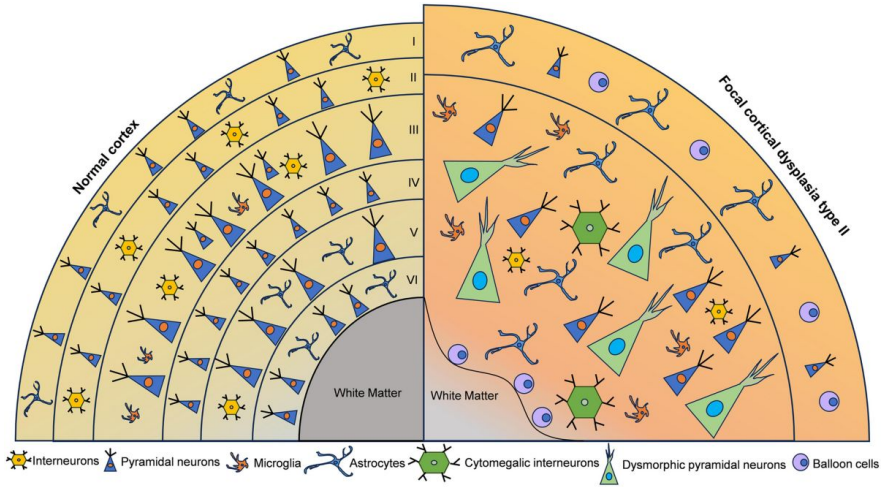
Detection of brain area with Focal Cortical Dysplasia

Butylin Andrey
Kashuk Ekaterina

Advisor: *Nadezhda Alsahanova*

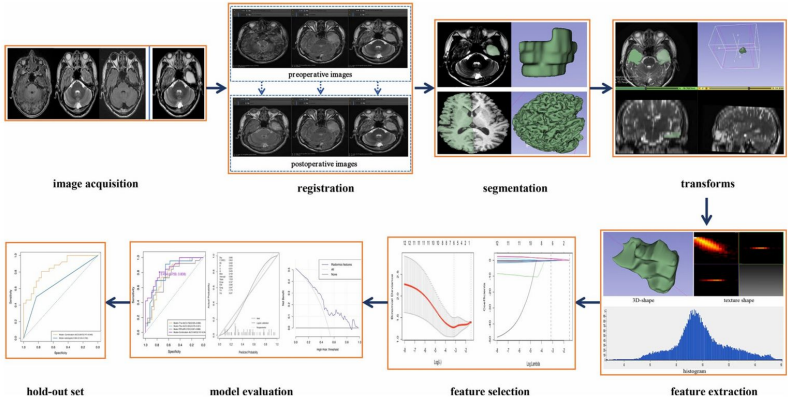
Focal cortical dysplasia (FCD) is one of the most common causes of *drug-resistant epilepsy*. However, its diagnosis remains extremely difficult, particularly in mild or "MRI-negative" cases. Traditional visual assessment methods, such as T1w, T2w, and FLAIR, *strongly depend on the radiologist's experience* and often fail to detect subtle microstructural changes.



- Lamination is lost except for an expanded layer I
- Neurons are dispersed through layers II–VI
- Often extend into the white matter, obscuring the gray–white boundary

Fang, Y., Zhang, Y., Huang, T. *et al.* Focal cortical dysplasia type II: review of neuropathological manifestations and pathogenetic mechanisms. *Acta Epileptologica* 7, 12 (2025).

A radiomics nomogram based on multiparametric MRI for diagnosing focal cortical dysplasia and initially identifying laterality (Chen et al. 2024)

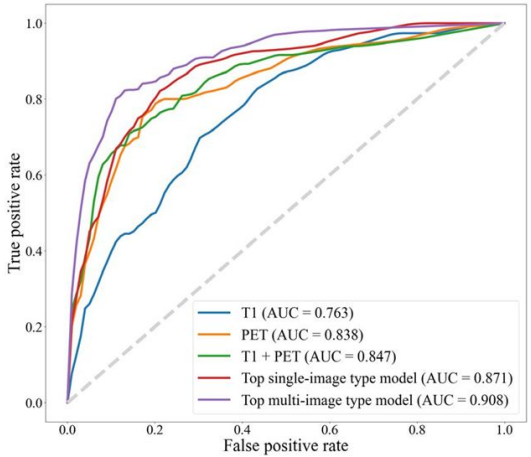


Modality	Sets	AUC(95%CI)
Model-T1w	Training	0.762(0.625–0.898)
	Validation	0.671(0.400-0.943)
Model-T2w	Training	0.805(0.679–0.931)
	Validation	0.814(0.606-1.000)
Model-T2FLAIR	Training	0.733(0.597–0.868)
	Validation	0.714(0.374-1.000)
Model-Combination	Training	0.847(0.747–0.946)
	Validation	0.857(0.618-1.000)
Model-radiologists	Training	0.664(0.543–0.784)
	Validation	0.521(0.148–0.893)
Model-Combination	Hold-out	0.828(0.619-1.000)
Model-radiologists	Hold-out	0.571 (0.363–0.778)

Most of useful features from T2w contrast

- 43 patients with histologically confirmed FCD.
- Comparison of radiomic signs of a healthy hemisphere and a hemisphere with diagnosed FCD.
- The signs stood out from three different contrasts, which led to an increase in ROC-AUC.
- The key model is Logistic regression.
- Finding a correlation between the signs and duration of the disease

Multi-level Fusion of FDG PET and MRI for Automated Epileptic Lesion Detection (Zijun et al. 2024)



- Combining two approaches: MRI (T1w) - object structure, PET - strict localization
- 93 radiomics features
- An approach based on separate approaches and based on a merget image
- Not only detection, but also surgical navigation

wA= 0.7
wD= 0.9

$$C_{A,\,fused} = w_A \cdot C_{A,\,PET} + (1-w_A) \cdot C_{A,\,T1}$$

Approximation components (CA) — contain a low-frequency, general structure (global shapes); Low-pass filter

DWT (Discrete Wavelet Transform)

$$C_{D,\,fused} = w_D \cdot C_{D,\,PET} + (1-w_D) \cdot C_{D,\,T1}$$

$$C_{D,\,fused} = \max\{|C_{D,\,PET}|, |C_{D,\,T1}|\}$$

Detailed components (CD) — contain high-frequency details (edges, textures, borders); High-pass filter

Goal: build a patient-wise machine-learning pipeline using the provided dataset, which enforce zero False Negatives while minimizing False Positives in FCD zones identification and provide clinical and feature-level interpretation of results.

Tasks:

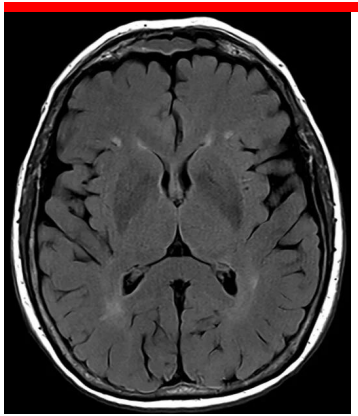
1. Split data for train-test by patients, not areas.
2. Choose and train model on train set correctly (choice of hyperparameters, model selection done using train set)
3. On train set, achieve 0 False Negatives and as small as possible False Positives (for each patient there should be at least one True Positive)
4. Assess the best model on test set
5. Make an interpretation of the results

Initial dataset information:

- Total samples: 11088
- Healthy samples: 10660
- FCD samples: 428
- Unique subjects: 168
- Unique brain areas: 66
- Total radiomics features: 93

Increasing the number of features

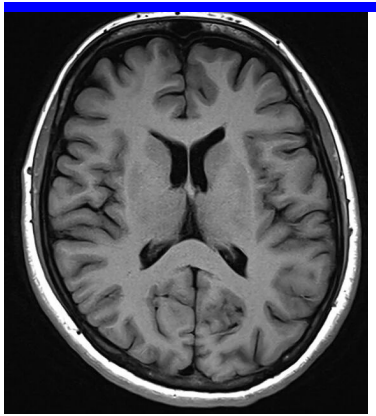
FLAIR



Initial dataset: 93 features

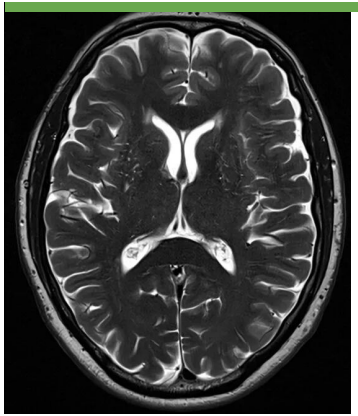
+

T1w



T1w data: 93 features

T2w

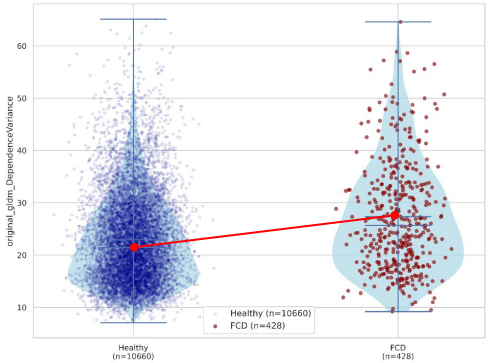


T2w data: 93 features

Top significant features by Mann-Whitney U test

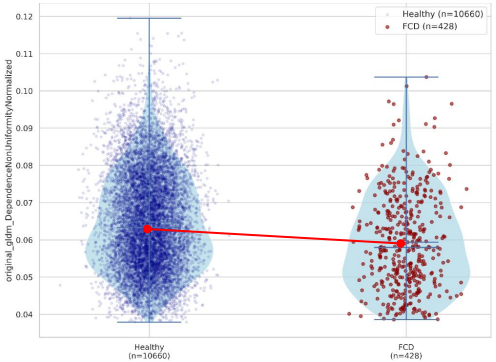
FLAIR

original_gldm_DependenceVariance



increased local cluster size variability consistent with patchy cortical organization

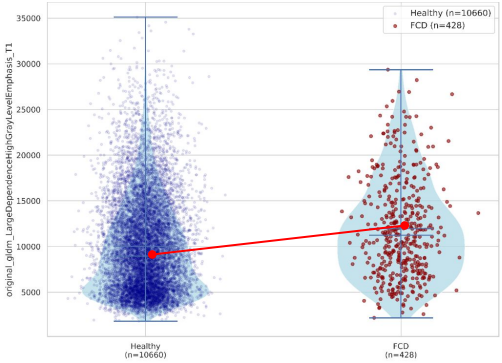
original_gldm_DependenceNonUniformityNormalized



reduced neighborhood dependency uniformity, consistent with overall heterogeneity of tissue

T1w

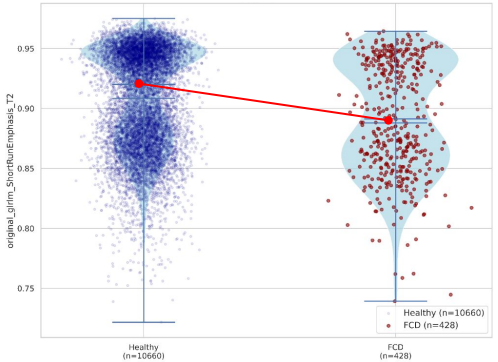
original_gldm_LargeDependenceHighGrayLevelEmphasis_T1



indicates large high-intensity homogeneous patches on T1

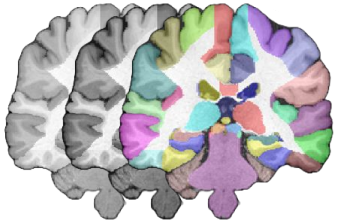
T2w

original_glrlm_ShortRunEmphasis_T2



indicates characteristic tissue disorganization and loss of normal lamination

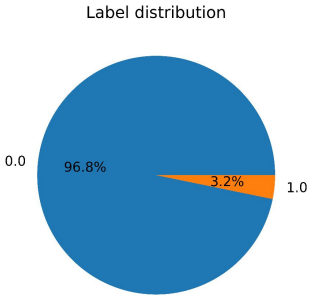
Training and optimization

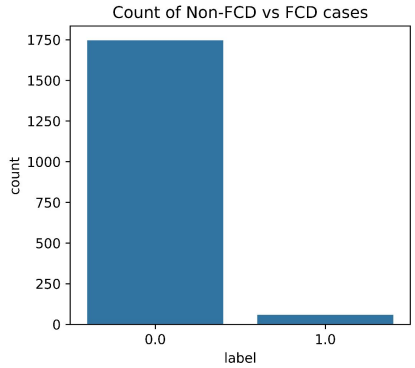


Applied a **StandardScaler** to 279 features to make radiomics measures comparable

Held out 30% of subjects as a final test set (no overlap with training set) to evaluate generalization

Train subjects: 117	**Test** subjects: 51
Train samples: 7722	**Test** samples: 3366





Dataset is highly imbalanced , **RandomUnderSampler** was used on the training portion of each fold to rebalance classes for model fitting

Model selection with a **GroupKFold** (n_splits=5) so that every fold contained only whole patients and no region from the same subject appeared in both training and validation sets.

Training and optimization

GridSearch for 6 different models

- Logistic Regression: C, penalty
- SVM: C ,kernel, gamma
- KNN: n_neighbors, p, weights
- Gradient Boosting: loss, learning_rate, n_estimators, subsample, criterion
- Random Forest: n_estimators , criterion , max_depth, max_features, class_weight
- XGBoost: n_estimators , max_depth, learning_rate

Thresholds that satisfy that constraint → selection the one that **minimize fold False Positives**

Optimized a decision threshold per fold using the rule **False Negative patients == 0**
"*every positive patient must have ≥1 True Positive detected region on that training fold*"

```
Testing params: {'C': 100.0, 'penalty': 'l1'}
 fold0: t=0.000, fp=1530, pfn=0, rec region=1.000, rec patient=1.000, spec region=0.000
 fold1: t=0.000, fp=1525, pfn=0, rec region=1.000, rec patient=1.000, spec region=0.000
 fold2: t=0.070, fp=1023, pfn=0, rec region=0.912, rec patient=1.000, spec region=0.294
 fold3: t=0.000, fp=1470, pfn=0, rec region=1.000, rec patient=1.000, spec region=0.000
 fold4: t=0.000, fp=1476, pfn=0, rec region=1.000, rec patient=1.000, spec region=0.000
  avg_fp: 1404.8 total_patient_fn: 0 avg_threshold: 0.014 avg_rec_region: 0.9824 avg_rec_patient: 1.0 avg_spec_region: 0.0589
```

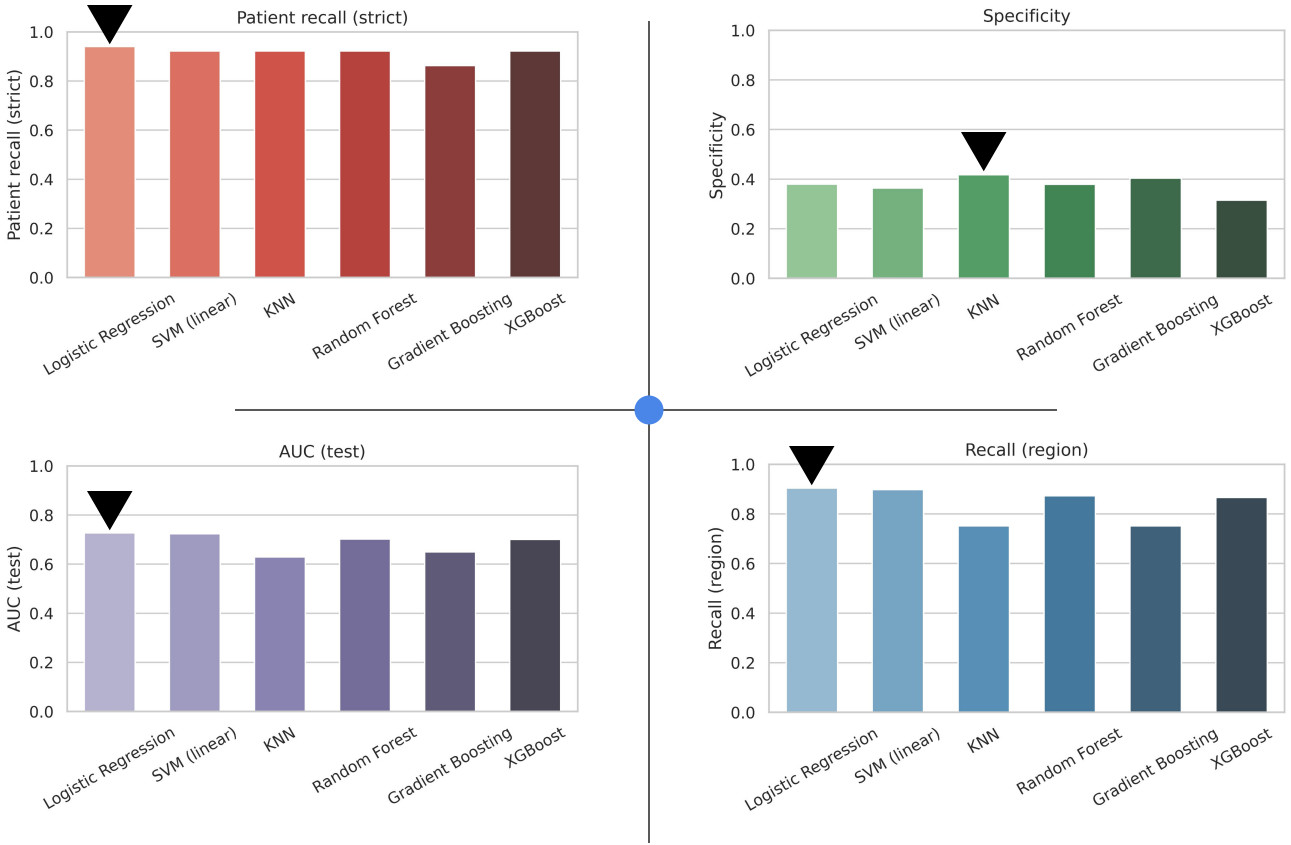
Per-fold thresholds were averaged and used for the final evaluation

Best parameters to test

```
BEST params: {'C': 0.01, 'penalty': 'l2'}
BEST avg fp: 885.6
BEST avg rec region: 0.9057813332489303
BEST threshold (mean from folds): 0.354
```

Results on test

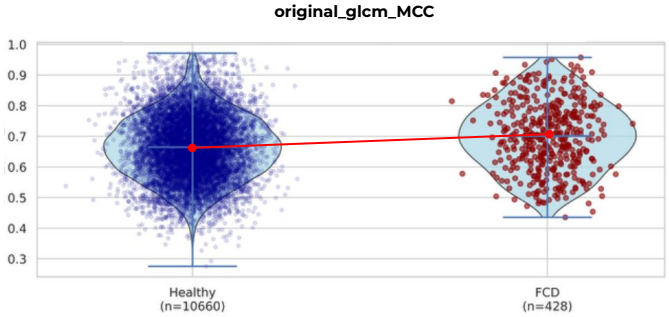
Best models: Logistic Regression and Random Forest



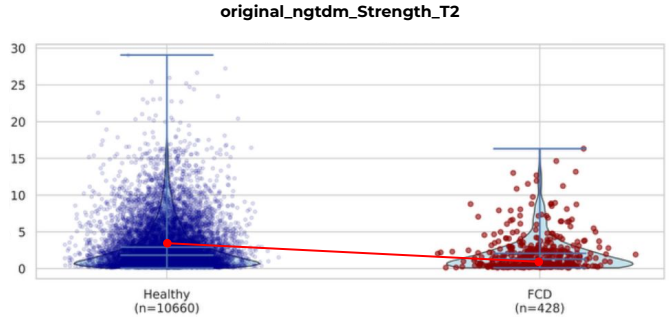
Results

Model	Best parameters	Specificity on Train	Region Recall on Train	Patient Recall on Train	Specificity on Test	Region Recall on Test	Patient Recall on Test	AUC
Logistic Regression	{'C':0.01,'penalty':'l2'}	0.4059	0.9058	1	0.38	0.904	0.941	0.727
SVM (linear)	{'C':0.01,'kernel':'linear','gamma':'scale'}	0.4004	0.9217	1	0.364	0.898	0.922	0.724
KNN	{'n_neighbors':50,'weights':'distance','p':2}	0.3825	0.8623	1	0.418	0.752	0.922	0.629
Random Forest	{'n_estimators':200,'criterion':'entropy','max_features':'sqrt','class_weight':'balanced_subsample'}	0.3872	0.9061	1	0.379	0.873	0.922	0.72
Gradient Boosting	{'n_estimators':300,'max_depth':3,'learning_rate':0.01}	0.4244	0.8515	1	0.404	0.752	0.863	0.65
XGBoost	{'n_estimators':100,'max_depth':3,'learning_rate':0.05}	0.9071	0.3444	1	0.315	0.866	0.922	0.701

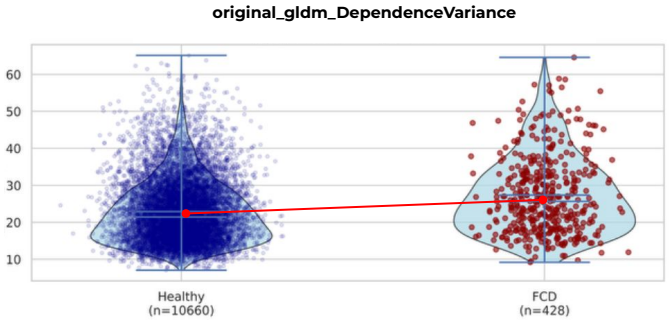
Top important features



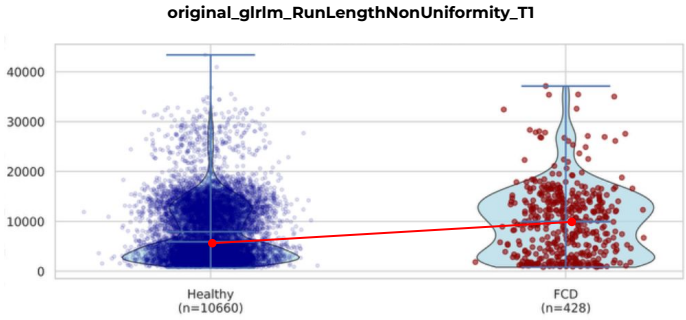
distorted but structured spatial organization



The tissue texture is **less coarse** and **more fine-grained**



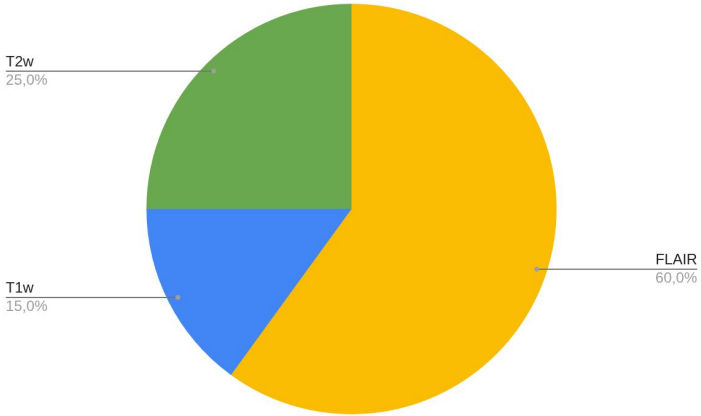
patchy voxel pattern in FCD,
mirroring its tissue disorganization



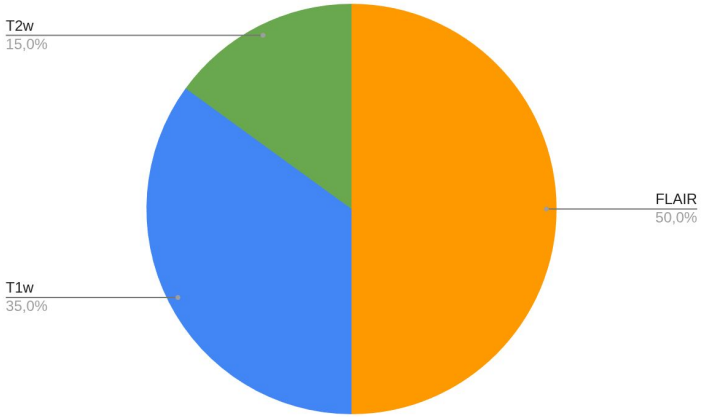
Patchy mix of disorganized and
abnormally smooth tissue

Feature modalities

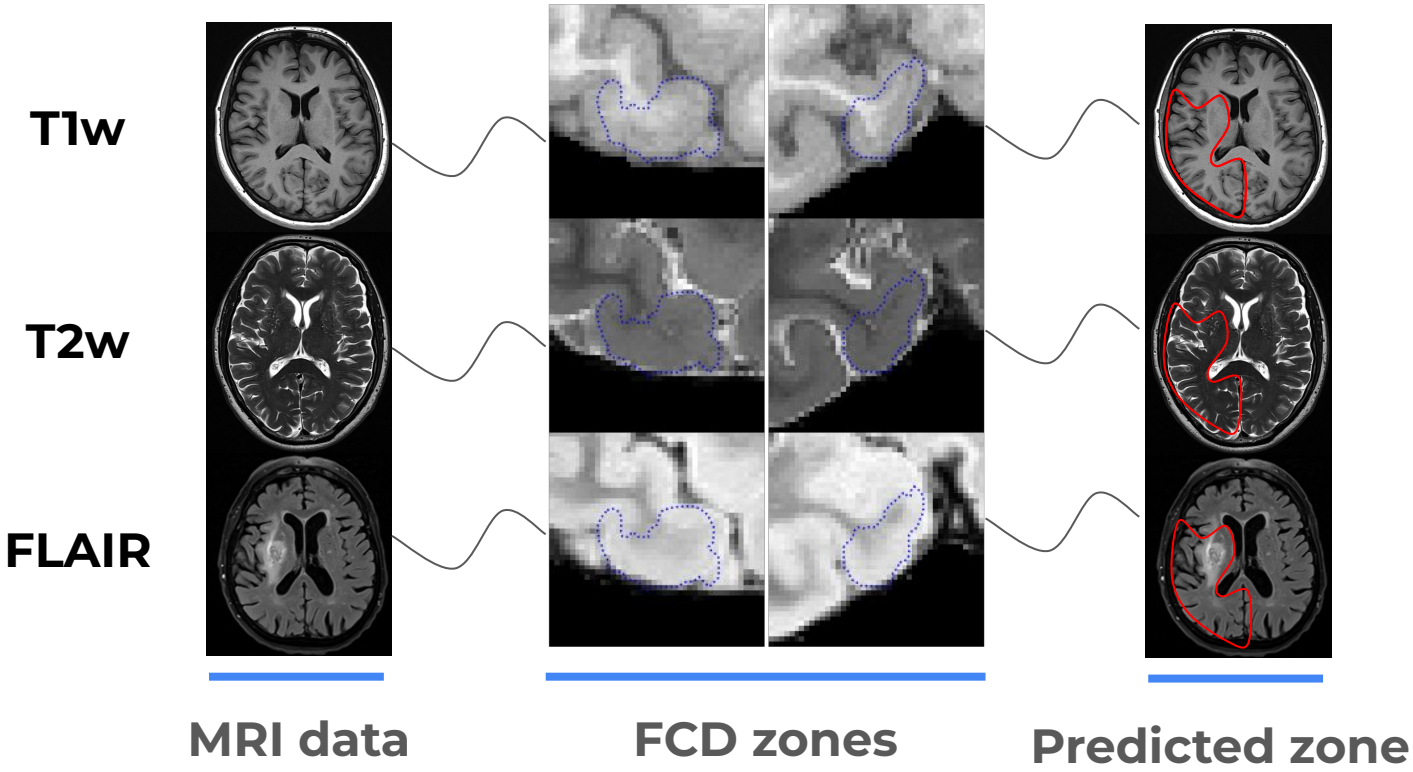
Logistic Regression top 20 features



Random Forest top 20 features

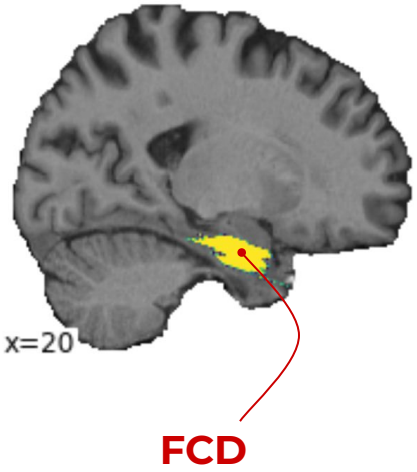


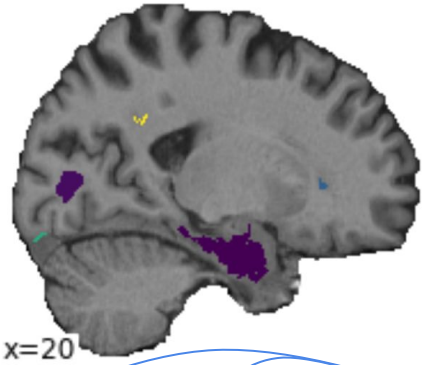
Usage



This solution reduces the area of manual data inspection by a radiologist.

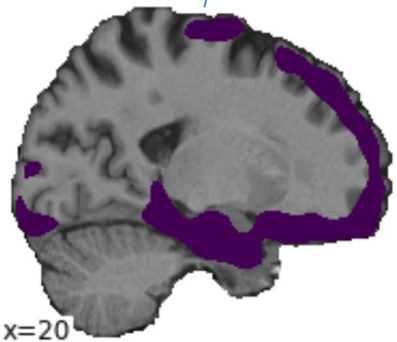
Usage





Initial Advisor's model

Average precision: 0.114
Average dice: 0.165



Our model

Combination

Average precision: 0.159
Average dice: 0.196

Conclusions

- Based on the literature review, the list of T1w and T2W features has been expanded with contrasts
- We achieved a high Recall on validation and test sample (~0.941 for LogReg)
- Logistic regression and Random Forest models showed the best results based on the selected hyperparameters
- The most significant features reflect the morphological characteristic of the FCD
- All the modalities (FLAIR, T1w, T2w) turned out to be significant features
- The model reduces the examination area for radiologists and can be used in combination with other models

What else could be done

- Advanced models, including deep learning, could reveal complex patterns beyond classical radiomics
- Feature interactions (for example, via SHAP) and clinical correlations could be explored

Contributions

Ekaterina Kashuk: EDA, development of the pipeline, hyperparameter tuning for LogRegression, KNN, SVM, XGBoost, biological interpretation of important features, report writing

Andrey Butylin: literature review, hyperparameter tuning for Random Forest and Gradient Boosting, presentation + GitHub project page

Thx!