



**Vilnius
universitetas**

**Vilnius
universitetas**

**Algoritmų, leidžiančių
nustatyti medžiagos
cheminę sudėtį, iš hiper-
spektrinių duomenų tyrimas**

Vytautas Paura

Darbo vadovas: dr. Virginijus Marcinkevičius

Darbo konsultantas: Dr. Valdas Rapševičius

Doktorantūros pradžios ir pabaigos metai: 2020–2024

Studijų planas

Studijų metai	Egzaminai		Dalyvavimas konferencijose		Publikacijos		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė
I (2020/2021)	2	0	0	0	0	0	
II (2021/2022)	2	0	0	0	0	0	
III (2022/2023)			1	0	1	0	
IV (2023/2024)			1	0	1	0	

Pusmečio planas

Egzaminai		Dalyvavimas konferencijose		Publikacijos	
Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta
Optimizavimo metodai ir jų taikymas	Planuoja ma išlaikyti iki 2021 rugsėjo mėn.	Dalyvavimas konferencijoje Lietuvoje	Dalyvavimas Open Readings 2021 http://www.openreadings.eu konferencijoje	Mokslinių tyrimų disertacijos tema apžvalga (konferencijos darbų medžiagoje)	

Disertacijos rengimo etapai

Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje):

1. Disertacijos tyrimo objekto detalizavimas.
2. Medžiagų cheminės sudėties nustatymo algoritmų apžvalga ir analizė.
3. Algoritmų naudojančių hiper-spektrinius duomenis medžiagų cheminei sudėčiai nustatyti apžvalga ir analizė.
4. Metodų analizės apibendrinimas ir mokslinių problemų susietų su tyrimo objektu identifikavimas, ir tyrimo tikslo ir uždavinių suformavimas.

Tyrimo objektas

Tyrimo objektas:

- Hiper-spektriniai vaizdai ir medžiagų cheminės sudėties iš hiper-spektrinių vaizdų nustatymo algoritmai.

Sprendžiamos problemos:

- Grynujų pikselių skaičius nustatytas iš pateikto hiper-spektrinio vaizdo ar duomenų rinkinio.
- Randant grynujų medžiagų spektrines žymes (angl. signatures).
- Suskaičiuojant atrastų grynujų medžiagų sumaišymo proporcijas visame hiper-spektriniame vaizde.

2020/2021 m. m. darbo planas

Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje)

- Disertacijos tyrimo objekto detalizavimas
- Medžiagų cheminės sudėties nustatymo algoritmų apžvalga ir analizė.
- Algoritmų naudojančių hiper-spektrinius duomenis medžiagų cheminei sudėčiai nustatyti apžvalga ir analizė.
- Metodų analizės apibendrinimas ir mokslinių problemų susietų su tyrimo objektu identifikavimas, ir tyrimo tikslo ir uždavinių suformavimas.

Išlaikyti egzaminus:

- „Optimizavimo metodai ir jų taikymas“
- „Informatikos ir informatikos inžinerijos tyrimo metodai ir metodika“

Publikacijų rengimas:

- Mokslinių tyrimų disertacijos tema apžvalga (konferencijos darbų medžiagoje)

Cheminės sudėties nustatymas

Sudėties nustatymo algoritmų tipai:

1. Neneigiamų matricų faktorizavimas.
2. Autoenkoderių tinklai.
3. Retų duomenų regresija.

Metrikos:

1. Vidutinės kvadratinės paklaidos šaknis (Root mean squared error - RMSE)
2. S rekonstravimo paklaida (Signal reconstruction error SRE)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2}$$

$$SRE = 10 \log_{10} \left(\frac{E[\|x\|_2^2]}{E[\|x - \hat{x}\|_2^2]} \right)$$

Cheminės sudėties nustatymo eksperimentas

Duomenų rinkiniai:

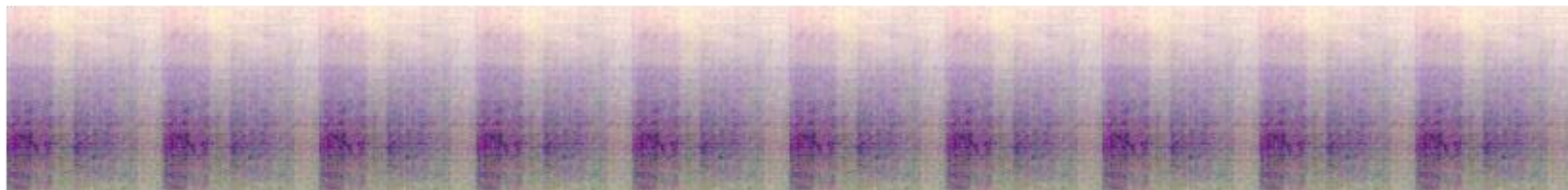
- Laboratorijoje sukurtas hiper-spektrinis kubas (Min Zhao et al., 2019)
- Dirbtinis kubas sudarytas iš USGS spektrinės bibliotekos duomenų (Raymond F. Kokaly et al., 2017)

Hiper-parametrų paieška:

- Kryžminė patikra
- Tinklelio paieška

Etaloninis eksperimentas:

1. Algoritmų atsparumas medžiagų kiekio pasikeitimams



Algoritmai naudoti eksperimente

Retų duomenų regresija:

- Sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL) [J. M. Bioucas-Dias and M. A. T. Figueiredo, 2010]
- SUnSAL and total variation (SUnSAL_TV) [J. M. Bioucas-Dias and M. A. T. Figueiredo, 2012]
- Spectral–spatial weighted sparse unmixing (S2WSU) [S. Zhang et al. 2018]
- Augmented linear mixing model (ALMM) [D. Hong et al., 2019]

Algoritmai naudoti eksperimente

Neneigiamų matricių faktorizavimas:

- Spatial group sparsity regularized NMF (SGSNMF) [X. Wang et al., 2017]
- Coupled nonnegative matrix factorization (CNMF) [N. Yokoya et al., 2012]
- Robust collaborative nonnegative matrix factorization (R-CoNMF) [J. Li et al., 2016]
- Total variation regularized reweighted sparse NMF (RSNMF) [W. He et al., 2017]

Autoenkoderių tinklai:

- Autoencoder network with adaptive abundance smoothing (AAS) [Z. Hua et al., 2020]

Išvados

- Tarp tirtų algoritmų nebuvo sukurta vieninga metodika šių algoritmų testavimui.
- Naudojamos metrikos galimai nėra patikimos ir yra sunkiai interpretuojamos.
- Skirtumai tarp rezultatų skirtingose metrikose silpnai koreliuoja.
- Algoritmai gauna skirtingus rezultatus priklausomai nuo naudojamo duomenų rinkinio.



**Vilniaus
universitetas**

Kontaktai

Akademijos g. 4

LT-08663 Vilnius

+370 6 256 79560

vytautas.paura@mif.stud.vu.lt