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

Gill RS [1,2*], Hong SJ [1,2], Fadaie F [1,2], Caldairou B [1,2], Lee HM [1,2], Bernhardt BC [1,3], Hall J [1], Veilleux M [1], Dubeau F [1], Eliane Kobayashi E [1], Frauscher B [1], Dudley R [4], Carmant L [5], Lortie A [5], Weil A [5], Nguyen D [6], Barba C [7], Brandt A [8], Coelho V [9], d'Incerti L [10], Lenge M [7], Semmelroch M [11], Bartolomei F [12], Guye M [13], Deleo F [10], Schulze-Bonhage A [8], Cendes F [9], Guerrini R [7], Jackson G [11], Bernasconi N [1,2], Bernasconi A [1,2]

[1] Department of Neurology and Neurosurgery, [2] Neuroimaging of Epilepsy Laboratory, [3] Multimodal Imaging and Connectome Analysis Lab, Montreal Neurological Institute, McGill University, Montréal, QC, Canada; [4] Department of Neuropediatrics, McGill University; [5] Department of Neuropediatrics and [6] Neurology, Université de Montreal; [7] Pediatric Neurology Unit and Laboratories, Children's Hospital A. MeyerUniversity of Florence, Italy; [8] Freiburg Epilepsy Center, Universitätsklinikum Freiburg, Germany; [9] Department of Neurology, University of Campinas, Brazil; [10] Istituto Neurologico Carlo Besta, Milano, Italy; [11] The Florey Institute of Neuroscience and Mental Health and The University of Melbourne, Melbourne, Australia; [12] Aix Marseille Univ, Inserm UMR 1106, INS, Institut de Neurosciences des Systèmes, Marseille, France; [13] Aix Marseille Univ, CNRS, CRMBM UMR 7339, Marseille, France

PURPOSE

Focal cortical dysplasia (FCD) is a surgically-amenable epileptogenic developmental malformation. Despite advances in MRI analytics, current algorithms [1-4] do not detect > 50% of FCD lesions [5]. Moreover, their use requires specialized expertise, thus precluding widespread clinical application. Our purpose was to develop a novel algorithm to distinguish FCD from healthy tissue directly on MRI voxels. We propose a method harnessing feature learning capability of convolutional neural networks (CNN) [6], a powerful deep learning paradigm. Our algorithm was trained and tested on data from Montreal Neurological Institute (Site 1) and tested on independent data from S1 and 6 sites worldwide (S2-S7), for a total of 230 individuals.

METHOD

Training. Volumetric T1-weighted 3D-MPRAGE (T1w) and 3D-FLAIR MRI were collected in 129 patients with histologically-verified FCD. Images underwent intensity inhomogeneity correction and standardization [7]. T1w images were linearly registered to the age-appropriate MNI152 symmetric template [8]. FLAIR images were linearly mapped to T1w images in MNI space. The training dataset included 40 patients (mean age: 28±9) evaluated at S1; routine MRI was initially reported as negative in 80%.

Classifier design. Patches extracted from volumetric datasets served as inputs to a two-stage cascaded CNN (**Figure 1**): the first CNN was designed to maximize sensitivity (*i.e.*, detecting a maximum number of lesional voxels), while 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 was assessed by testing the model on 38 healthy controls (age: 30±7) and 63 TLE-HS patients as disease controls (age: 31±8). For validation, sensitivity was tested in an independent cohort of 89 histologically-verified FCD patients (47 adults, age: 32±11; 42 children, age: 8±5) across S1 and 6 other sites.

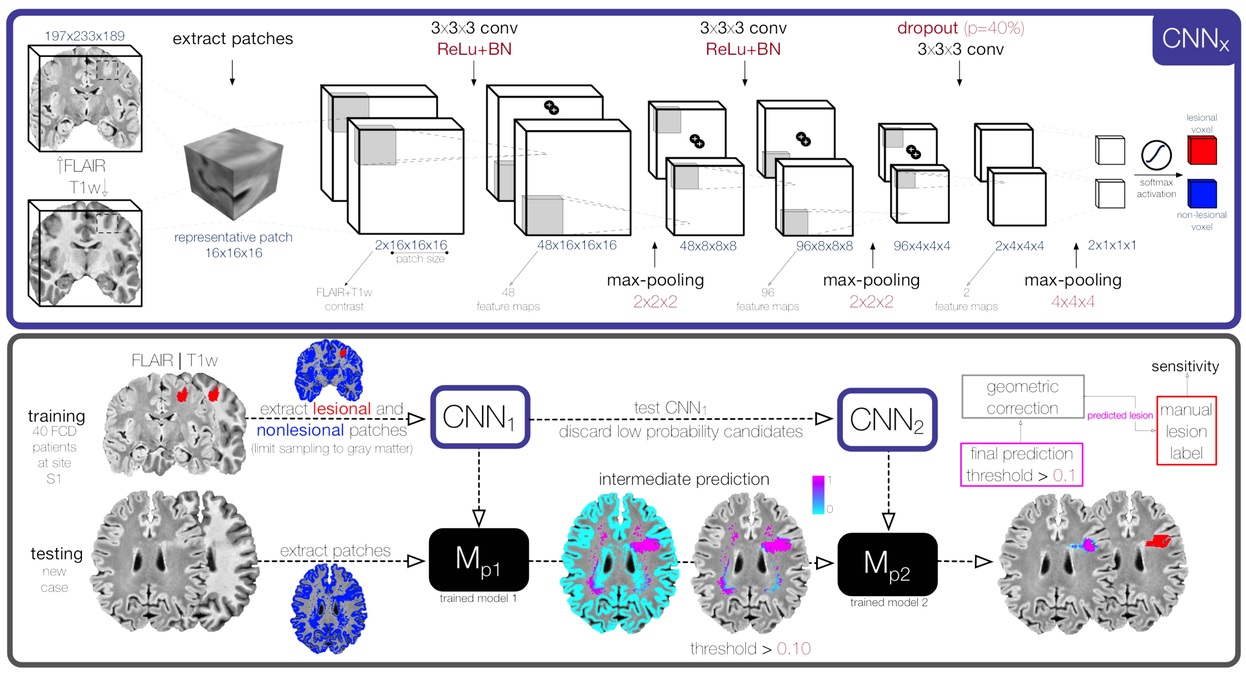

Figure 1. **Upper panel**: Convolutional network architecture (CNN_x) for two-label classification with three consecutive convolutions and max-pooling units, followed by a voxel-wise softmax classification using multimodal (FLAIR+T1w) patches. **Lower panel**: Training/testing schema using two-stage cascade (CNN1/CNN2). Multiomodal patches (size: 2x16x16x16) 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.

RESULTS

For S1, sensitivity was 87±4% (average of 35/40 FCD 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 92% (82/89; 4±3 in 60). Per-site sensitivity in S1-S7 is listed in **Table**, and **Fig. 2** shows test case examples.

Site	Sensitivity	FP clusters in patients
S1	100% (14/14)	2±2
S2	85% (17/20)	4±2
S3	89% (8/9)	2±1
S4	75% (6/8)	2±1
S5	100% (8/8)	4±2
S6	96% (22/23)	4±4
S7	100% (7/7)	2±1

Table. Per-site sensitivity in sites S1-S7. The accompanying incidence of false positive (FP) clusters (mean±S.D.) at each site is also listed.

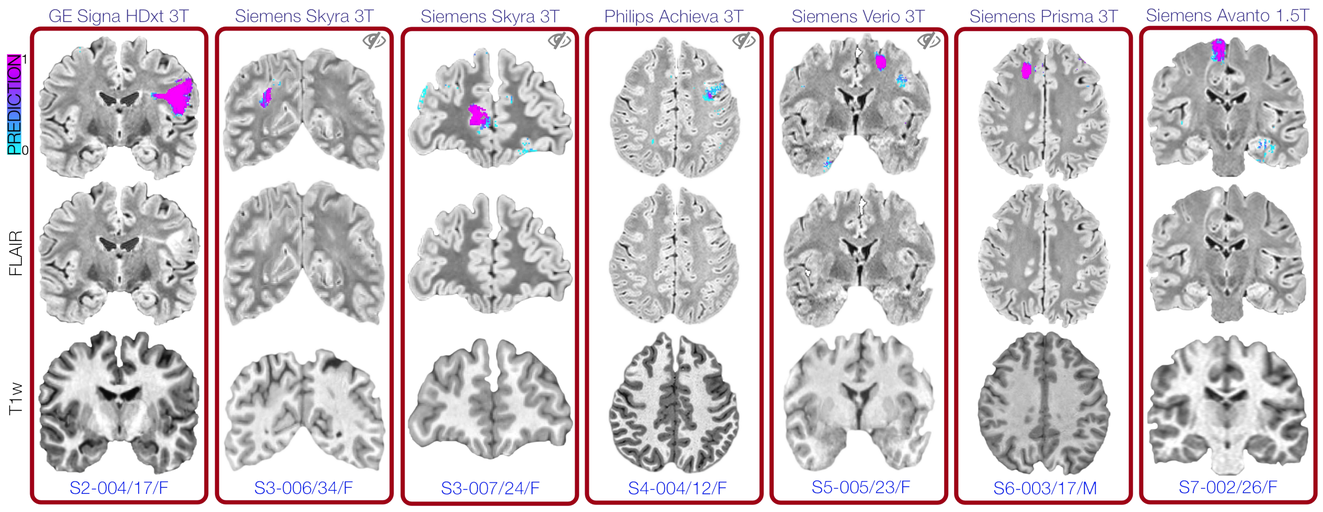

Figure 2. Classification results using the cascaded CNNx trained on 40 FCD patients at site S1 (Siemens TrioTim 3T) to demonstrate generalizability for lesion detection along three axes of heterogeneity: scanner type, field strength (top labels), and age (bottom labels). The seven cases obtained using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

CONCLUSION

We present the first deep learning multicentre study for automated FCD detection based on histologically-confirmed lesions. Operating on routine multi-contrast MRI in voxel-space, our algorithm provides the highest performance to date. We demonstrated generalizability by showing robust performance across independent cohorts with different age, scanner hardware and sequence parameters. Notably, ~50% of FCD lesions were missed by conventional MRI visual inspection. Easy implementation, minimal pre-processing, performance gains, and inference time of < 6 min/case make this classifier the ideal platform for large-scale clinical use, particularly in "MRI-negative" FCD.

REFERENCES

[1] Adler S., et al. NeuroImage Clin 15, 95-105 (2017), [2] Hong, S.-J. et al. Neurology 83, 48–55 (2014), [3] Gill RS, et al. DLMIA/ML-CDS, MICCAI Proceedings 349–356 (2017), [4] Tan YL, et al. Neuroimage 166, 10–18 (2017), [5] Kini, LG, et al. Neuroimage: Clin 11, 515–529 (2016), [6] LeCun Y, et al. Nature 521, 436–444 (2015), [7] Sled JG et al. IEEE Trans Med Imaging17, 87-97 (1998), [8] Fonov V et al. Neuroimage 54, 313-327 (2011)

neuroimaging of epilepsy laboratory

noel

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

pdf