Deep convolutional neural networks for detection of cortical dysplasia: a multicenter validation

Gill RS [1*], Hong SJ [1], Fadaie F [1], Caldairou B [1], Bernhardt BC [1,2], Barba C [3], Coelho V [4], Lenge M [3], Semmelroch M [5], Bartolomei F [6], Guye M [7], Cendes F [4], Guerrini R [3], Jackson G [5], Bernasconi N [1], Bernasconi A [1]

[1] Neuroimaging of Epilepsy Laboratory, [2] Multimodal Imaging and Connectome Analysis Lab, Montreal Neurological Institute, McGill University, Montréal, QC, Canada, [3] Pediatric Neurology Unit and Laboratories, Children's Hospital A. Meyer University of Florence, Florence, Italy, [4] Department of Neurology, University of Campinas, Campinas, Brazil, [5] The Florey Institute of Neuroscience and Mental Health and The University of Melbourne, Melbourne, Australia, [6] Aix Marseille Univ, Inserm UMR 1106, INS, Institut de Neurosciences des Systèmes, Marseille, France, [7] Aix Marseille Univ, CNRS, CRMBM UMR 7339, Marseille, France

PURPOSE

Focal cortical dysplasia (FCD) is a prevalent surgically-amenable epileptogenic malformation of cortical development. On MRI, FCD typically presents with cortical thickening, hyperintensity, and blurring of the gray-white matter interface. These changes may be visible to the naked eye, or subtle and be easily overlooked [1]. Despite advances in MRI analytics, current surface-based algorithms [2-5] do not detect FCD in up to 50% of cases [6]. We propose a novel algorithm to distinguish FCD from healthy tissue directly on MRI voxels. Our method harnesses feature learning capability of convolutional neural networks (CNN) [7] with minimal data pre-processing. Our algorithm was trained and tested on data from the Montreal Neurological Institute (S1), and tested on independent data from S1 and four sites worldwide (S2-S5), for a total of 185 individuals.

METHOD

Figure 1 summarizes our method.

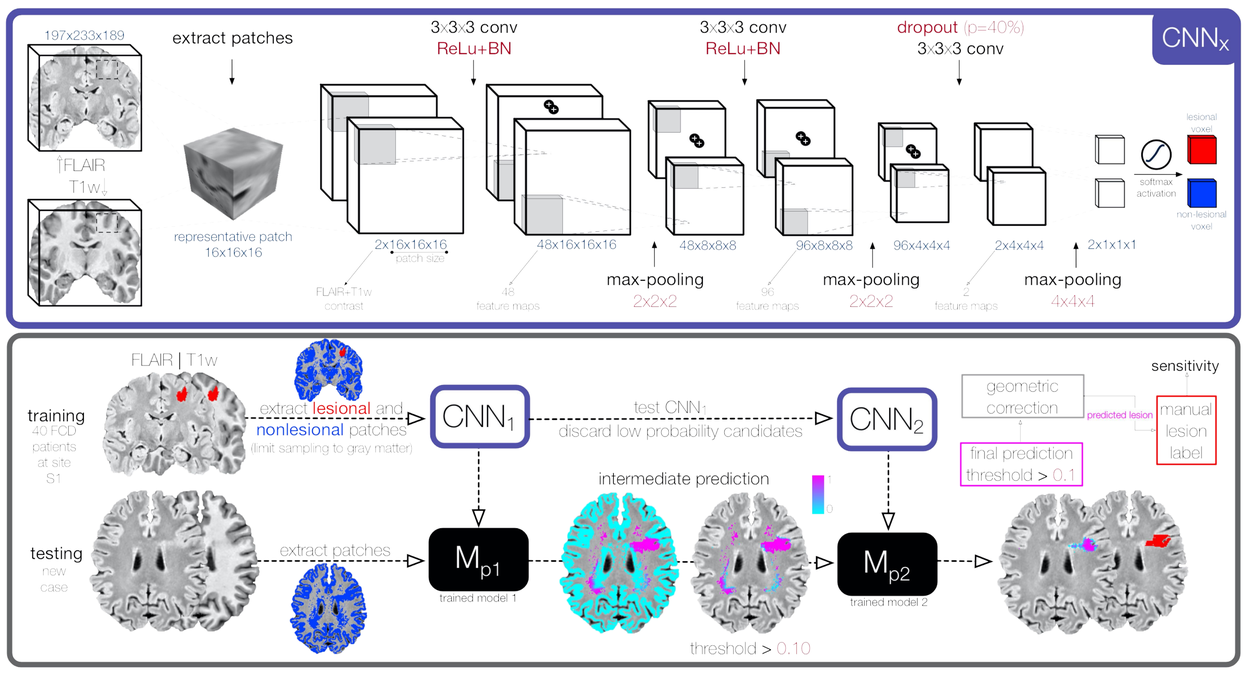

Figure 1. Upper panel: Convolutional network architecture (CNN_x) for two-label classification with three consecutive convolutions and max-pooling units, followed by a voxel-wise softmax classification using multimodal (FLAIR+T1w) patches. Each convolution is followed by rectified linear units (ReLu) to introduce non-linearity. Batch normalization (BN) and dropout serve as regularizers. Adadelta serves as the optimizer to minimize the binary crossentropy loss.
Lower panel: Training and testing schema using two-stage CNN_x cascade (CNN1/CNN2). To reduce training times, multimodal patches (2x16x16x16; centered around the voxel to classify) are sampled from a gray matter (GM) mask. This mask is generated using intra-subject z-score of FLAIR contrast, discarding hypointense voxels (z<0.1). Performance (sensitivity) is measured relative to the expert manual labels.

Training. Volumetric 3T T1-weighted 3D-MPRAGE and 3D-FLAIR MRI were collected in 40 patients (mean age: 28±9) with histologically verified FCD. Routine MRI was initially reported as unremarkable in 80% (33/40). Images underwent intensity inhomogeneity correction and standardization. T1w images were registered linearly to the MNI152 template. FLAIR images were linearly registered to T1w MRI, and subsequently to MNI152 based on the previously estimated registration.
Classifier design. Volumetric datasets served as inputs to a two-stage cascaded CNN (Fig. 1). The first CNN was designed to maximize sensitivity (i.e., detecting a maximum number of lesional voxels), the second optimized specificity (i.e., reducing false positives).
Validation. At S1, a 5-fold cross validation repeated 20 times tested sensitivity (prediction co-localizing with manual FCD labels). Specificity (false positives) was assessed by testing the model on 38 age-/sex-matched healthy controls (age: 30±7) and 63 temporal lobe epilepsy (TLE) disease controls (age: 31±8). Sensitivity was tested in a separate cohort of 44 histologically-confirmed FCD (30 adults, age: 32±11; 14 children, age: 9±5) across sites with different scanners, field strengths (1.5T, 3T) and head-coils (8, 16, 32 channels).

RESULTS

For S1, sensitivity was 87±4% (an average of 35/40 lesions detected). In these cases, 2±1 extra-lesional clusters were detected. Specificity was 95% in healthy controls (3±1 clusters in 2/38) and 90% in TLE (1±0 in 7/63). For cross-dataset classification, overall sensitivity was 89% (39/44) with 4±2 extra-lesional clusters in 27 cases. Per-site sensitivity in S1-S5 was 100% (8/8 FCD detected, 2±2 extra-lesional clusters), 86% (12/14, 4±2), 89% (8/9, 2±1), 75% (6/8, 2±1), and 100% (5/5, 5±2), respectively.
Figure 2 shows test case examples.

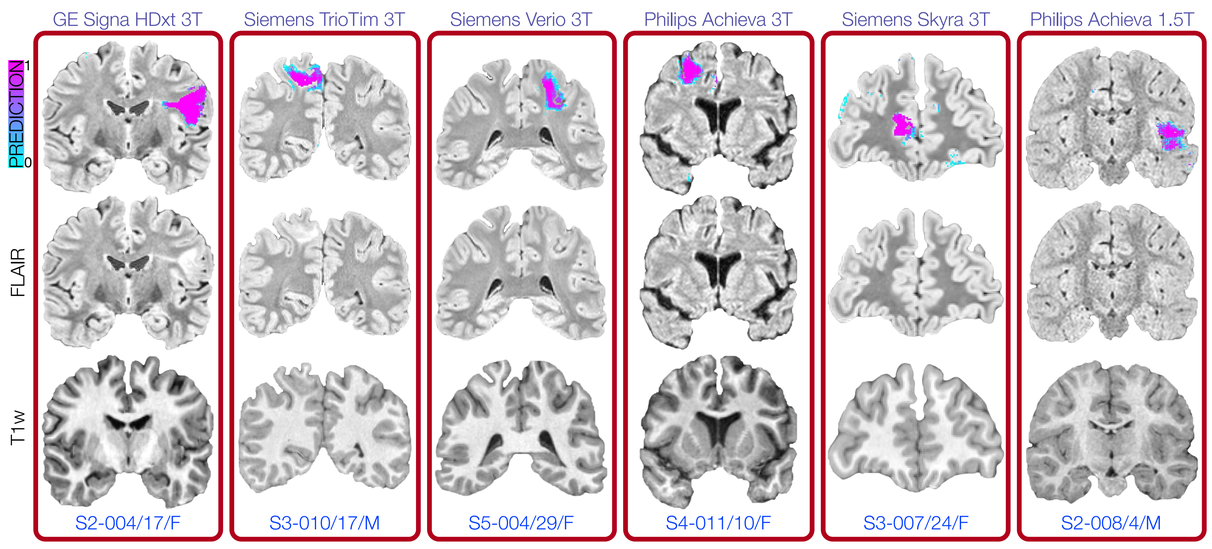

Figure 2. Classification results using the cascaded CNN_x trained on 40 FCD patients at site S1 (Siemens TrioTim 3T) to demonstrate generalizability for lesion detection along three axes of heterogeneity: scanner type, field strength (top labels), and age (bottom labels). The six columns represent six cases obtained using different scanners at five sites (excluding site S1). The top row indicates the strength of prediction overlaid on the FLAIR, while the second and third rows show the corresponding slices for the FLAIR and T1w contrasts, respectively. The bottom labels are read as site-patient-ID/age/gender.

CONCLUSION

We present the first multicenter study on automated FCD detection based on histologically-confirmed lesions based on deep learning. Operating on routine multi-contrast MRI in voxel-space, our algorithm provides the highest performance to date. We demonstrated generalizability of a model trained on a single-site dataset by showing robust performance across independent cohorts from various centers worldwide with different age, scanner hardware and sequence parameters. Notably, >50% of lesions were missed by conventional radiological inspection. Easy implementation, minimal pre-processing, significant performance gains [6], and inference time of <6 minutes/case make this classifier an ideal platform for large-scale clinical use, particularly in "MRI-negative" FCD.

REFERENCES

1. Bernasconi A, et al. Advances in MRI for 'cryptogenic' epilepsies. Nat Rev Neurol 7, 99–108 (2011)
2. Adler S., et al. Novel surface features for automated detection of focal cortical dysplasias in paediatric epilepsy. NeuroImage: Clin 14, 18–27 (2017)
3. Hong, S.-J. et al. Automated detection of cortical dysplasia type II in MRI-negative epilepsy. Neurology 83, 48–55 (2014)
4. Gill RS, et al. Automated detection of epileptogenic cortical malformations using multimodal MRI. DLMIA/ML-CDS, MICCAI Proceedings (eds. Cardoso MJ, et al.) 349–356 (2017)
5. Tan YL, et al. Quantitative surface analysis of combined MRI and PET enhances detection of focal cortical dysplasias. Neuroimage 166, 10–18 (2017)
6. Kini, LG, et al. Computational analysis in epilepsy neuroimaging: A survey of features and methods. Neuroimage: Clin 11, 515–529 (2016)
7. LeCun Y, et al. Deep learning. Nature 521, 436–444 (2015)

*ravnoor.gill@mail.mcgill.ca
www.bic.mni.mcgill.ca/~noel