



VILNIUS UNIVERSITY
INSTITUTE OF DATA SCIENCE AND DIGITAL
TECHNOLOGIES



2025/2026 mokslo metų pirmojo pusmečio ataskaitinė konferencija

Doktorantas: Raimondas Juškys

Darbo vadovė: Prof. dr. Audronė Jakaitienė

- **Disertacijos pavadinimas:** Mašininio mokymosi metodai klasifikavimo ir segmentavimo uždaviniams neuroonkologijoje (application of machine learning algorithms for classification and segmentation tasks in neurooncology).
- **Darbo vadovas:** prof. Audronė Jakaitienė
- **Doktorantūros pradžios ir pabaigos metai:** 2024 – 2028
- **Studijų metai:** 2.

Doktorantūros studijų planas ir jo vykdymo suvestinė (2025/2026, I pusmetis)

Studijų metai	Egzaminai	
	Planas	Įvykdyta
I (2024/2025)	1	1
II (2025/2026)	2	1
III (2026/2027)	1	0
IV (2027/2028)	-	-
Iš viso:	4	0

Mokslinių tyrimų planas ir jo vykdymo suvestinė (2025/2026, I pusmetis)

Studijų metai	Dalyvavimas konferencijose				Publikacijos					
	Tarptautinėse		Nacionalinėse		Su citav. rodikliu			Be citav. rodiklio		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė	Planas	Įvykdyta	Būklė
I (2024/2025)	0	0	0	0	0	0		0	0	
II (2025/2026)	0	0	0	0	0	0		0	0	
III (2026/2027)	1	0	0	0	1	0		0	0	
IV (2027/2028)	1	0	0	0	1	0		0	0	
Iš viso:	2	0	0	0	2	0		0	0	

Ataskaitinio pusmečio planas ir jo vykdymas (2025/2026, I pusmetis)

Egzaminai 2025/2026 (I pusmetis)		
Planas	Įvykdyta	Būklė
Fundamentalieji informatikos ir informatikos inžinerijos metodai	+	Išlaikyta

Dalyvavimas konferencijose 2025/2026 (I pusmetis)		
Planas	Įvykdyta	Konferencijos tipas
Dalyvavimas konferencijose nenumatytas	-	-

Publikacijos 2024/2025 (I pusmetis)			
Planas	Įvykdyta	Būklė	Publikacijos tipas
Publikacijos nenumatytos	-	-	-

Mokslinių tyrimų ir disertacijos rengimo etapai

Darbo pavadinimas		Atlikimo terminai	Pastabos
1.	<p>Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje):</p> <hr/> <p>1.1. Sistemine mokslines literatūros apžvalga: a) Literatūros paieškos strategijos sudarymas. b) Studijų atrinkimo procesas</p> <p>1.2. Sistemine mokslines literatūros apžvalga: a) Publikuotų studijų ir metodų analizė bei palyginimas b) Publikuotų studijų šališkumo vertinimas</p>	<p>2024 m. spalio mėn. – 2025 m. balandžio mėn.</p> <p>2025 m. balandžio mėn. – 2025 m. liepos mėn.</p>	<p>Parengta sistemine literatūros apžvalga.</p>
2.	<p>Mokslinio tyrimo vykdymas:</p> <hr/> <p>2.1. Tyrimo metodikos sudarymas:</p> <p>a) Disertacijos tikslo formulavimas. b) Disertacijos uždavinių formulavimas.</p> <hr/> <p>2.2. Teorinis tyrimas:</p> <p>a) Klasifikavimo ir segmentavimo modelių pritaikymas. b) Algoritmų palyginimas ir tobulinimas, jų optimizacija.</p>	<p>2025 m. liepos mėn. – 2025 m. spalio mėn.</p> <p>2025 m. spalio mėn. – 2026 m. balandžio mėn.</p>	<p>Suformuoti disertacijos tikslas ir uždaviniai.</p> <p>Suformuotas duomenų rinkinys, atliktas pirminis duomenų apdorojimas, radiomikos požymių išskyrimas bei pradinis modeliavimas.</p>

Tyrimo objektas: Mašininio mokymosi metodų taikymas automatiniam galvos smegenų navikų segmentavimui ir radiomikos pagrindu veikiančiam glioblastomos bei solitarinių galvos smegenų metastazių diferenciacijai, naudojant įprastines MRT sekas (T1, T1 su kontrastu, T2, FLAIR).

Tyrimo tikslas: Sukurti, validuoti ir optimizuoti mašininio mokymosi metodikas, skirtas automatiniam galvos smegenų navikų segmentavimui ir tolesniam radiomikos pagrindu vykdomam glioblastomos ir solitarinių galvos smegenų metastazių klasifikavimui naudojant įprastines MRT sekas.

Tyrimo uždaviniai:

1. Atlikti sisteminę literatūros apžvalgą apie automatinius galvos smegenų navikų segmentavimo metodus, įvertinant dažniausiai naudojamą architektūras bei jų taikymo rezultatus.



2. Atlikti sisteminę literatūros apžvalgą apie klasifikavimo metodus, naudojamus glioblastomos ir solitarinių galvos smegenų metastazių diferenciacijai, įvertinant taikytus modelius ir jų diagnostinius rodiklius.



3. Sukurti ir validuoti automatinio galvos smegenų navikų segmentavimo metodiką MRT vaizdams.

4. Sukurti radiomikos požymių išskyrimo metodiką iš automatiškai segmentuotų naviko sričių įvairiose MRT sekose.



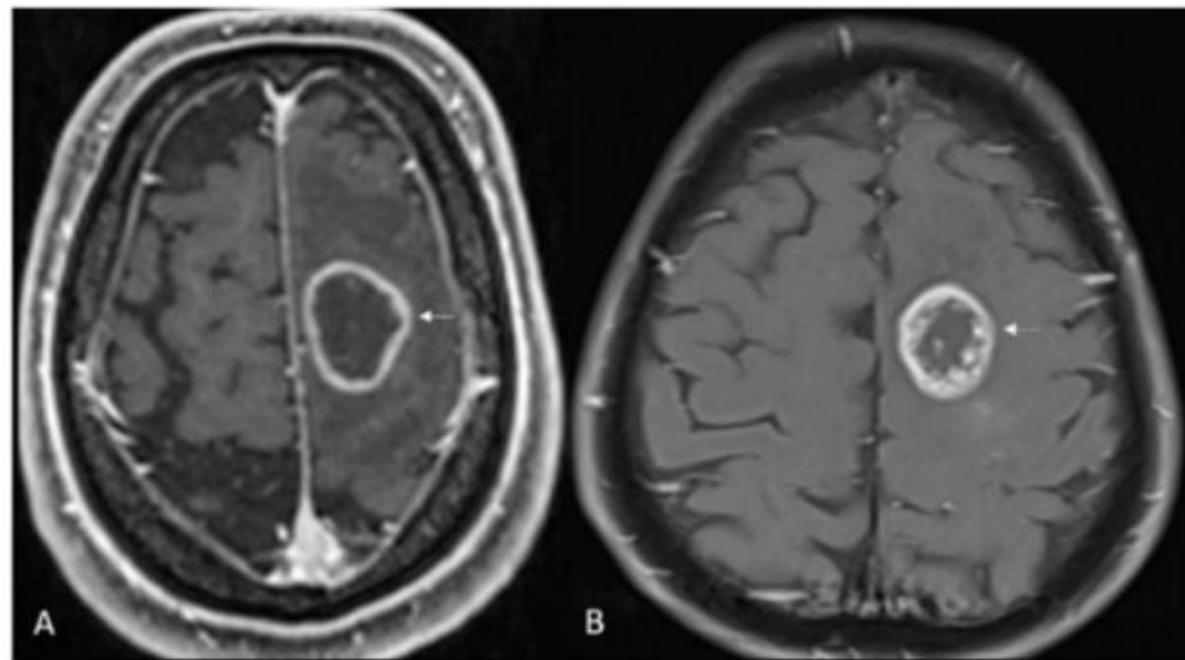
5. Pritaikyti išgautus radiomikos požymius skirtingiems mašininio mokymosi klasifikatoriams palyginti atliekant glioblastomos ir solitarinės metastazės binarinį klasifikavimą.

6. Atlikti požymių svarbos analizę, identifikuojant labiausiai diskriminuojančius radiomikos rodiklius GBM ir SBM klasifikacijai.

2025/2026 mokslo metų pirmojo pusmečio moksliniai rezultatai ir veikla

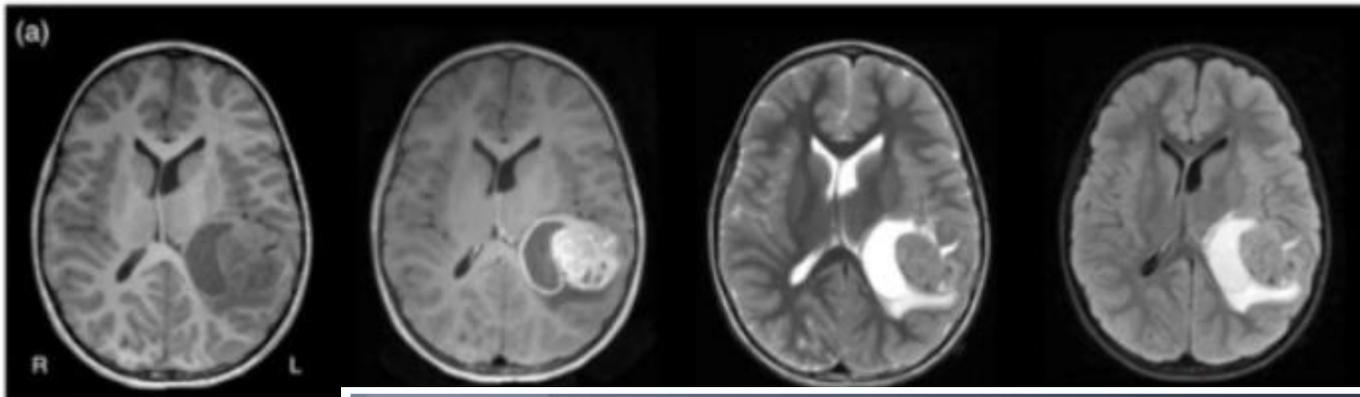
- Suformuotas galutinis tyrimo duomenų rinkinys, apimantis galvos smegenų gliomų ir metastazių magnetinio rezonanso vaizdus. Atliktas duomenų pirminis apdorojimas (angl. preprocessing) bei užtikrinta jų kokybė tolimesnei analizei.
- Atliktas radiomikos požymių išskyrimas (angl. feature extraction) ir požymių inžinerija (angl. feature engineering) iš segmentuotų MRT vaizdų naviko sričių.
- Įgyvendintas pradinis mašininio mokymosi modelių sudarymas ir testavimas binarinio klasifikavimo uždaviniui.

Problem



Data – open access from BraTS 2025 Lighthouse challenge

- **2 Datasets:** BraTS-GLI (n=1251) and BraTS-METS (n=1296).
- **Sequences:** T1, T1Gd, T2, and T2-FLAIR.
- **Annotation Schema:** Rigid four-compartment hierarchy: NETC (Label 1), SNFH (Label 2), ET (Label 3), and RC (Label 4).



Conventional MRI (T1, T1CE, T2, FLAIR)
Accuracy ~65%

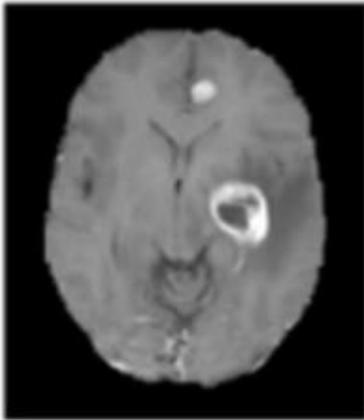
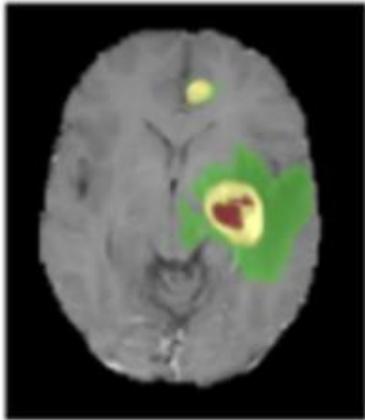
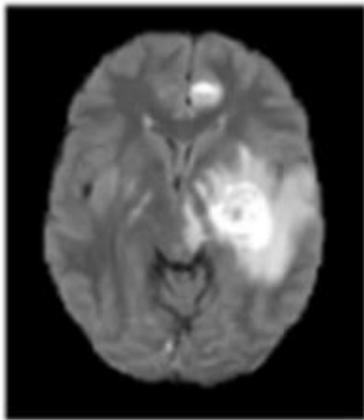
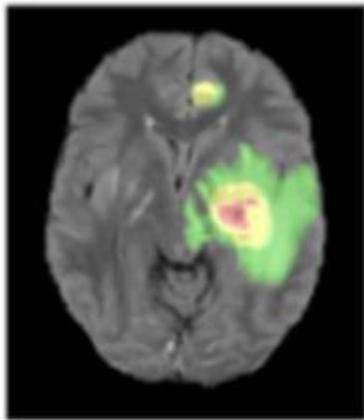
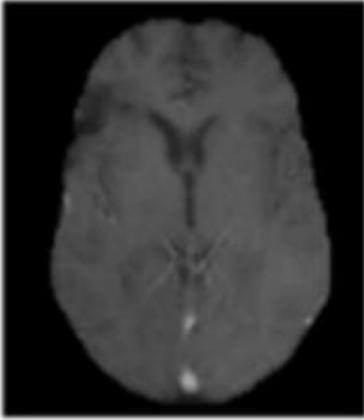
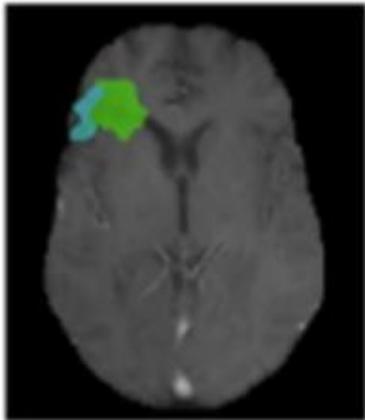
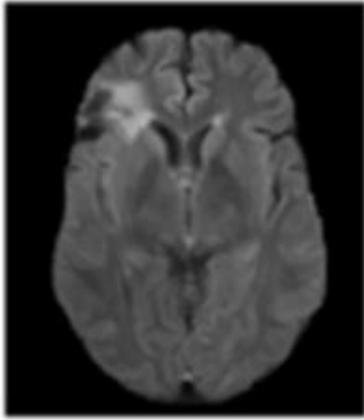
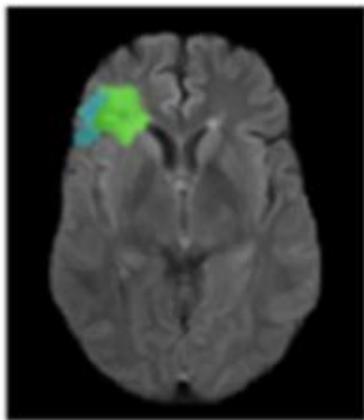


MICCAI BraTS Lighthouse
Cluster of Challenges

Thanks to our sponsors:

MICCAI SIG-CompPath SIG-Challenges SageBionetworks NIH NATIONAL CANCER INSTITUTE Lacuna Fund ML Commons INDIANA UNIVERSITY

Label structure (ROIs)

	T1 post-contrast	T1 post-contrast (annotated)	FLAIR	FLAIR (annotated)
Sample case #1				
Sample case #2				

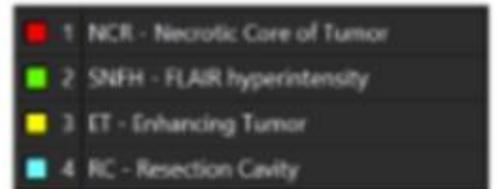
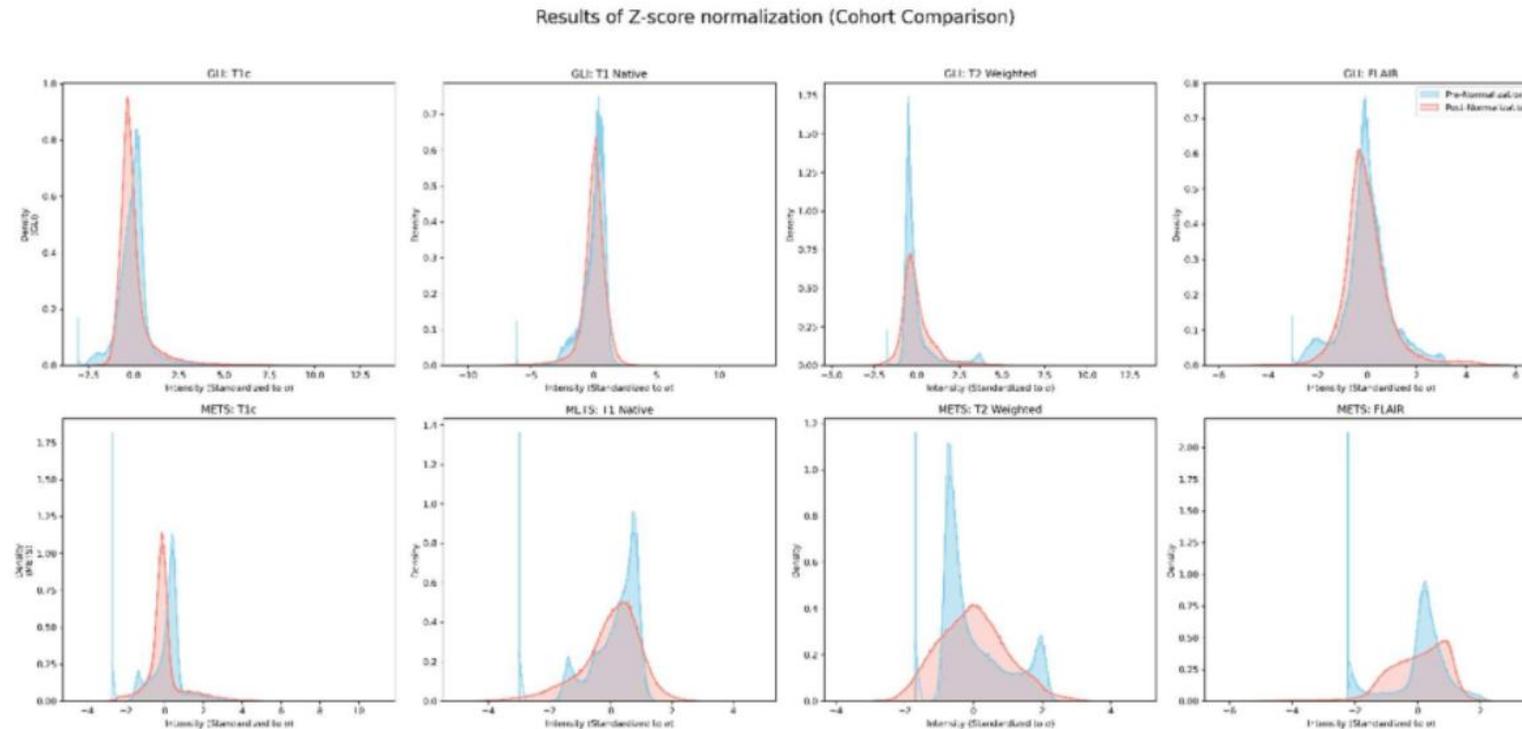


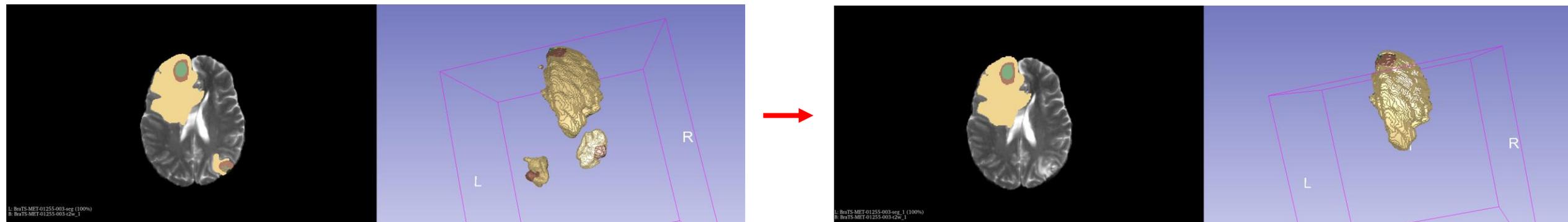
Image-level preprocessing

- **Spatial Unification:** Conversion of heterogeneous orientations (RAS, LAS, LPS) to a unified LPS coordinate system.
- **Isotropic Resampling:** 1.0 mm³ resolution using B-spline for MRI and Nearest Neighbor for labels.
- **Signal Homogenization:** N4 Bias Field Correction followed by Z-score normalization.

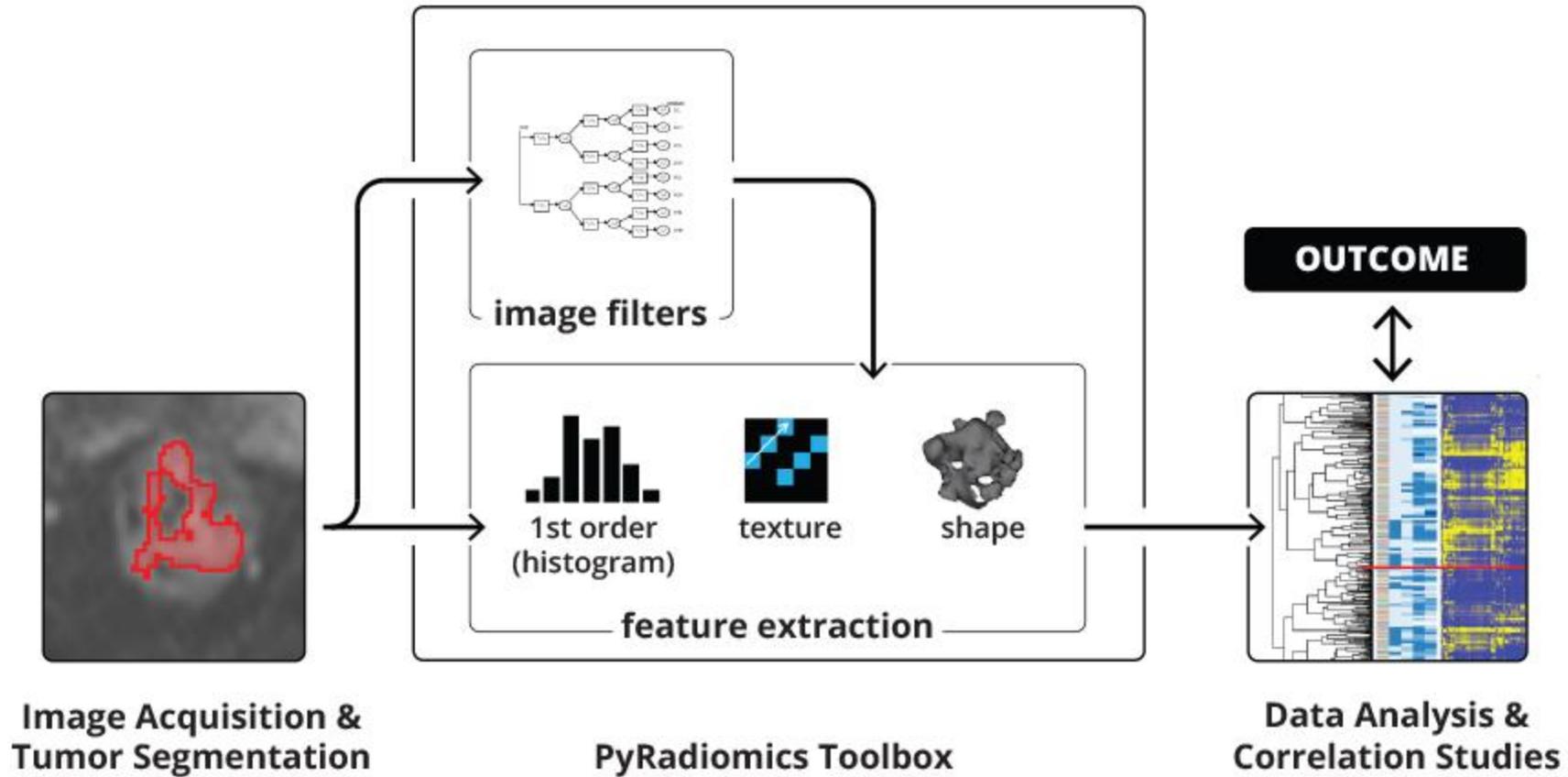


Data selection/filtering

- **Index Lesion Isolation:** Connected Components Analysis (CCA) used to isolate the largest contiguous tumor mass.
- **Resection cavity exclusions:** Cases with label 4 (resection cavity) were excluded
- **Thresholding:** Minimum 200 mm³ active tumor core required for stable texture calculation.
- **Final Cohort:** 1126 GLI cases and 606 METS cases.
- **Stratification:** 80/20 split (Train n=1384, Test n=348), preserving class balance.



Feature extraction



Feature-level preprocessing

- **Dimensionality:** 1744 initial features (107 descriptors x 4 sequences x 4 ROIs + 16 interface (engineered) features).
- **Domains:** First-order statistics, Shape descriptors, and Second-order texture (GLCM, GLRLM, GLSZM, NGTDM, GLDM).
- **Interface Features:** Quantifying the transition zone (e.g., core-to-edema volume ratio) to study tumor-edema margin zone.
- **Dimensionality reduction:** 2 steps:
 - 1) **Correlation filter** – Spearman correlation $|r| > 0.95$ reduced feature map to 845.
 - 2) **Recursive feature elimination cross validation (RFECV)** – identified optimal feature vector (S_{opt}) of 107 features

Internal validation

Table 14: Summary of internal validation results

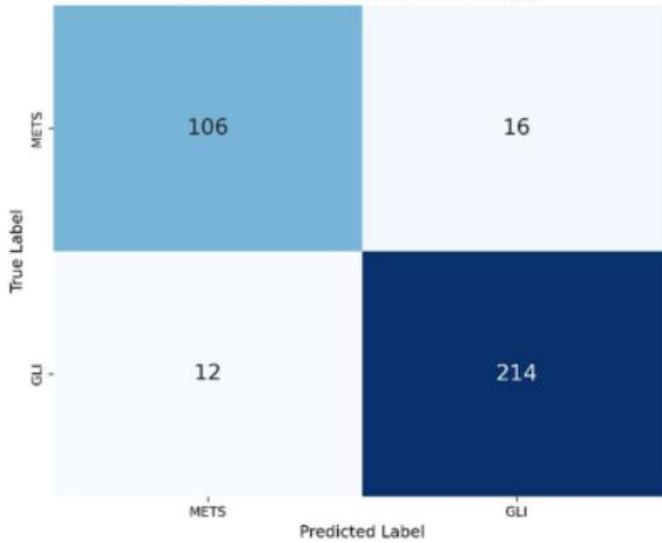
Metric	Random Forest (Mean \pm SD)	Support Vector Machine (Mean \pm SD)
Accuracy	0.9400 \pm 0.0234	0.9523 \pm 0.0127
AUC	0.9819 \pm 0.0096	0.9873 \pm 0.0082
Precision	0.9406 \pm 0.0239	0.9528 \pm 0.0128
Recall	0.9400 \pm 0.0234	0.9523 \pm 0.0127
F1-Score	0.9396 \pm 0.0239	0.9522 \pm 0.0127
NPV	0.9336 \pm 0.0345	0.9362 \pm 0.0248

External validation

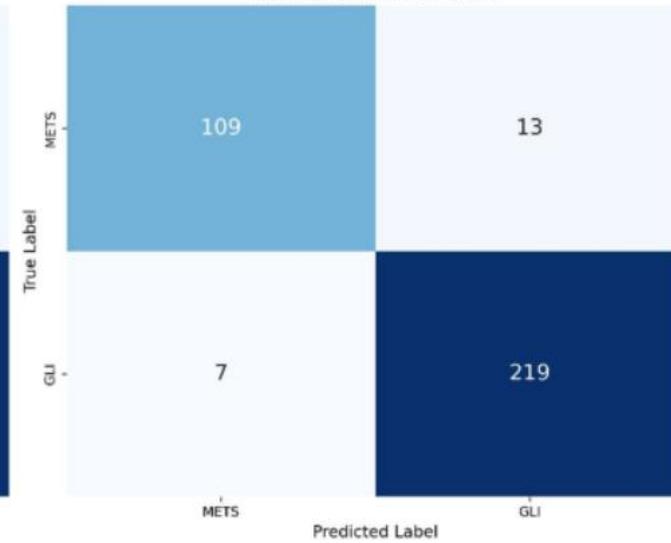
Metric	Random Forest	Support Vector Machine
Accuracy	0.9195 ± 0.0229	0.9425 ± 0.0120
AUC	0.9782 ± 0.0096	0.9833 ± 0.0072
Precision	0.9192 ± 0.0234	0.9425 ± 0.0137
Recall	0.9195 ± 0.0229	0.9425 ± 0.0126
F1-Score	0.9192 ± 0.0234	0.9422 ± 0.0126
NPV	0.9331 ± 0.0345	0.9439 ± 0.0256

External validation classification results

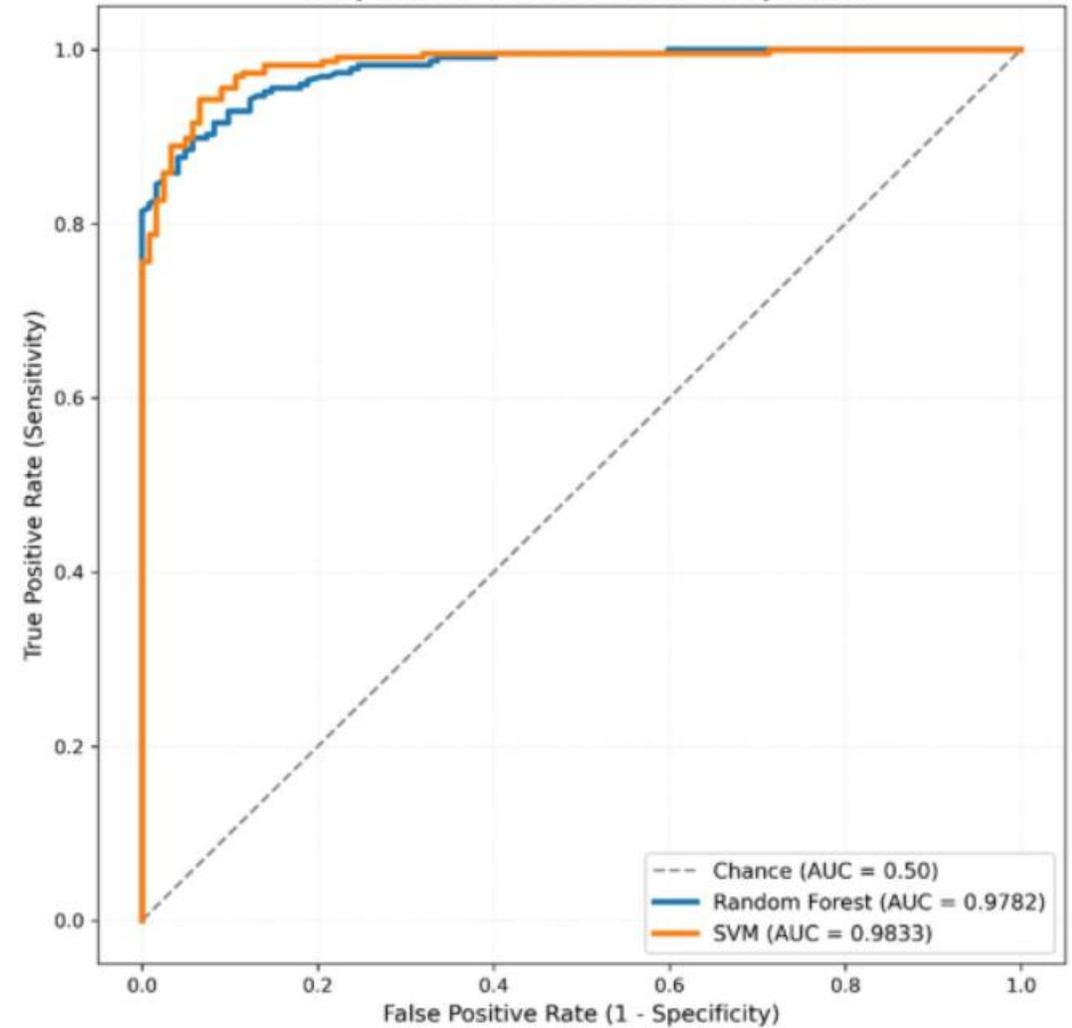
Confusion Matrix: Random Forest



Confusion Matrix: SVM

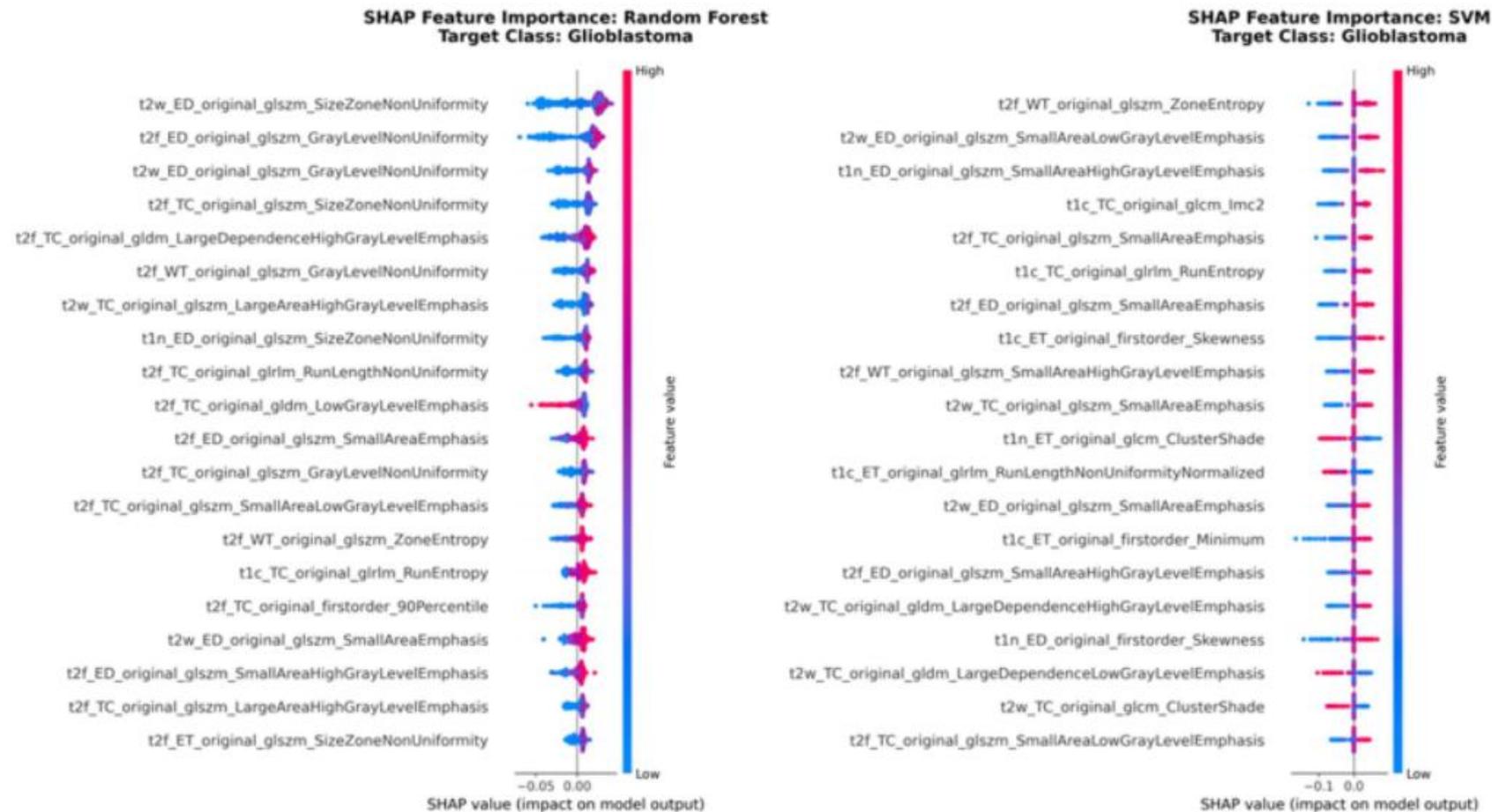


Independent Test Set ROC Comparison

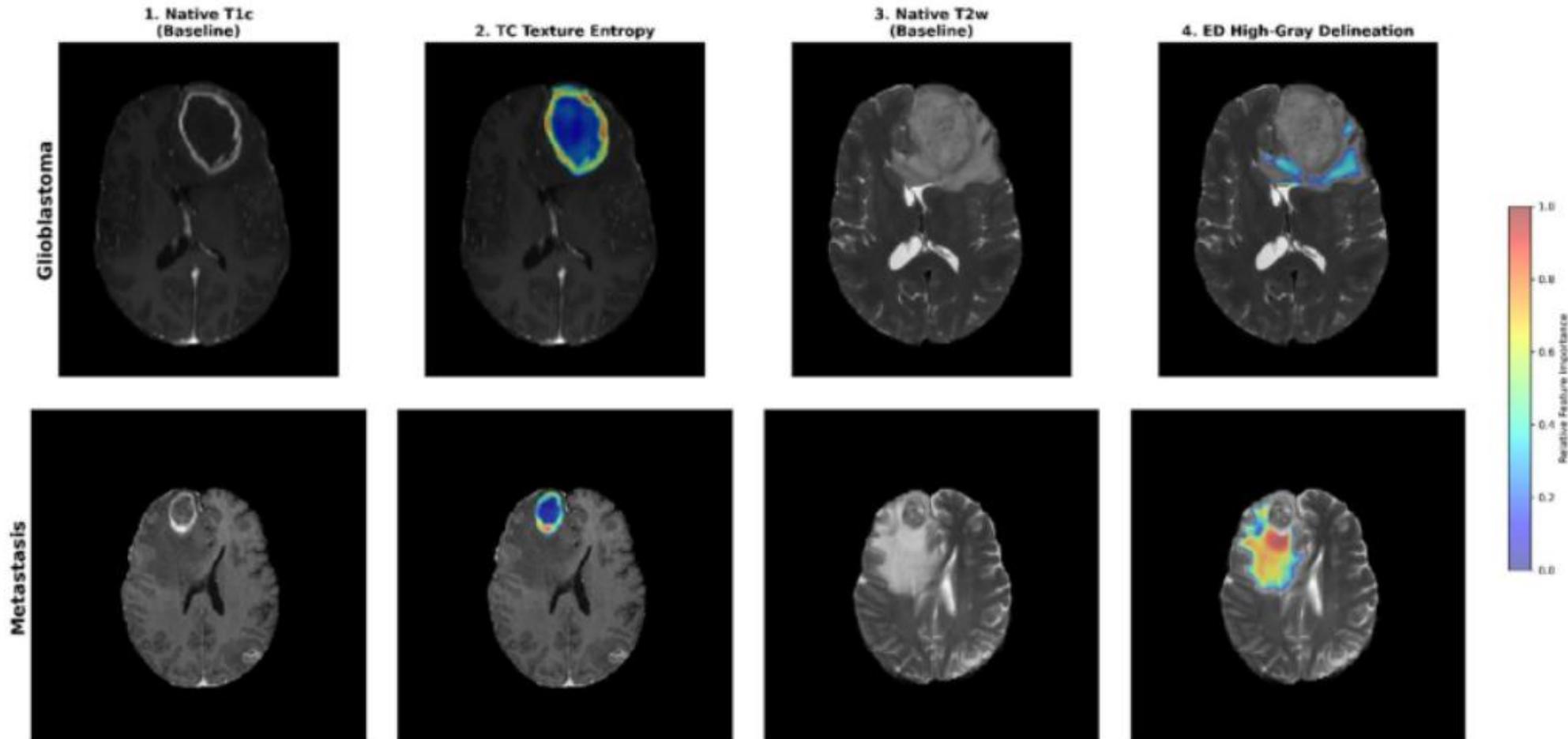


Feature importance analysis

Figure 7: Summary of SHAP values for RF (left) and SVM (right) models.



Heatmap analysis



Kitas semestras

Studijų planas:

- Optimizuoti ir standartizuoti vaizdų pirminio apdorojimo, radiomikos požymių atrankos ir dimensijų mažinimo procesus, užtikrinant išgaunamų duomenų stabilumą ir atkuriamumą (angl. reproducibility) skirtingose MRT sekose (galvos smegenų gliomų ir metastazių atvejais).
- Atlikti klasifikavimo algoritmų hiperparametrų optimizaciją, fokusuojantis į modelių tikslumo (angl. accuracy) ir jautrumo (angl. sensitivity) balansą bei interpretuojamumo didinimą naudojant XAI (angl. Explainable AI) metodus.
- Atlikti išsamų sukurtų klasifikavimo modelių validavimą naudojant testavimo duomenų rinkinį, siekiant įvertinti jų apibendrinamumą (angl. generalizability) diferencijuojant galvos smegenų gliomas ir metastazes.