



**Vilniaus Universitetas**  
**Matematikos ir informatikos institutas**  
**LIETUVA**



---

INFORMATIKOS INŽINERIJA (07 T)

---

**Ataskaita už 2016-2017m.**  
**mokslo metus**

Doktorantas Mindaugas Morkūnas

2017, Vilnius

# Bendra informacija

- **Disertacijos pavadinimas:** Modelis navikų genetinių pokyčių ir biožymenų raiškos duomenų integravimui
- **Disertacijos vadovas:** dr. Povilas Treigys
- **Doktorantūros pradžia:** 2016 m.
- **Planuojama doktorantūros pabaiga:** 2020 m.

# Informacija apie tyrimą

- **Tyrimo objektas:**
  - Skaitmeniniai pilno pjūvio patologijos vaizdai.
- **Tyrimo tikslas:**
  - Pasiūlyti naujus būdus vėžio ląstelių subpopuliacijų atpažinimui ir heterogeniškumo vertinimui.
- **Tyrimo uždaviniai:**
  - Sukurti ir iširti naujus skaitmeninių patologijos vaizdų analizės metodus integruojant genetinius duomenis.
  - Sukurti programų sistemos prototipą, kuriame būtų įgyvendinti pasiūlyti vėžio ląstelių subpopuliacijų atpažinimo ir heterogeniškumo vertinimo būdai.
- **Planuojami rezultatai:**
  - Nauji vėžio ląstelių subpopuliacijų atpažinimo, jų heterogeniškumo vertinimo būdai ir juos įgyvendinantis programų sistemos prototipas.

# 2016 – 2017 metų darbo planas

## ○ **Studijų planas:**

- Išlaikyti egzaminą „Informatikos inžinerijos matematiniai metodai“
- Išlaikyti egzaminą „Daugiamačių duomenų vizualizavimo metodai“

## ○ **Mokslinių tyrimų planas:**

- Disertacijos tyrimo objekto detalizavimas;
- Navikų genetinio heterogeniškumo (kloninės struktūros) predikcijos metodų apžvalga
- Skaitmeninių patologijos vaizdų analizės (navikų ląstelių subpopuliacijų identifikavimo, audinių atpažinimo) metodų apžvalga

# 2016 – 2017 metų darbo planas

- **Mokslinių tyrimų publikavimas:**
  - Mokslinių tyrimų disertacijos tema apžvalga (recenzuojamame leidinyje)
- **Pranešimai seminaruose ir konferencijose:**
  - Disertacijos tyrimo objekto pristatymas konferencijoje Lietuvoje

# Ataskaita už 2016-2017 metus

- **Išlaikyti egzaminai:**

- **„Informatikos inžinerijos matematiniai metodai“**

Vertinimo komisija: prof. habil. dr. Laimutis Telksnys, prof. habil. dr. Kazys Kazlauskas, dr. Gintautas Tamulevičius

- **„Daugiamačių duomenų vizualizavimo metodai“**

Vertinimo komisija: prof. habil. dr. Gintautas Dzemyda, doc. dr. Olga Kurasova, dr. Viktor Medvedev

# Ataskaita už 2016-2017 metus

- **Disertacijos rengimas:**

- Atlikta navikų ląstelių subpopuliacijų identifikavimo, audinių atpažinimo, heterogeniškumo vertinimo metodų apžvalga

- **Straipsniai:**

- Parengta ir žurnalo „Informatica“ redakcijai įteikta publikacija „Machine Learning Based Classification of Colorectal Cancer Tumor Tissue in Whole-Slide Images“

# Ataskaita už 2016-2017 metus

## ○ Pranešimai konferencijose:

- „8th International Workshop of Data Analysis Methods for Software Systems“. 2016-12-01/03 Druskininkai, Lietuva. M.Morkūnas, P.Treigys, A. Laurinavičius. **„An Overview of Methods to Spatially Map Intra-tumor Genetic Heterogeneity in Whole Slide Pathology Images“**. Stendinis pranešimas.
- „NEUBIAS2020: network of european bioimage analysts“, Lisbon, Portugal 15-17.02.2017. M.Morkūnas, P.Treigys, A.Laurinavičius, J.Bernatavičienė. **„Whole-slide Pathology Images Spatial Mapping of Intra-tumor Genetic Heterogeneity“**. Stendinis pranešimas.



# Ataskaita už 2016-2017 metus

## ○ Pranešimai konferencijose:

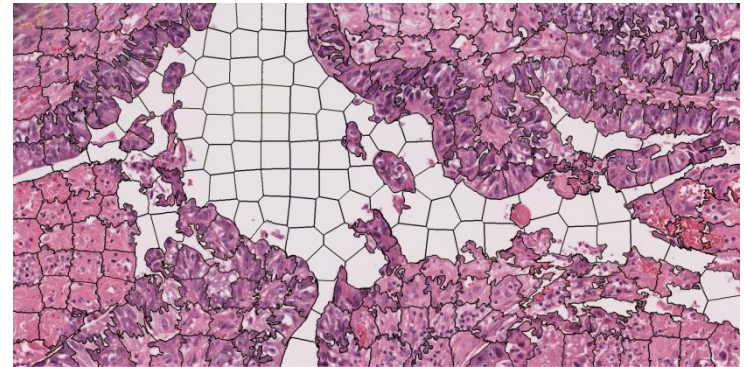
- „9th European Regional Conference on Thoracic Oncology“. 2017-06-16/17, Vilnius, Lietuva. M.Morkūnas. **„Intra-tumor Genetic Heterogeneity of Lung Adenocarcinoma as Investigated by Next Generation Sequencing“**. Žodinis pranešimas.
- „Kompiuterininkų dienos – 2017“. 2017-09-21/22, Kaunas, Lietuva. M.Morkūnas, P.Treigys, A. Laurinavičius. **„Machine Learning Based Classification of Colorectal Cancer Tumor Tissue in Whole-Slide Images“**. Žodinis pranešimas.

# Įvadas

- Daugumai vėžio tipų būdinga kloninė evoliucija ir dėl jos atsirandantis navikinio audinio heterogeniškumas.
- Naviko heterogeniškumas - skirtingai mutuotos ląstelių subpopuliacijos navike pasižymi skirtingomis funkcijomis ir morfologija.
- Navikinio audinio heterogeniškumo problema yra kritinė diagnozuojant ligą, skiriant gydymą.
- Baltyminių biožymenų bei šviesinės mikroskopijos pagalba tam tikromis biologinėmis savybėmis pasižyminčias naviko ląstelių subpopuliacijas galima atvaizduoti aukštos raiškos skaitmeniniuose vaizduose.
- Skaitmeninėje patologijos analizėje viena iš pirmųjų užduočių ir dažnu atveju tarpinis tyrimo tikslas yra atskirų audinio grupių išskyrimas naviko vaizde.

# Naviko audinių grupių klasifikavimas

- Naviko audinių požymių klasifikavimas
  - Naviko zonos pilno pjūvio vaizduose sudalinamos į superpikslius



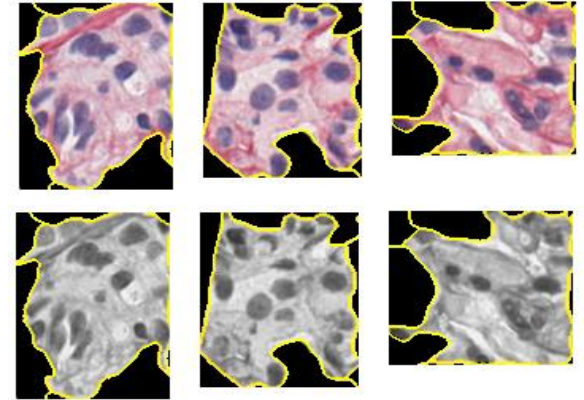
- Superpikslių kontūrai atvaizduojami naviko vaizde ir anotuojami eksperto

	Epitelinis audinys	Jungiamasis audinys
Apmokymo rinkinys	28000	28000
Validavimo rinkinys	2000	2000
Tikrinimo rinkinys	7068	3929

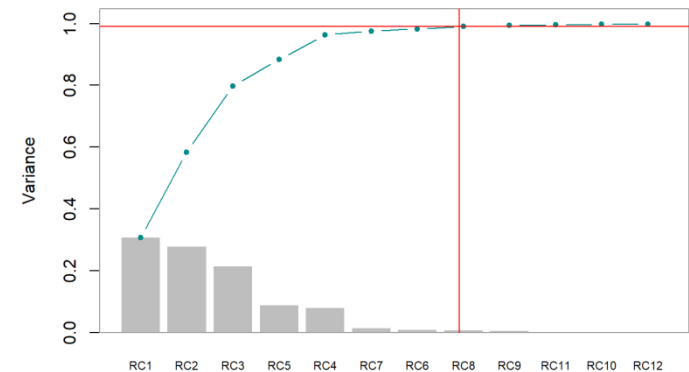
# Naviko audinių grupių klasifikavimas

- Naviko audinių požymių klasifikavimas

- Iš kiekvieno superpikselio išskiriami tekstūros ir spalvos požymiai



- Naudojant pagrindinių komponentų analizę sumažinamas superpikselių požymių rinkinių dimensijų skaičius.



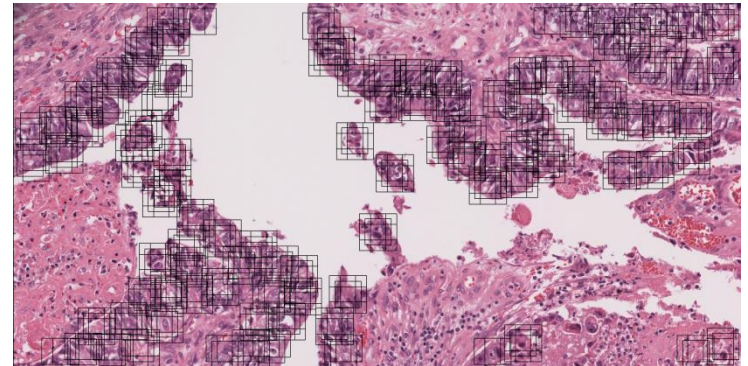
# Naviko audinių grupių klasifikavimas

- Naviko audinių požymių klasifikavimas
  - Superpikseliai klasifikuojami pagal išskirtus požymius naudojant skirtingus klasifikavimo algoritmus.

Superpikselių paruošimo būdas	Klasifikavimo algoritmas	Požymiai		
		Spalva	Spalva ir tekstūra	Tekstūra
2D „ištiesinti“ superpikseliai	RDF	0,9569	0,9577	0,9367
	SVM	0,9611	0,9613	0,9498
	MLP	<b>0,9668</b>	<b>0,9660</b>	0,9591
	DL	0,9608	0,9580	0,9397
2D RGB superpikseliai	RDF	0,9466	0,9533	0,9385
	SVM	0,9515	0,9620	0,9488
	MLP	0,9577	<b>0,9666</b>	0,9542
	DL	0,9556	0,9575	0,9449
2D pilki superpikseliai	RDF	0,7083	0,8553	0,8694
	SVM	0,7817	0,8688	0,8861
	MLP	0,7967	0,8765	0,8885
	DL	0,7967	0,8694	0,8777

# Naviko audinių grupių klasifikavimas

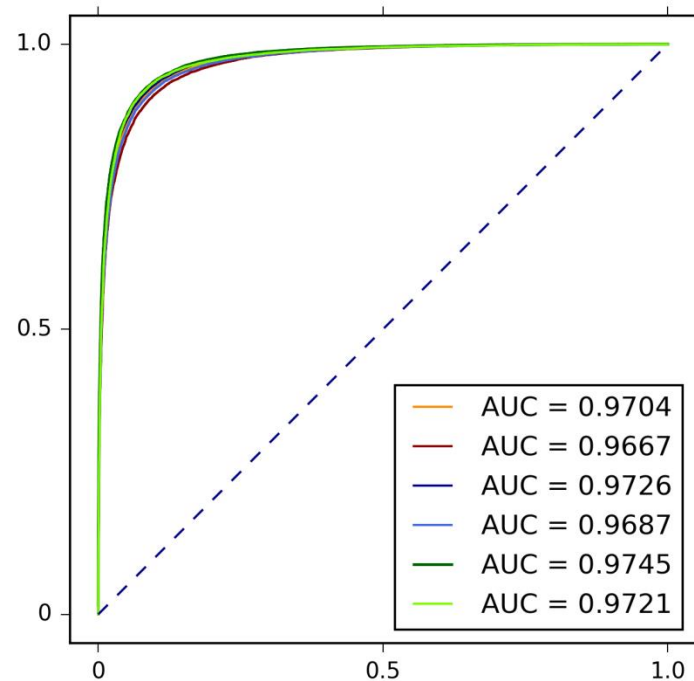
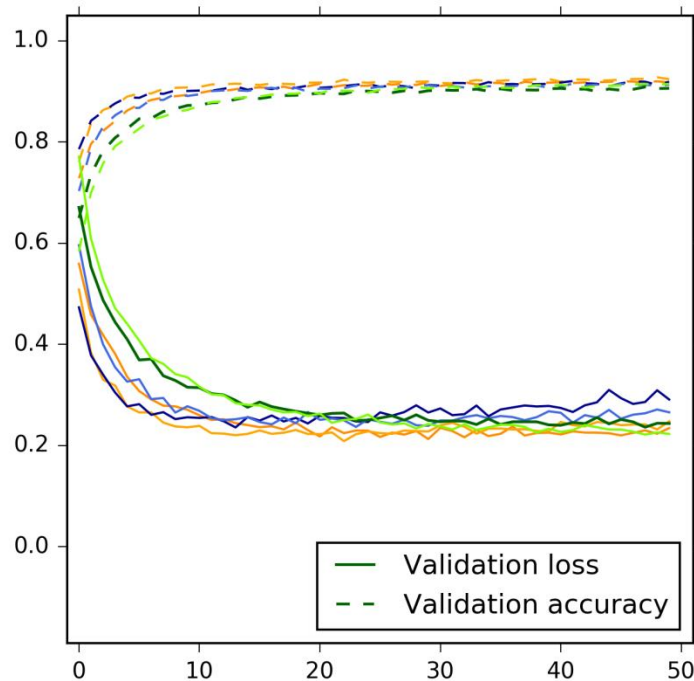
- Naviko audinių klasifikavimas konvoliucinio tinklo pagalba
  - Vaizdo blokai tinklo apmokymui parenkami apie superpikselių masės centrus.
  - Iš kiekvieno superpikselio gaunami 5 persidengiantys 72 x 72 pikselių dydžio vaizdo blokai.



	Epitelinis audinys	Jungiamasis audinys
Apmokymo rinkinys	40000	40000
Validavimo rinkinys	25000	5000
Tikrinimo rinkinys	26000	6000

# Naviko audinių grupių klasifikavimas

- Naviko audinių klasifikavimas konvoliucinio tinklo pagalba
  - Gauti vaizdo blokai naudojami apmokyti konvoliucinį neuroninį tinklą.



# Naviko audinių grupių klasifikavimas

- Naviko audinių klasifikavimas konvoliucinio tinklo pagalba
  - Skirtingų architektūrų konvoliucinio neuroninio tinklo rezultatai.

AUC	conv2d_1	conv2d_2	conv2d_3	conv2d_4	dense_1
<b>0,9704</b>	32	32	64	64	384
<b>0,9667</b>	32	32	64	64	1024
<b>0,9727</b>	64	64	128	128	384
<b>0,9687</b>	64	64	128	128	1024
<b>0,9745</b>	<b>96</b>	<b>96</b>	<b>192</b>	<b>192</b>	<b>384</b>
<b>0,9721</b>	96	96	192	192	1024



# Tiksliai anototų duomenų poreikis skaitmeninėje patologijoje



# Išvados

- Aukščiausias superpikselių klasifikavimo tikslumas pasiektas naudojant daugiasluoksnį perceptroną spalvotiems superpikseliams (AUC =  $\sim 0.967$ ). Kitų autorių pasiekti rezultatai tam pačiam vėžio tipui yra labai artimi (AUC = 0,97 - 0,99).
- Aukščiausias pasiektas konvoliucinio neuroninio tinklo tikslumas (AUC =  $\sim 0.975$ ) yra artimas kitų autorių publikuotiems rezultatams gautiems naudojant kitų tipų navikų duomenis - krūties vėžio (AUC = 0.9316) ir prostatos vėžio (AUC = 0.965 - 0.99).

# 2017-2018 metų darbo planas

## ○ **Studijų planas:**

- Išlaikyti egzaminą „Skaitmeninis vaizdų apdorojimas“
- Išlaikyti egzaminą „Duomenų analizės strategijos ir sprendimų priėmimas“

## ○ **Mokslinių tyrimų planas:**

- Naviko kloninės struktūros predikcijos iš genetinių duomenų metodų tyrimas
- Skaitmeninės patologijos vaizdų analizės naviko ląstelių subpopuliacijoms metodų tyrimas

## ○ **Mokslinių publikacijų planas:**

- Publikacija viename recenzuojamame periodiniame leidinyje

## ○ **Rezultatų pristatymo planas:**

- Dalyvavimas bent vienoje konferencijoje Lietuvoje

Ačiū už dėmesį