



**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	Dalyvavimas konferencijose				Publikacijos					
	Tarprautinėse		Nacionalinėse		Su citavimo rodikliu			Be citavimo rodiklio		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė	Planas	Įvykdyta	Būklė
I (2020/2021)			1	1				1	1	Publikuota
II (2021/2022)			1	1				1	1	Publikuota
III (2022/2023)	1	1			1	1	Publikuota			
IV (2023/2024)	1	1			1	0	Įteikta	1	1	Publikuota
Iš viso:	2	2	2	2	2	1		2	2	

Pusmečio ataskaita

Publikacijos IV (2023/2024) (I pusmetis)			
Pavadinimas	Įvykdyta	Būklė	Publikacijos tipas
Crop Hyperspectral Dataset Unmixing Using Modified U-Net Model		Įteikta konferencijoje: "Baltic DB&IS 2024" Publikuota: "Digital Business and Intelligent Systems", 2024	
Hyperspectral Unmixing of Agricultural Images taken from UAV Using Adapted U-Net Architecture		Įteikta žurnale: "Fundamenta Informaticae"	

Disertacijos rengimo etapai

Gautų duomenų analizė, apibendrinimas, išvadų parengimas:

1. Tikslų, uždavinių, tyrimo metodikos, ginamųjų teiginių patikslinimas.
2. Analitinės disertacijos dalies parengimas.
3. Teorinės disertacijos dalies parengimas.
4. Eksperimentinės disertacijos dalies parengimas.
5. Bendrųjų išvadų formulavimas.

Tyrimo objektas

Tyrimo objektas:

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

Sprendžiamos problemos:

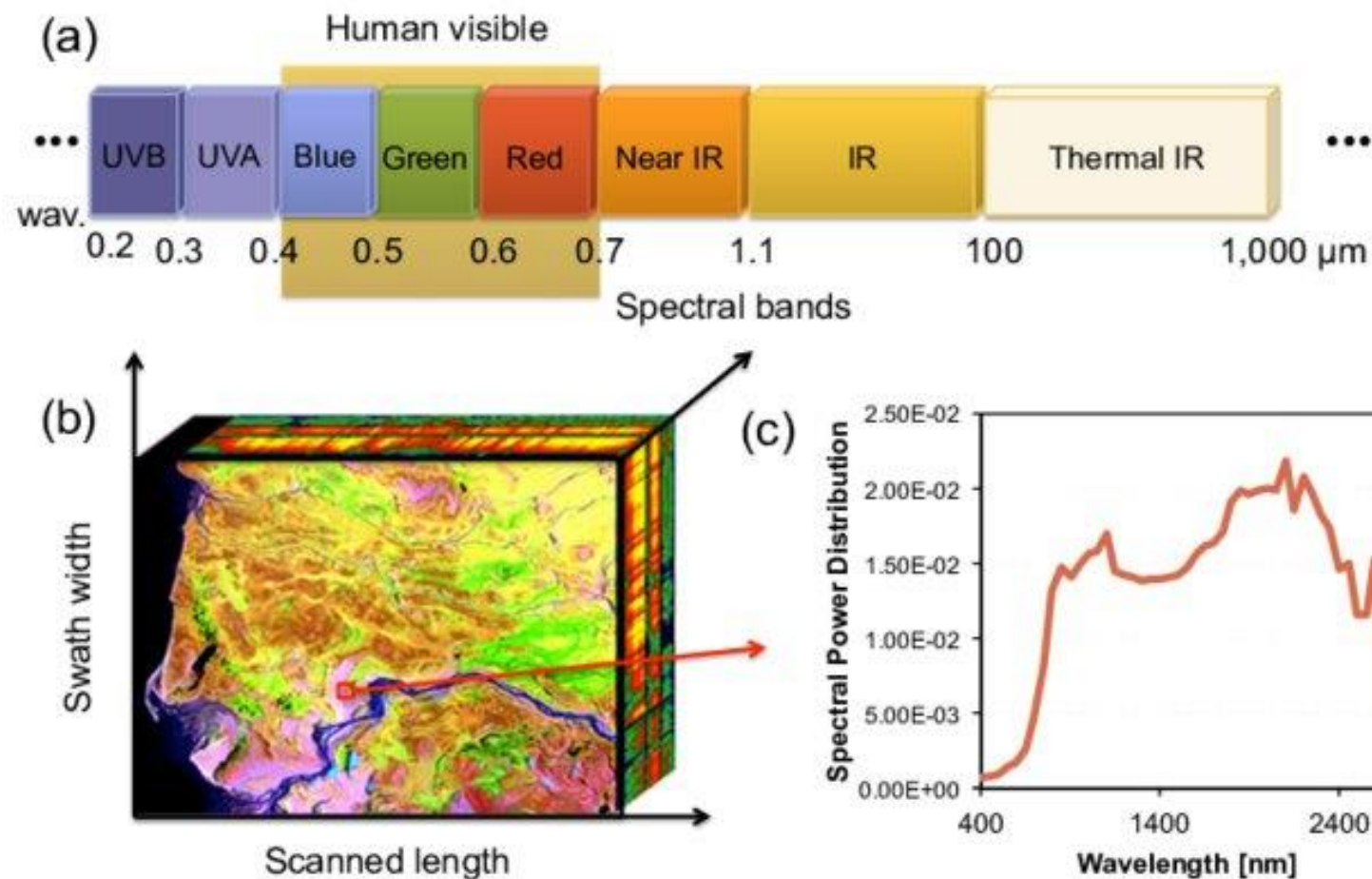
- Grynujų pikselių skaičius nustatytas iš pateikto hiper-spektrinio vaizdo ar duomenų rinkinio.
- Randant grynujų medžiagų spektrines žymes (angl. signatures).
- Suskaičiuojant atrastų grynujų medžiagų sumaišymo proporcijas visame hiper-spektriniame vaizde.

2023/2024 m. m. darbo planas

- Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje):
 - Hiper-spektrinių duomenų rinkinio, naudojamo cheminių medžiagų nustatymui, sudarymo metodikos kūrimas ir algoritmų optimizavimas.
 - Žemės ūkio srities ir pasėlių hiper-spektrinių duomenų rinkimas iš bepiločių orlaivių, jų apdorojimas ir paruošimas duomenų rinkinio sudarymui.
 - Hiper-spektrinių duomenų rinkinio, skirto naudoti medžiagų nustatymo algoritmams tirti, sudarymas.
 - Sukurtas ir tobulinamas medžiagų cheminės sudėties nustatymo algoritmas, paremtas U-Net tipo modelių architektūra ir kitais publikuotais giliųjų neuroninių tinklų algoritmais.
 - Medžiagų cheminės sudėties nustatymo algoritmų eksperimentinis tyrimas panaudojant naujai sukurtą hiper-spektrinių duomenų rinkinį.
- Publikacijų rengimas:
 - Hiper-spektrinių duomenų analizės algoritmo kūrimas naudojant savo sukurtus ir atvirus duomenų rinkinius.
 - Sukurto analizės algoritmo ir duomenų rinkinių publikavimas (recenzuojamame leidinyje, CA WoS su Impact Factor).

Tyrimo objektas

Hiper-spektrinių
vaizdų analizė
specializuojantis į
žemės ūkio
duomenis pasėlių
stebėjimui.



Tyrimo objektas

Cheminių medžiagų nustatymo iš hiper-spektrinių vaizdų algoritimų dalys:

- Grynųjų medžiagų skaičiaus nustatymas hiper-spektriniame vaizde;
- Grynųjų medžiagų spektrinių žymių nustatymas;
- Rastų medžiagų, susimaišymo proporcijų pikseliuose, matricos radimas.

Dažniausia naudojamų algoritimų tipai:

- Neneigiamų matricų faktorizavimas.
- Autoenkoderių tinklai.
- Retų duomenų regresija.

Sprendžiamos problemos

- Cheminių medžiagų nustatymo algoritmų taikymas žemės ūkio tipo duomenims.
- Atvirų hiper-spektrinių duomenų trūkumas.
- Giliųjų neuroninių tinklų panaudojimas cheminių medžiagų nustatymui nelaboratoriniuose duomenyse.

Etaloninis eksperimentas

Etaloninio eksperimento dalys:

1. Algoritmų atsparumas skirtingų medžiagų kiekiui hiper-spektriniame vaizde.
2. Algoritmų atsparumas triukšmui.
3. Algoritmų gebėjimas atskirti medžiagas priklausomai nuo įvesties hiper-spektrinių duomenų kiekio.

Etaloninis eksperimentas

Retoji 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]

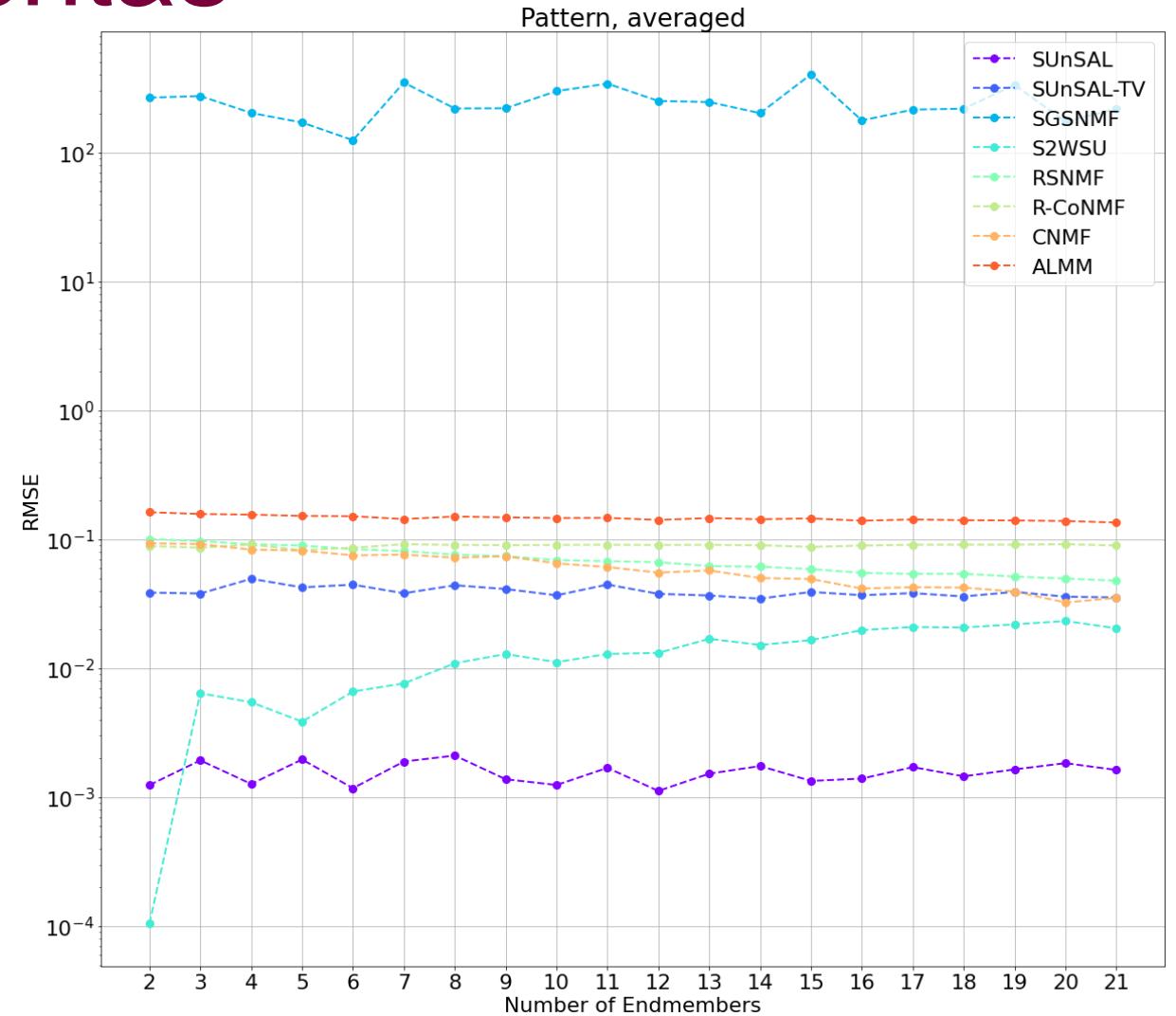
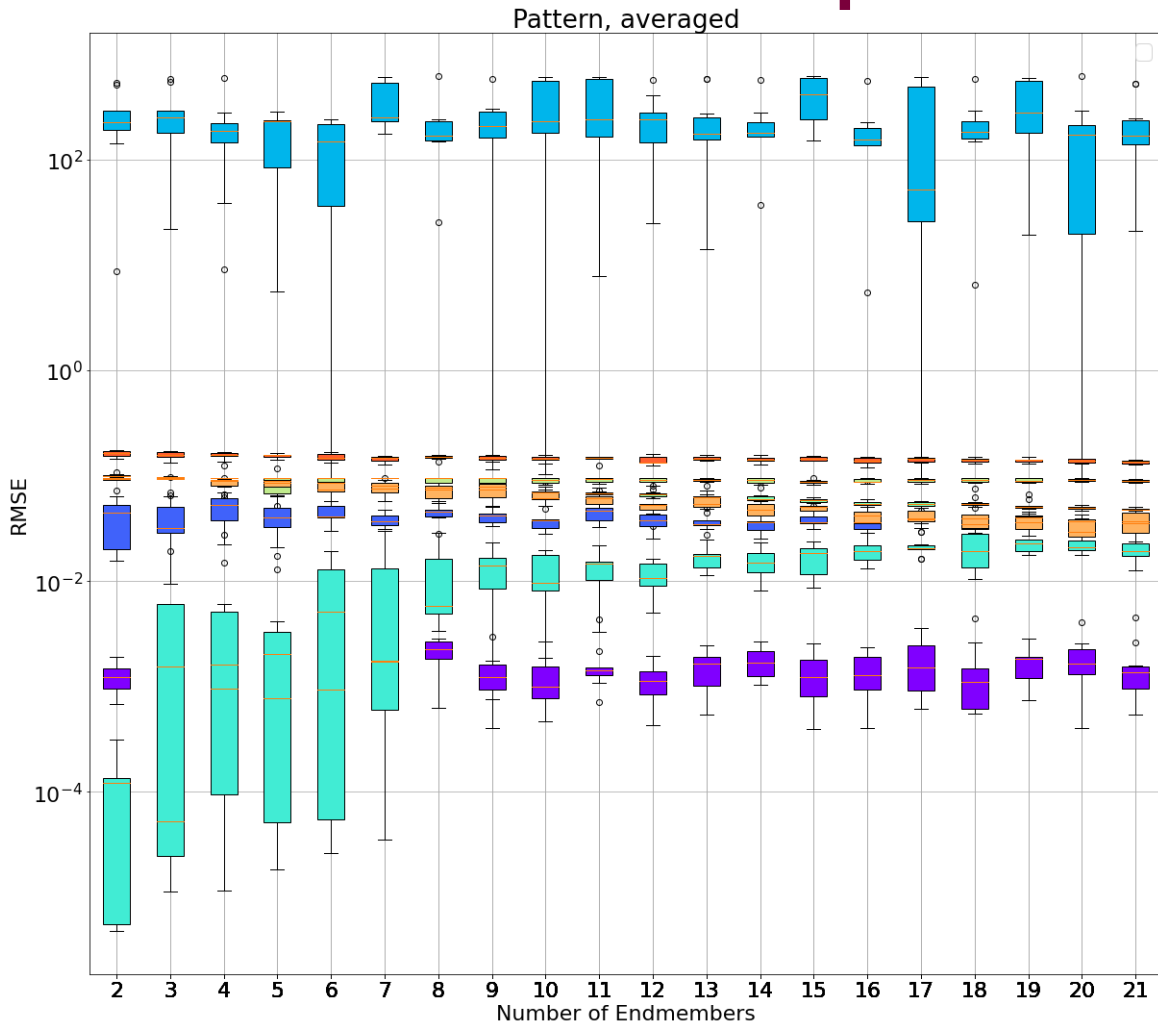
NMF:

- 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]

Autoencoderių tinklai:

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

Etaloninis eksperimentas



Hiper-spektrinių duomenų rinkinys

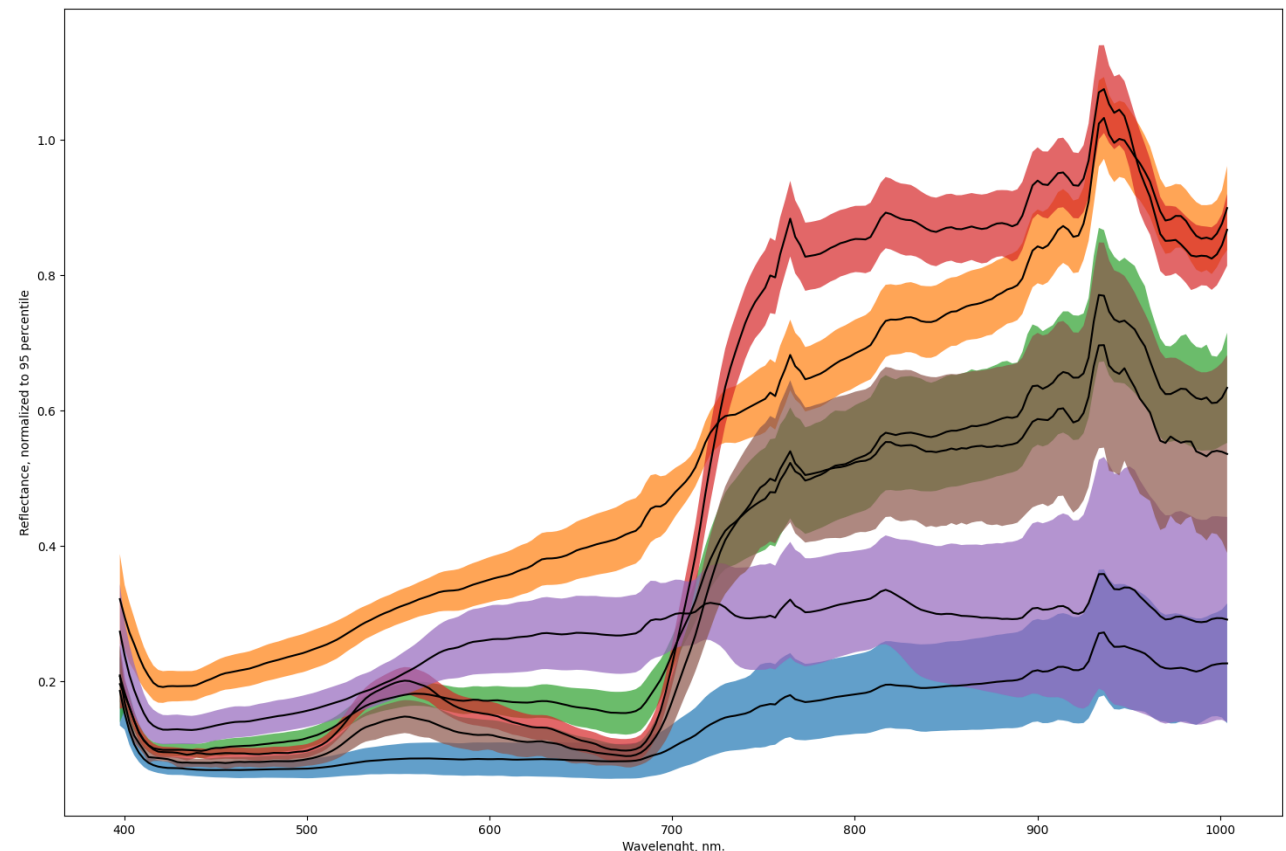
- Duomenų rinkinio sudarymui surinkti trys hiper-spektriniai kubai.
- Kubų dydžiai yra apie 3000 x 1024 pikselių ir 224 spektriniai ruožai.
- Spektriniai ruožai fiksuoti sensoriaus apima nuo 400 iki 1000 nanometrų ilgio bangas.
- Atlikta kalibracija iš šviesos intensyvumo vienetų į šviesos atspindžio procentus.



Hiper-spektrinių duomenų rinkinys

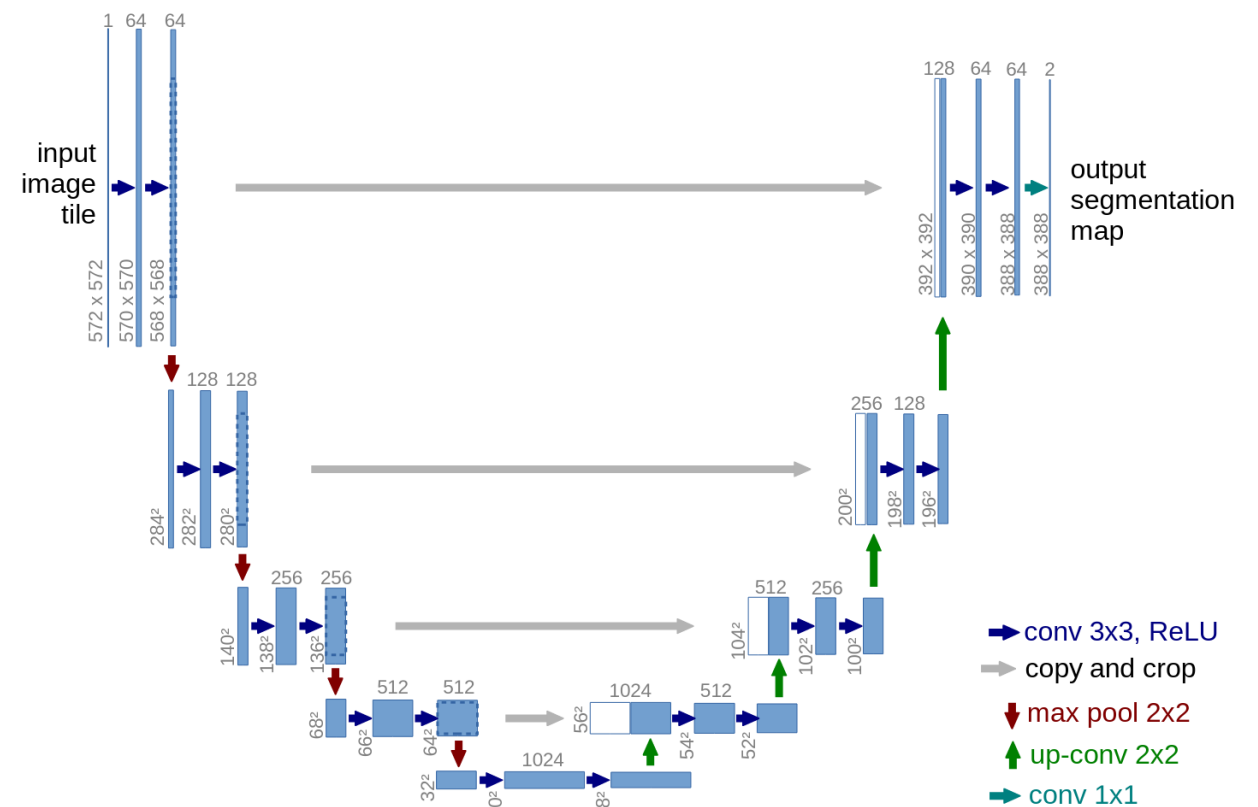
- Pasirinktas VCA algoritmas nustatyti pradinėms klasėms (nustatyti medžiagoms).
- Pagal atliktus eksperimentus su skirtingu klasių kiekiu atrinktu VCA, buvo pasirinkta 6 klasės.
- Visas vaizdas buvo klasifikuotas pagal mažiausią RMSE vertę.
- Kiekvienai klasei nubraižytas vidurkio ir standartinio nuokrypio grafikas.
- Pagal šias klases sukurtas trijų hiper-spektrinių vaizdų duomenų rinkinys.

<https://doi.org/10.5281/zenodo.13856357>



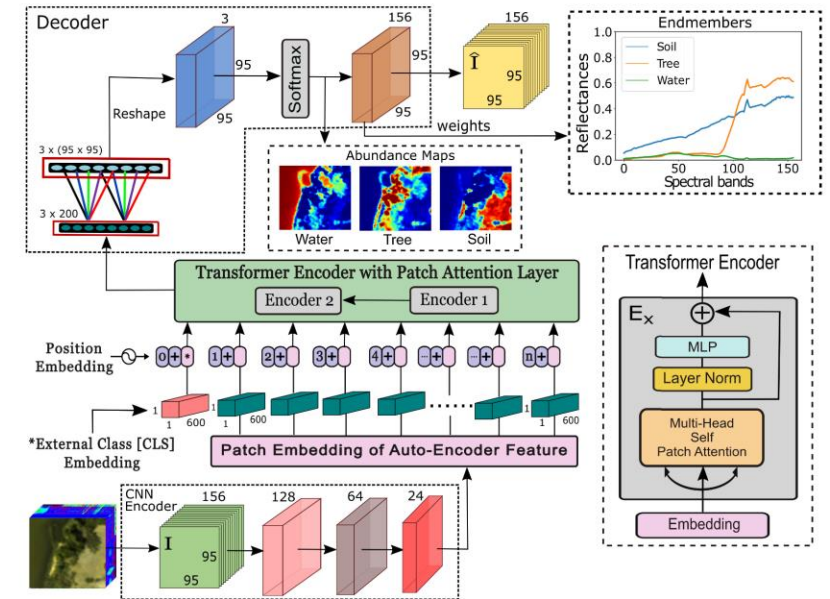
Hiper-spektrinių duomenų analizės algoritmas

- U-Net modelis originaliai skirtas medicinoje vaizdų analizei.
- Šio tipo modelis pasirinktas dėl jo galimybės įsisavinti informaciją skirtingose duomenų skalėse (dėl duomenų suspaudimo ir konvoliucijos sluoksnių).
- Modelis modifikuotas išskirti medžiagų ir jų susimaišymo proporcijų matricas.

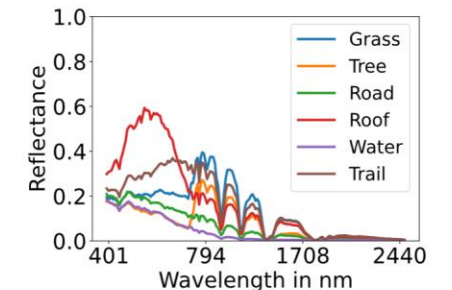
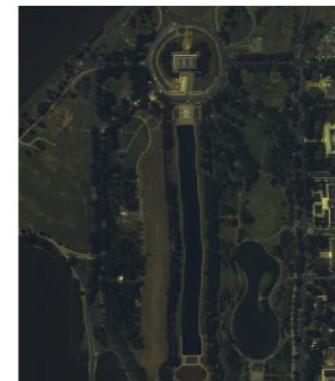


Eksperimentas

- Palyginamuoju modeliu pasirinktas aukštas tikslumo rezultatus pasiekęs transformerių tinklais paremtas modelis.
- Sukurtas ir palyginamasis modeliai buvo testuoti ant šešių hiper-spektrinių kubų: DC Mall, Samson, Apex ir trys kubai iš naujai sukurto duomenų rinkinio.



Šaltinis: <https://doi.org/10.48550/arXiv.2203.17076>



Šaltinis: <https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html>

Eksperimento rezultatai

	Proposed model				Transformer model			
	mRMSE	mSAD	RE	Epochs	mRMSE	mSAD	RE	Epochs
Apex	0.4705	0.1737	0.0990	<i>1001</i>	0.5555	0.2025	0.1048	<i>1000</i>
DC	0.3971	0.3764	0.0480	<i>1001</i>	0.3918	0.3009	0.0232	<i>1000</i>
Samson	0.4301	0.1507	0.0526	<i>1001</i>	0.6031	0.2400	0.1675	<i>1000</i>
Blueberry Cube 1	0.3112	0.2737	0.0752	<i>3001</i>	0.4845	0.3951	0.3012	<i>1000</i>
Blueberry Cube 2	0.3740	0.2591	0.1263	<i>3001</i>	0.4511	0.4012	0.2860	<i>1000</i>
Blueberry Cube 3	0.3088	0.2214	0.0978	<i>3001</i>	0.4232	0.3852	0.2645	<i>1000</i>

Išvados

- Sukurtas hiper-spektrinių duomenų rinkinys iš duomenų surinktų šilauogių laukuose.
- Sukurtas U-Net tinklu paremtas medžiagų kiekio hiper-spektriniuose duomenyse nustatymo algoritmas.
- Sukurto algoritmo rezultatas pagal RE ir RMSE metrikas beveik visada geresnis nei palyginamojo.
- Pagal atliktus tyrimus algoritmo tikslumui didelę įtaką daro pasirinktas duomenų rinkinys ir hiper-parametrų rinkinys pritaikytas duomenims.
- Transformeriais paremtas tinklas reikalauja mažiau epochų apmokymui, bet daugiau atminties resursų.



**Vilnius
universitetas**

Kontaktai

Akademijos g. 4

LT-08663 Vilnius

+370 6 256 79560

vytautas.paura@mif.stud.vu.lt