbers of channels (different scan times).

Proc. Intl. Soc. Mag. Reson. Med. 26 (2018)  
5378



# q-Space Novelty Detection

## *in Short Diffusion MRI Scans of Multiple Sclerosis*

Vladimir Golkov<sup>1</sup> Aleksei Vasilev<sup>1</sup> Francesco Pasa<sup>1,2</sup> Ilona Lipp<sup>3</sup> Wassim Boubaker<sup>1</sup>  
Eleonora Sgarlata<sup>3,4</sup> Franz Pfeiffer<sup>2</sup> Valentina Tomassini<sup>2,5</sup> Derek K. Jones<sup>3</sup> Daniel Cremers<sup>1</sup>

<sup>1</sup> Department of Informatics, Technical University of Munich, Germany

<sup>2</sup> Physics Department, Technical University of Munich, Germany

<sup>3</sup> CUBRIC, Cardiff University, Cardiff, UK

<sup>4</sup> Department of Neurology and Psychiatry, Sapienza University of Rome, Italy

<sup>5</sup> Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, UK



## Diffusion MRI Processing Methods

### Handcrafted Representations

Examples: DTI, DKI, NODDI, ...

- **information loss**
- requires disease-specific preliminary studies
  - response to unseen microstructural changes unknown

### No Handcrafted Representations

Direct estimation of tissue properties from raw q-space signal

- + no information loss [1,2]
- + **very short scans**

### Fitting

**(currently used almost everywhere)**

- unstable
- ⇒ **simplistic models** [5]
- ⇒ **long scans**

### No Fitting

Fitting replaced by deep learning [1,2] or closed-form formulas [6,7]

- + **very short scans**
- + stable [1,2]

### Supervised Learning

Deep learning to estimate tissue type directly [1,2,3]

- + **optimal**
- requires disease-specific **labels**

### Novelty Detection (this work)

Detection of any deviations [4] from a reference (healthy) database

- + **prior knowledge not required**
- + **abnormal data not required**

[1,2,3] Golkov et al., MICCAI 2015, IEEE TMI 2016, ISMRM 2018 (#1580)

[4] Golkov et al., ISBI 2016

[5] Ferizi et al., NeuroImage 2015

[6,7] Hansen et al., MRM 2013, NeuroImage 2016



# Methods

## q-Space Novelty Detection

- q-Space measurements  $x$  in a voxel are the “feature vector” for that voxel
- Given: Database  $Y$  of voxels  $y$  from healthy controls
- Novelty detection: How dissimilar is patient voxel  $x$  from the healthy voxels?
- Novelty score based on distance in feature space:  $r(x) = \min_{y \in Y} \|x - y\|$

# q-Space Novelty Detection

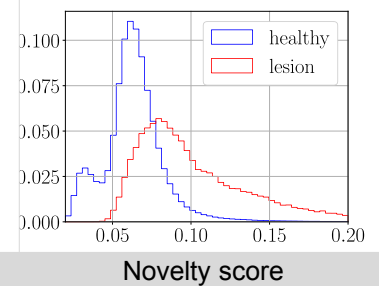
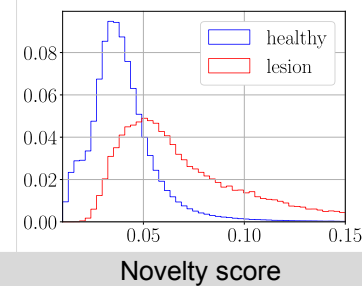
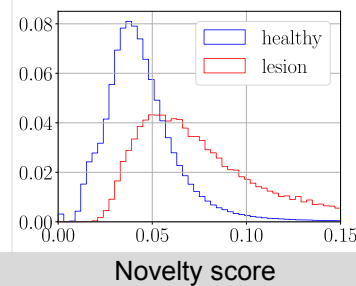
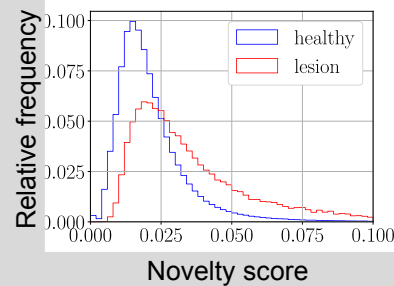
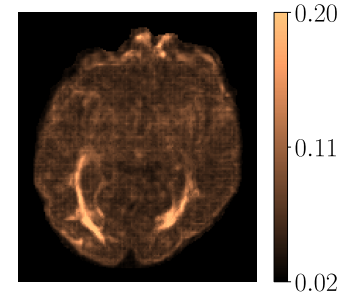
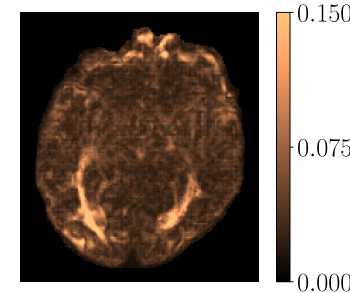
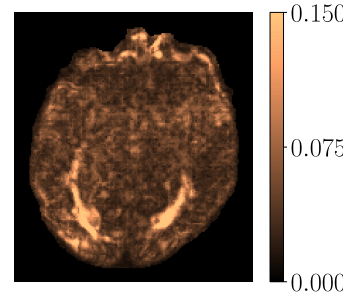
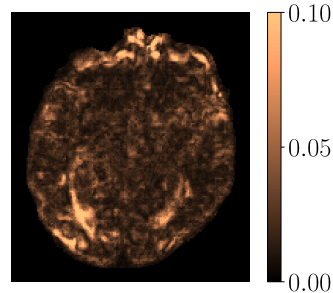
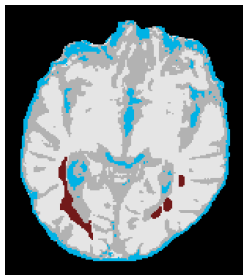
Manual  
lesion labels

8 q-samples  
(2 minutes scan)

12 q-samples  
(3 minutes scan)

16 q-samples  
(4 minutes scan)

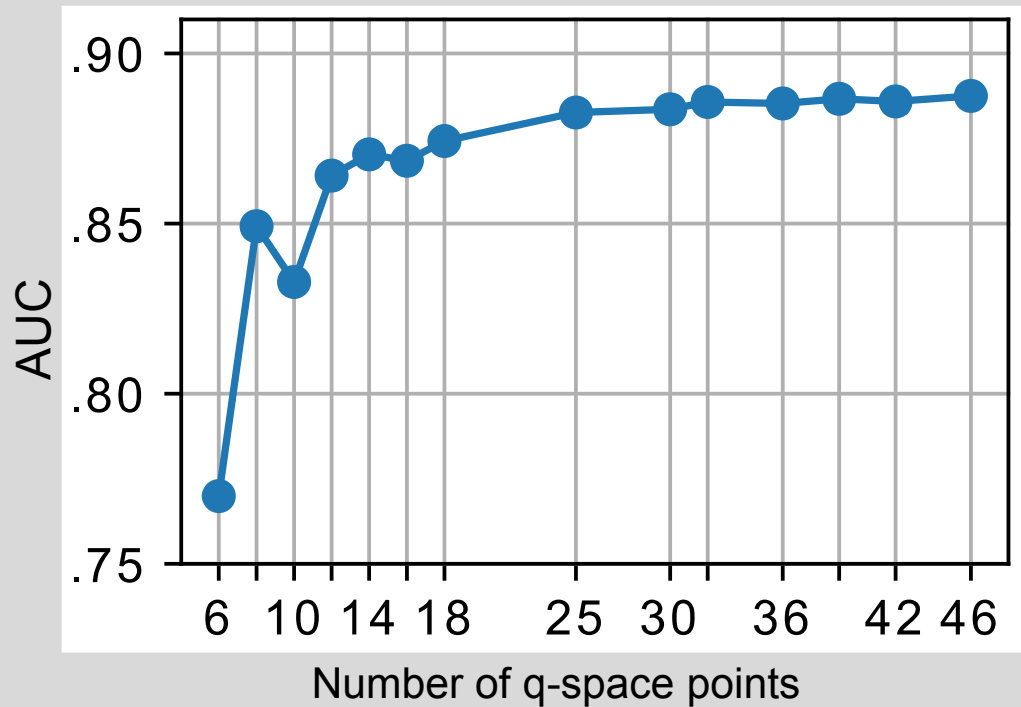
46 q-samples  
(13 minutes scan)



- Novelty scores coincide with MS lesions
- Only healthy data required for training – no lesion data and no lesion labels
- Direct usage of raw q-space data, no drawbacks of handcrafted representations



# Fast Diffusion MRI Scans



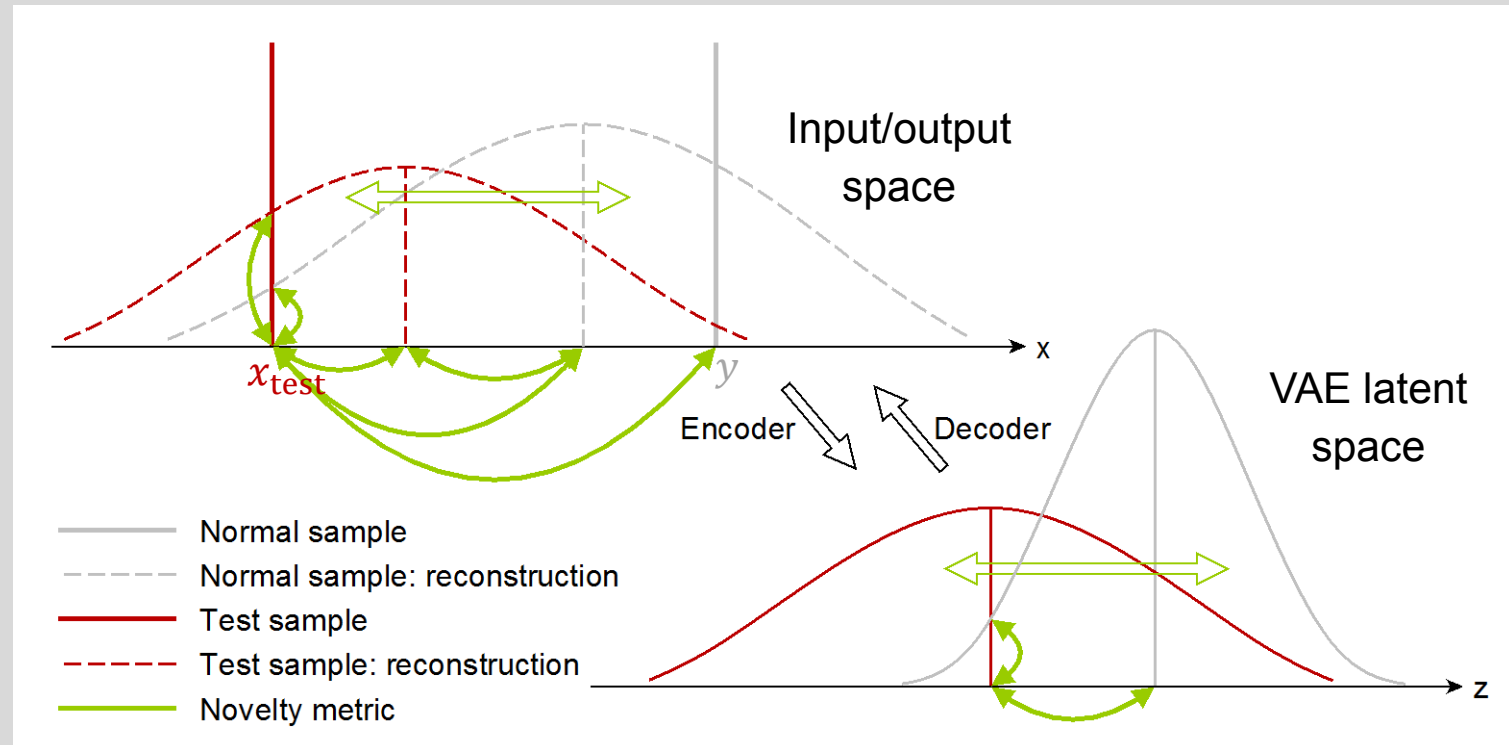
- Much information about tissue is contained in few q-space points



# Deep Novelty Detection

- Variational autoencoder (VAE) trained on voxels of healthy volunteers
  - Cannot reconstruct abnormal voxels well (novelty score := reconstruction error)
  - Abnormal different from normal in input/latent/output spaces

[Vasilev et al. “q-Space novelty detection with variational autoencoders” submitted to NIPS’18, arXiv’18]



# q-Space Novelty Detection: Summary

## Avoid the usual problems

Technical University of Munich  
Cardiff University  
Sapienza University of Rome

q-Space Novelty Detection | Golkov et al. | #5378

**Diffusion MRI Processing Methods**

**Handcrafted Representations**  
Examples: DTI, DKI, NODDI, ...

- Information loss
- requires disease-specific preliminary studies
- response to unseen microstructural changes unknown

**No Handcrafted Representations**  
Direct estimation of tissue properties from raw q-space signal

- + no information loss [1,2]
- + very short scans

**Fitting (currently used almost everywhere)**

- unstable
- ⇒ simplistic models [5]
- ⇒ long scans

**No Fitting**  
Fitting replaced by deep learning [1,2] or closed-form formulas [6,7]

- + very short scans
- + stable [1,2]

**Supervised Learning**  
Deep learning to estimate tissue type directly [1,2,3]

- + optimal
- requires disease-specific labels

**Novelty Detection (this work)**  
Detection of any deviations [4] from a reference (healthy) database

- + prior knowledge not required
- + abnormal data not required

[1,2,3] Golkov et al., MICCAI 2015, IEEE TMI 2016, ISMRM 2018 (#1560)  
[4] Golkov et al., ISBI 2016  
[5] Fennel et al., NeuroImage 2015  
[6,7] Hansen et al., MRSA 2013, ISMRM 2014

## Use raw q-space data directly

Technical University of Munich  
Cardiff University  
Sapienza University of Rome

q-Space Novelty Detection | Golkov et al. | #5378

**Methods**

q-Space Novelty Detection

- q-Space measurements  $x$  in a voxel are the "feature vector" for that voxel
- Given: Database  $Y$  of voxels  $y$  from healthy controls
- Novelty detection: How dissimilar is patient voxel  $x$  from the healthy voxels?
- Novelty score based on distance in feature space:  $r(x) = \min_y \|x - y\|$

## Detect any q-space oddity

Technical University of Munich  
Cardiff University  
Sapienza University of Rome

q-Space Novelty Detection | Golkov et al. | #5378

**q-Space Novelty Detection**

Manual lesion labels	8 q-samples (2 minutes scan)	12 q-samples (3 minutes scan)	16 q-samples (4 minutes scan)	48 q-samples (13 minutes scan)

- Novelty scores coincide with MS lesions
- Only healthy data required for training – no lesion data and no lesion labels
- Direct usage of raw q-space data, no drawbacks of handcrafted representations

## Fast scans

Technical University of Munich  
Cardiff University  
Sapienza University of Rome

q-Space Novelty Detection | Golkov et al. | #5378

**Fast Diffusion MRI Scans**

Number of q-space points	AUC
6	0.75
8	0.85
10	0.82
12	0.86
14	0.87
16	0.87
18	0.87
20	0.88
25	0.88
30	0.88
36	0.88
42	0.88
46	0.88

- Much information about tissue is contained in few q-space points

## Deep learning

Technical University of Munich  
Cardiff University  
Sapienza University of Rome

q-Space Novelty Detection | Golkov et al. | #5378

**Deep Novelty Detection**

- Variational autoencoder (VAE) trained on voxels of healthy volunteers
- Cannot reconstruct abnormal voxels well (novelty score := reconstruction error)
- Abnormal different from normal in input/latent/output spaces

[Vesilev et al. "q-Space novelty detection with variational autoencoders" submitted to NIPS'18, arXiv'18]