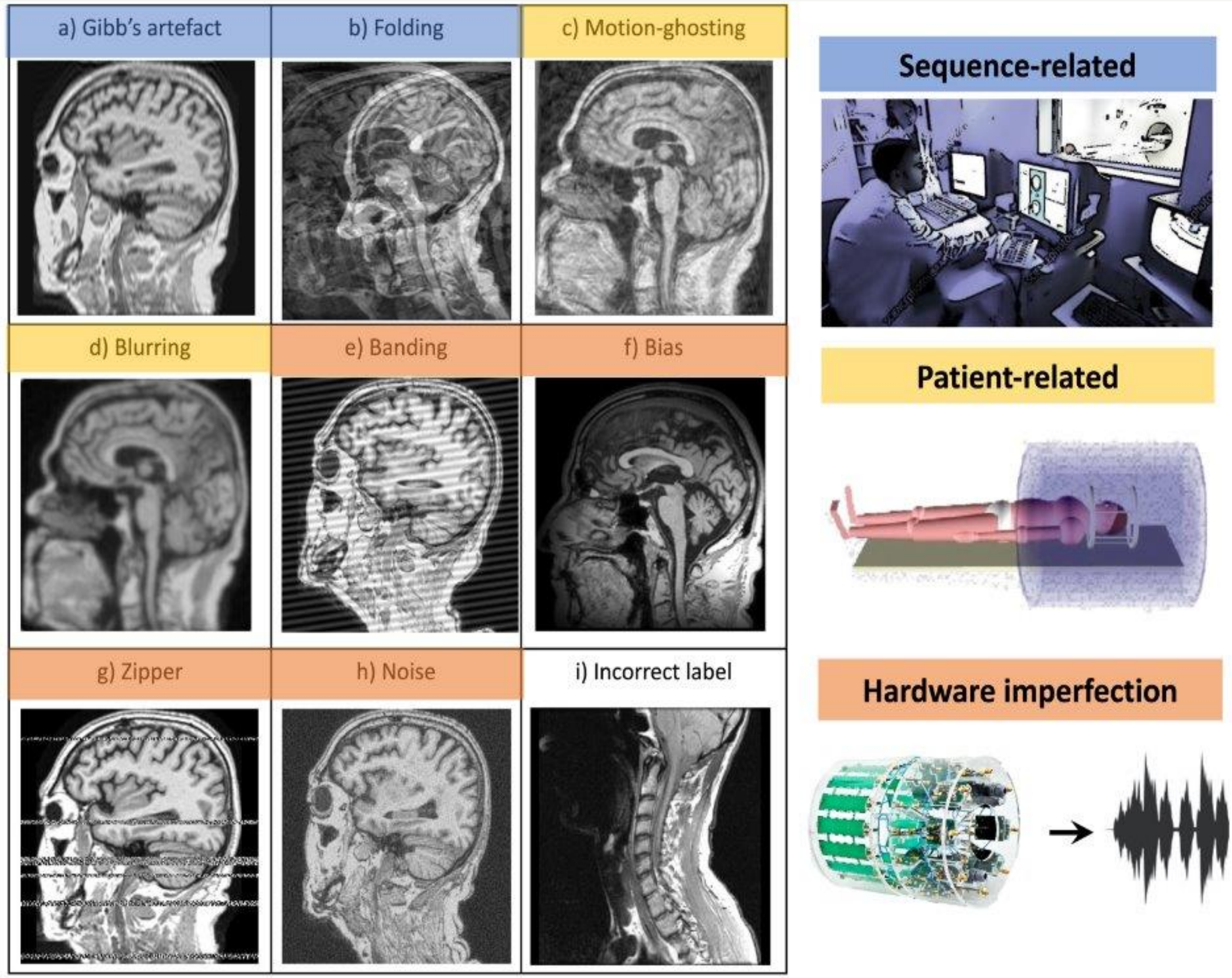
IMPROVING QUALITY CONTROL OF MRI IMAGES USING SYNTHETIC MOTION DATA

C Bricout, K Cho, M Harms, O Pasternak, C Bearden, PD McGorry, RS Kahn, JM Kane, B Nelson, SW Woods, ME Shenton, S Bouix, S Ebrahimi Kahou

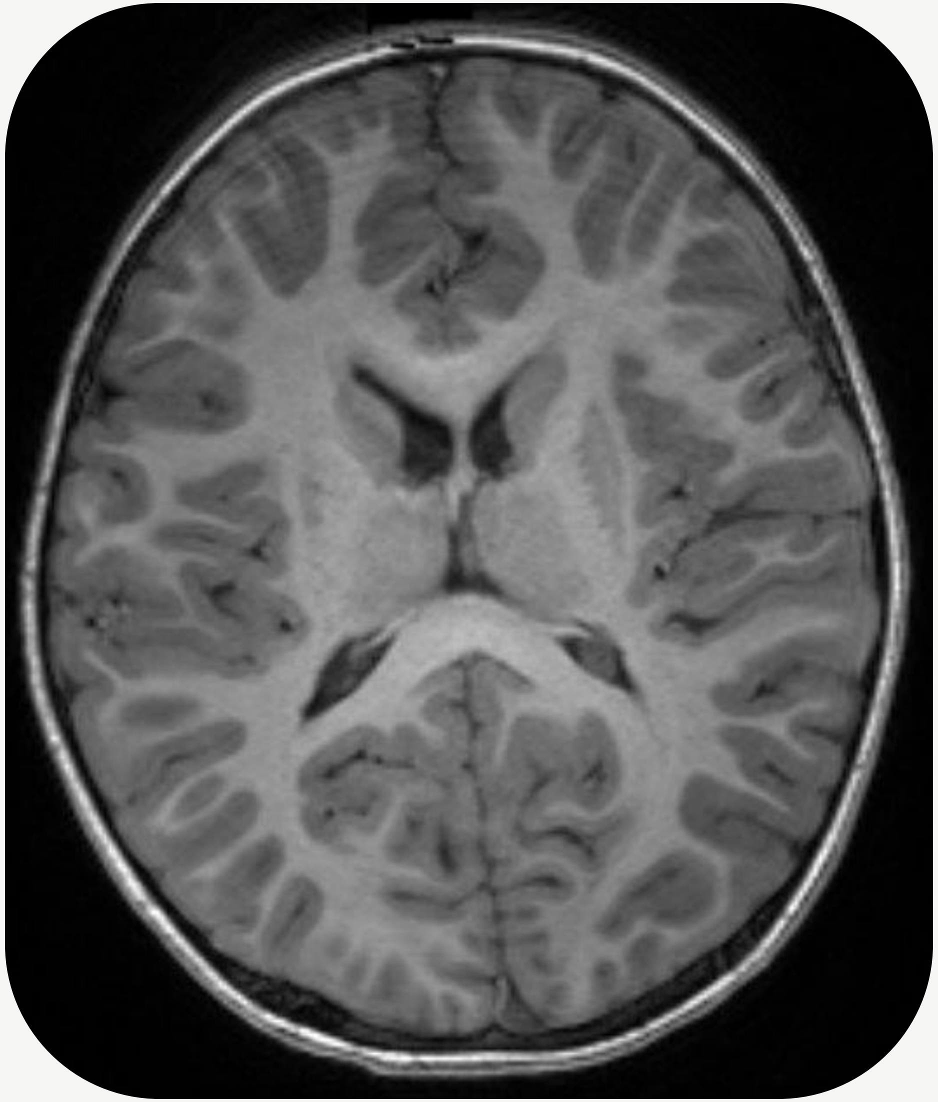
Brain MRI Artifacts

- Wide variety of artifacts affecting the **usability** of the data

- Automatic Detection becomes necessary as acquisition study scales up (100,000 subject in UK Biobank)



Motion Artifacts



- Extremely **common**
- Large range of intensity
- Multiple effects
- Impact anatomical measurement even at **levels undetectable by human eye** (cortical thickness)

Problem

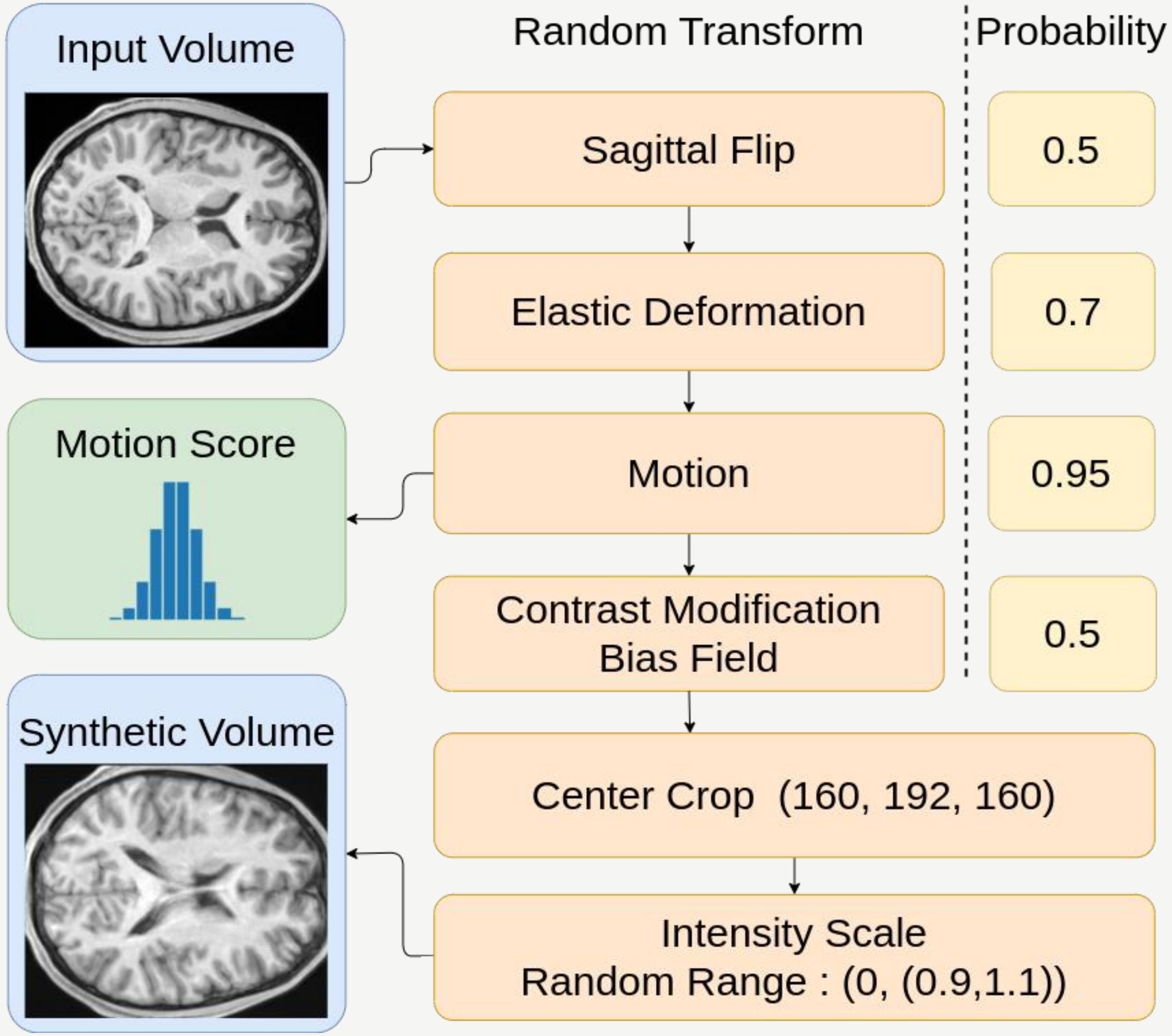
Training a network to detect / quantify motion artifacts is particularly **difficult** as :

- Most datasets **do not** include volumes with significant problems
- Datasets with **Quality Control** (QC) or **Motion** label are usually small and highly **unbalanced**
- QC / Motion is usually pass/fail or **subjective** grading (1-4)
- QC is only a proxy of motion
- Motion labels **do not necessarily relate** with physical motion
- Deep Learning in 3D is **computationally expensive**

Training a network to quantify motion objectively would allow for **statistical correction of anatomical measurements** and would be **an explainable proxy for QC**.

Proposed Method

- Learn a motion score using only **synthetically generated artifacts**[2].
- Quantify motion using a **physic-based metric**, the average root mean square deviation[1] between all affine matrices used for generation
- Assess relationship between motion quantification and QC scoring, by training **QC scoring network using the pretrained network's embeddings**



Data

Two datasets with a **4-point scale QC** :

- *Human Connectome Project Early Psychosis*[3] (*HCPEP*) with **390 volumes**
- *Accelerating Medicine Partnership Schizophrenia*[4] (*AMPSCZ*) with **1,048 volumes** over 44 acquisition sites. **26 sites** are reserved for synthetic data and **7 sites** for transfer learning

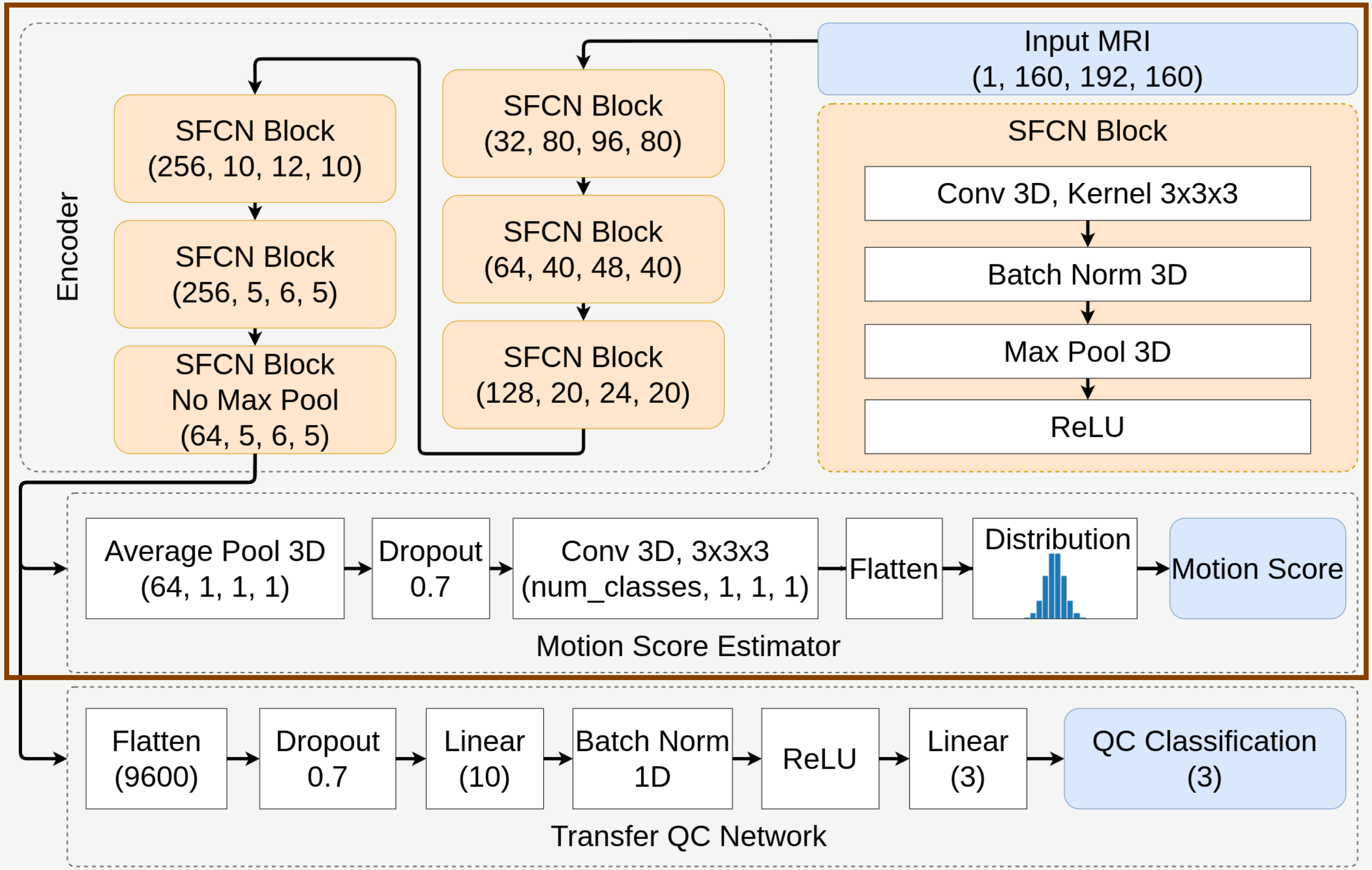
Keeping only the best quality data for the synthetic dataset, we get :

- **Synthetic Dataset** : 319 *AMPSCZ* and 143 *HCPEP* volumes
- **Transfer Dataset** : 378 *AMPSCZ volumes*

Models

Simple Fully Convolutional Network[5] (**SFCN**) is a **memory efficient** architecture introduced first for brain age estimation and used by Pollak et al. for motion estimation based on camera tracking.

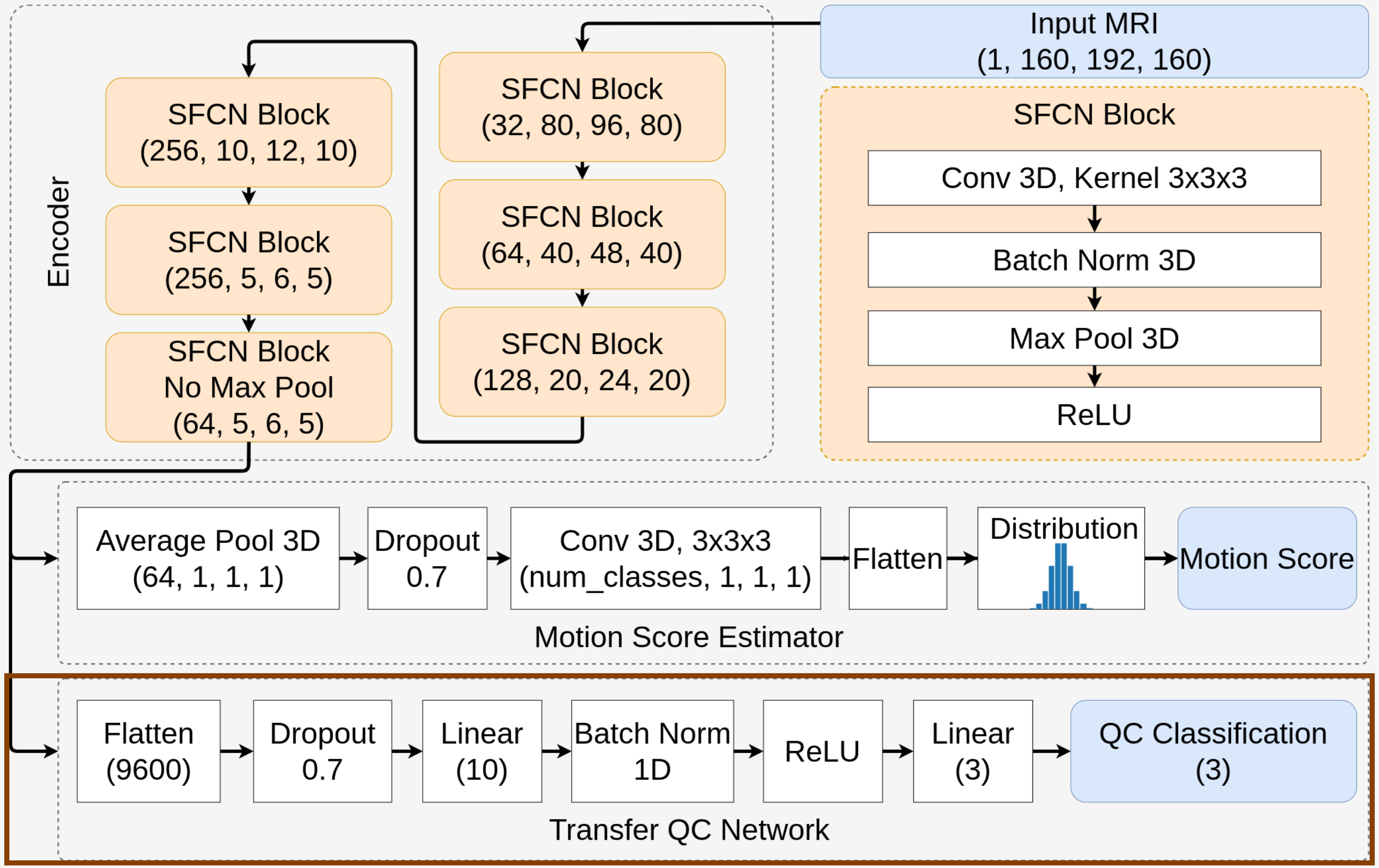
Outputs a **distribution** over 50 bins learned by optimizing **KL-Divergence**. This is easier than vanilla regression over a single neuron.



Models

Transfer QC Network is a simple multi-layer perceptron model mapping embeddings of the pretrained SFCN to a 3 classes output.

We use this instead of the original classifier to get more **functional expressivity** as we **completely froze** the weight of the encoder, pretrained on a different task.

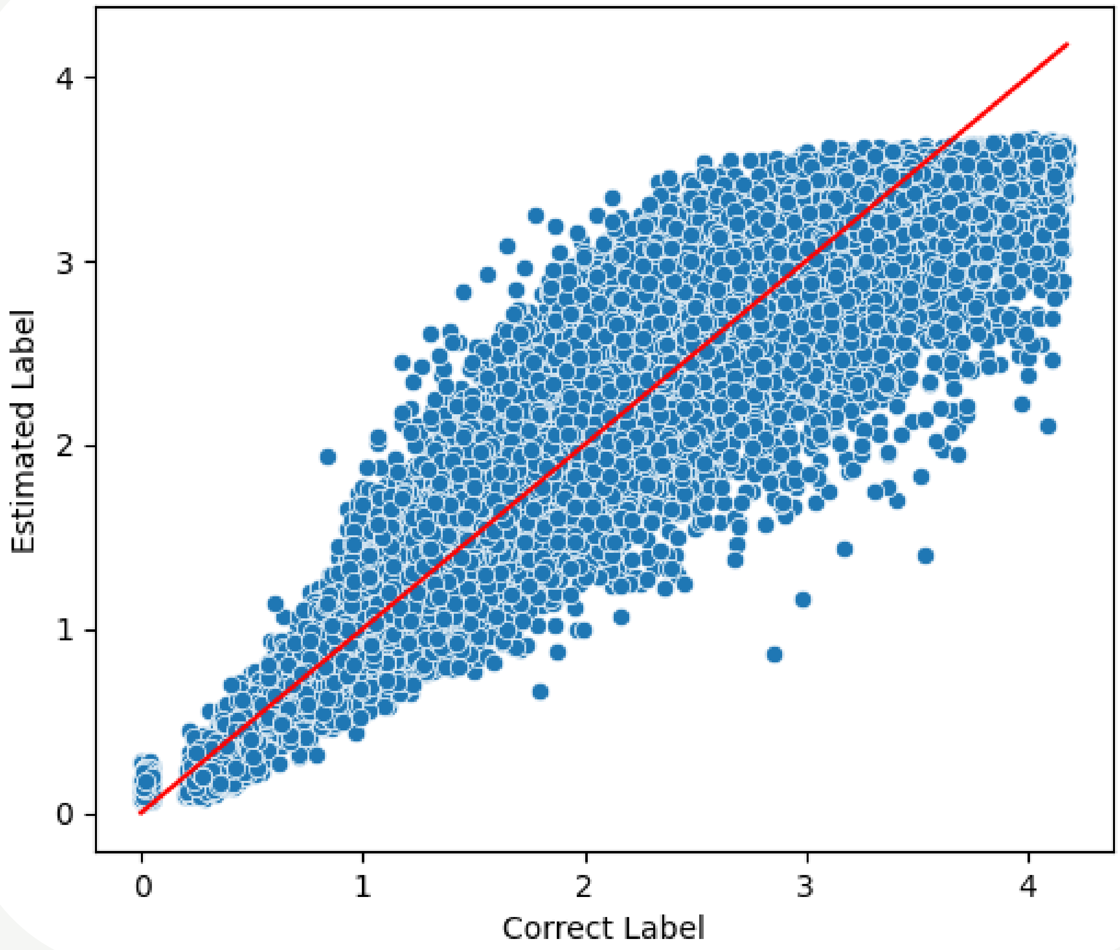


Pretrain Results

We produced 300 synthetic sample for each real volume, resulting in :

- 110,100 training volume
- 14,100 validation volume
- 13,801 testing volume

We obtain a **R^2 of 0.89** on test set



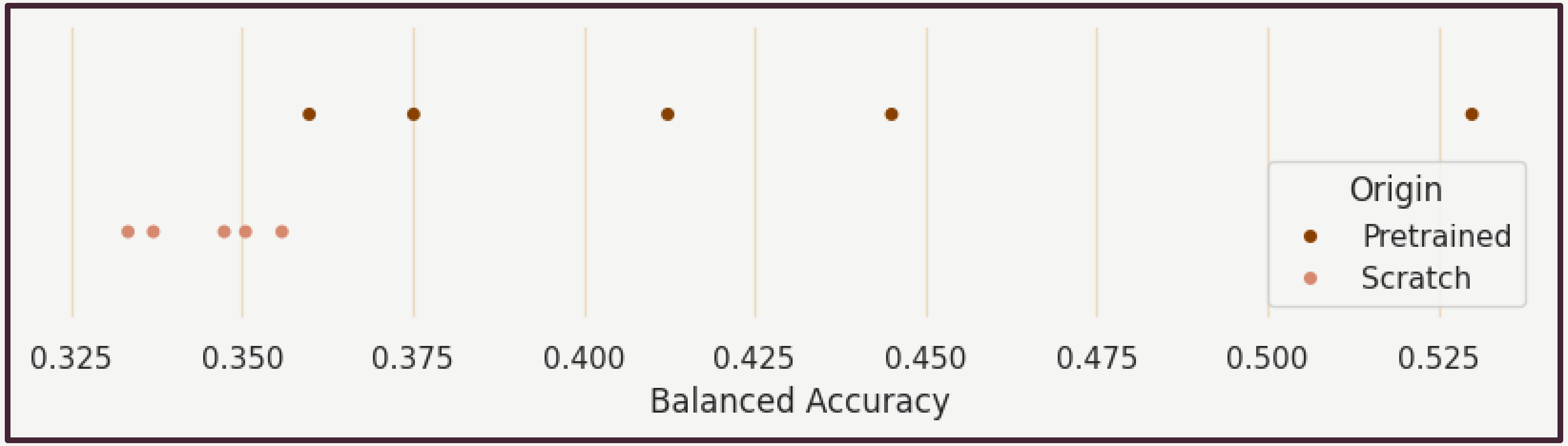
QC Task Results

We compare the results of two training setting on 5 random seeds:

- Training the SFCN from **Scratch**
- Training the **Transfer network** on the embedding of our pretrained model

Model	Setting	Balanced Accuracy	F1 Score By Class 1/2	3	4
SFCN	Pretrained	**0.41**	**0.13**	**0.34**	0.68
	Scratch	0.35	0.00	0.20	**0.70**

Median Results for Scratch and Transfer Learning



QC Score	1/2	3	4
Train	7	38	70
Test	1	12	26
Validation	9	90	125

Data Split for QC Task

Efficiency

	Max GPU Ram (GB)	Max GPU Power (Watt)	Duration (hh:mm:ss)
Pretrain	39.54	403.94	25:28:56
Transfer	2.50	76.79	00:03:02
Scratch	37.41	373.29	00:19:29
Decreased (%)	93.32%	79.43%	94.00%

Differences in Resource Usage Between Settings

- Using transfer learning on our pretrained network **does not require access to powerful GPUs** to be trained and takes significantly **less time to train**. Enabling even **small laboratory** to train their own.

- However, if we include the pretraining cost, it would necessitate **33 uses** for the overall approach to become more time-efficient.

Discussion

Our results show that :

- We can learn to **predict a motion score** from synthetically motion-corrupted MRI
- Models trained to quantify synthetic motion appear to **learn meaningful embeddings** that can be leveraged to perform QC classification on real volumes
- Training on synthetic data can help with **highly unbalanced datasets**
- Transfer learning is significantly more **resource efficient** (if the pretrained model is used multiple time)

Our best model has low performance, this could be improved by :

- Pretraining on multiple objectives reflecting different kinds of artifacts
- Data augmentation strategies

Conclusion

To the best of our knowledge, this is the first attempt to learn regression for subject motion with synthetically generated artifacts and to transfer this motion-specific knowledge to a more general QC classification task

More broadly, our research hints that :

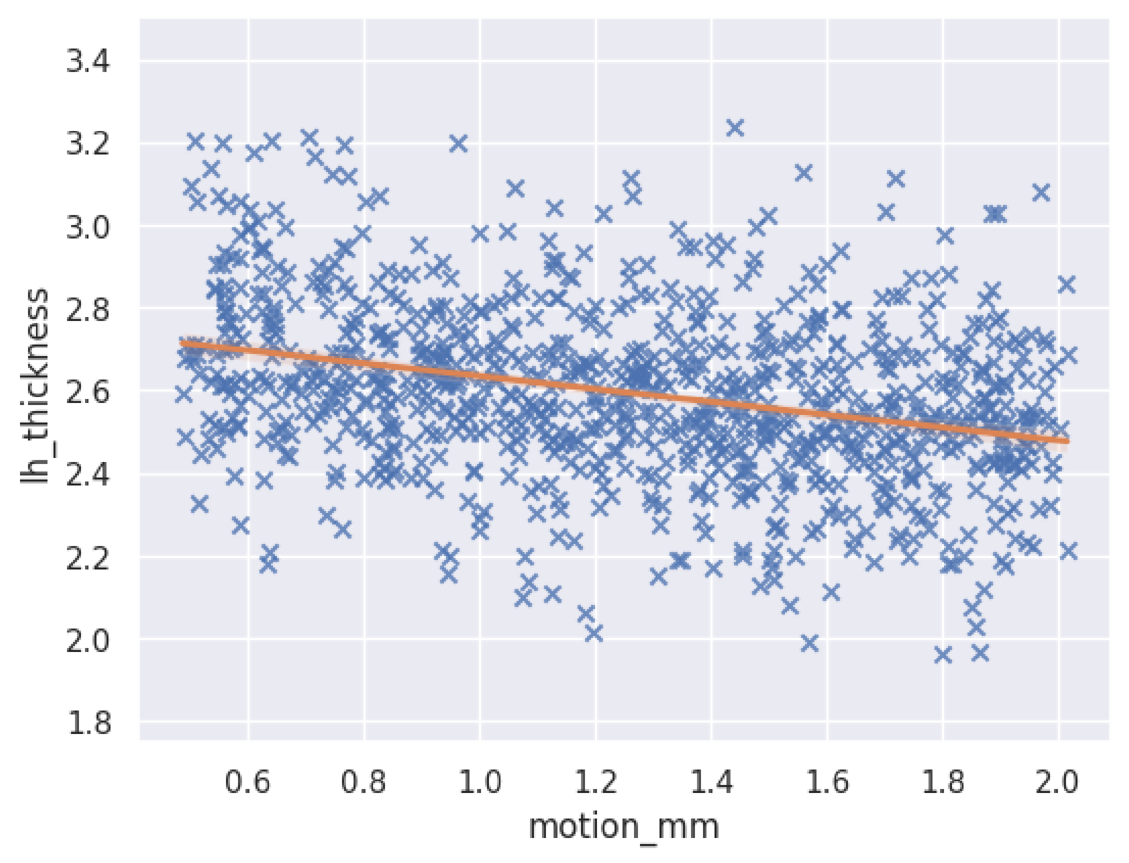
- Synthetic data can help with extreme data scarcity
- Pretraining a model on an **objective quantitative task** before fine tuning on a **subjective qualitative task** is a promising approach

Future Research – Statistical Correction

It has been shown that motion artifacts, even invisible ones, impact **automated anatomical measurements**.

We want to try to correct this statistical bias using our model. In doing so, we hope to :

- Provide **retrospective motion correction** for neurological research dataset
- Potentially **validate the knowledge learned on synthetic data** if our statistical correction display a similar behavior between synthetic and real volumes.



Early result on estimating a relationship between our predicted motion and cortical thickness

Bibliography

[1] M. Jenkinson, "Measuring Transformation Error by RMS Deviation," Technical TR99MJ1.
[2] F. Pérez-García et al., "TorchIO: A Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning," Computer Methods and Programs in Biomedicine, vol. 208, pp. 106236, Sept. 2021.
[3] G. R Jacobs et al., "An Introduction to the Human Connectome Project for Early Psychosis," Schizophrenia Bulletin, p. sbae123, July 2024.
[4] C. Wannan et al., "Accelerating Medicines Partnership® Schizophrenia (AMP® SCZ): Rationale and Study Design of the Largest Global Prospective Cohort Study of Clinical High Risk for Psychosis," Schizophrenia Bulletin, vol. 50, no. 3, pp. 496–512, May 2024.
[5] C. Pollak et al., "Estimating Head Motion from MRImages," Feb. 2023, arXiv:2302.14490 [cs, eess].