

Магістерська
кваліфікаційна робота
на тему:

“РОЗРОБКА
АЛГОРИТМІЧНОГО ТА
ПРОГРАМНОГО
ЗАБЕЗПЕЧЕННЯ МЕТОДУ
ВИПАДКОВИХ ЛІСІВ В
ЗАДАЧАХ КЛАСИФІКАЦІЇ
ЗОБРАЖЕНЬ”

Керівник: к.т.н.,
ст.викл. Маслій Р.В.

Виконала:
ст. гр. 1КСУ-14мн
Філіпчук О.Ю.

Метою дослідження є підвищення вірогідності класифікації зашумлених зображень та зменшення часу на навчання класифікатора при використанні випадкових лісів.

Об'єкт досліджень – процеси навчання класифікатора та класифікація зображень на основі випадкових лісів.

Предмет досліджень – метод випадкового лісу, що забезпечує збільшення вірогідності класифікації.

Практична цінність роботи полягає у розробці алгоритмічного та програмного забезпечення методу класифікації зображень на основі методу випадкових лісів.

Наукова новизна: запропоновано новий метод класифікації зображень, який на відміну від існуючих використовує випадкові ліси та ЛБШ у якості дескрипторів зображення, що дозволив поліпшити класифікацію зображення для гаусівського білого шуму у діапазоні $[0,01; 0,03]$, для імпульсного шуму у діапазоні $[0,04; 0,06]$ та для мультиплікативного шуму - $[0,03; 0,05]$.

Задачі дослідження:

- 1) розглянути існуючі методи класифікації даних;
- 2) оцінити переваги та недоліки реалізації розглянутих методів для досягнення поставленої мети;
- 3) огляд методу випадкового лісу і можливості його застосування для класифікації зображень;
- 4) огляд та вибір дескрипторів зображень;
- 5) розробка моделі перетворення зображень;
- 6) розробка методу класифікації зображень з використанням методу випадкового лісу;
- 7) програмна реалізація розробленого методу;
- 8) дослідження розробленого методу при використанні зображень Aerial Image Classification;
- 9) дослідження розробленого методу класифікації на зашумлених зображеннях при використанні Aerial Image Classification;

Класифікація – впорядковане по деякому принципу множина об'єктів, які мають подібні класифікаційні ознаки (одне або декілька властивостей), обраних для визначення схожості або відмінності між цими об'єктами.

Мета процесу класифікації полягає в тому, щоб побудувати модель, яка використовує прогнозовані атрибути в якості вхідних параметрів і отримує значення залежного атрибута. Процес класифікації полягає в розбитті множини об'єктів на класи за певним критерієм.

Класифікатором називається деяка сутність, яка визначає, якому з визначених класів належить об'єкт по вектору ознак.

Для проведення класифікації за допомогою математичних методів необхідно мати формальний опис об'єкта, яким можна оперувати, використовуючи математичний апарат класифікації. Таким описом в нашому випадку виступає зображення. Кожний об'єкт (зображення) несе інформацію про деякі властивості об'єкта.

Існує досить велика кількість методів щодо розв'язання задачі класифікації, зокрема, основні з них:

- класифікація за допомогою дерев рішень;
- класифікація методом випадкових лісів;
- класифікація методом AdaBoost;
- класифікація за допомогою штучних нейронних мереж;
- класифікація методом опорних векторів;
- класифікація за допомогою методу найближчого сусіда.

Приклад випадкового лісу

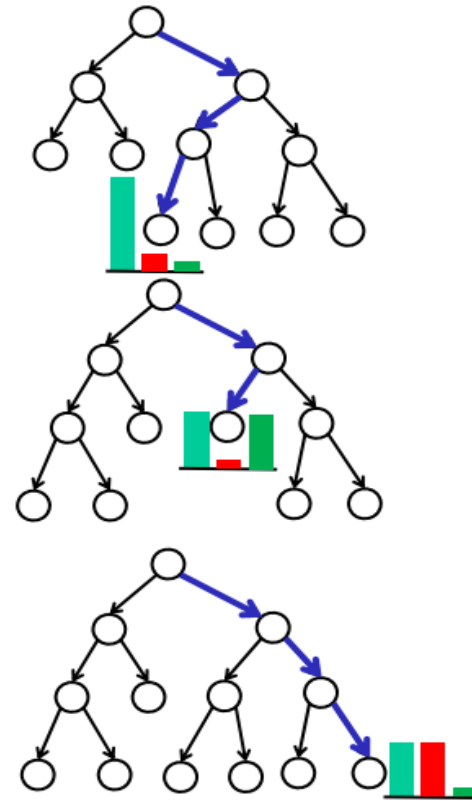
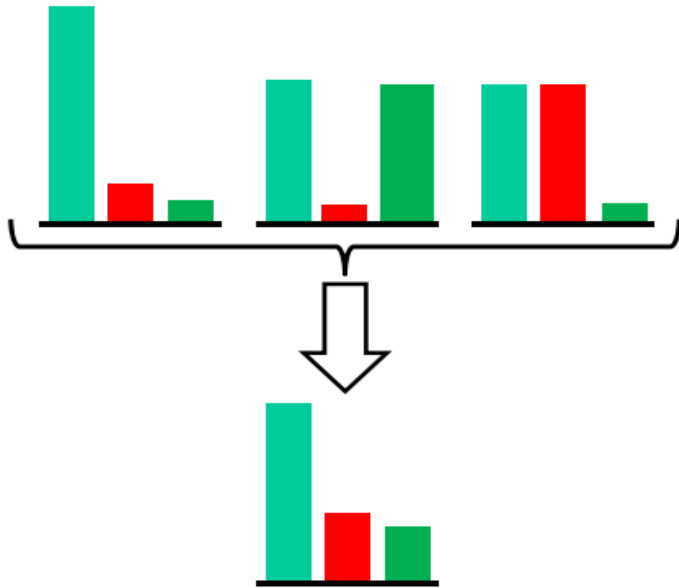


Схема програми отримання гистограми ЛБШ



Схема роботи системи

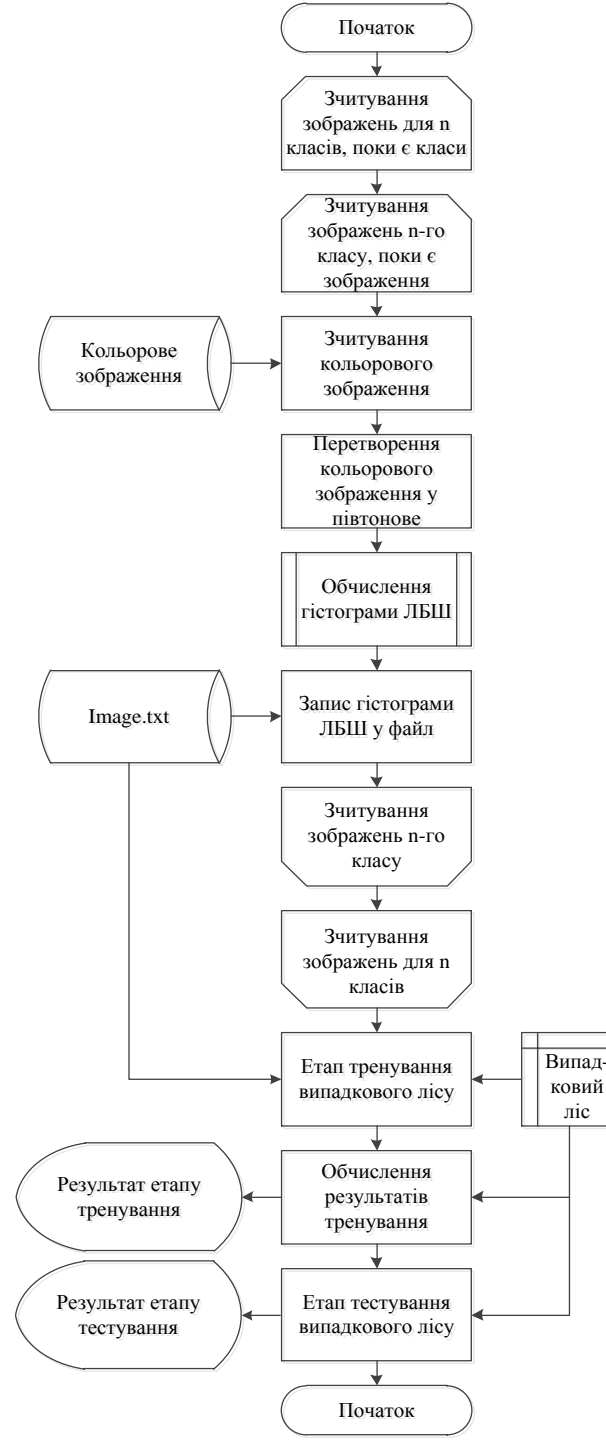
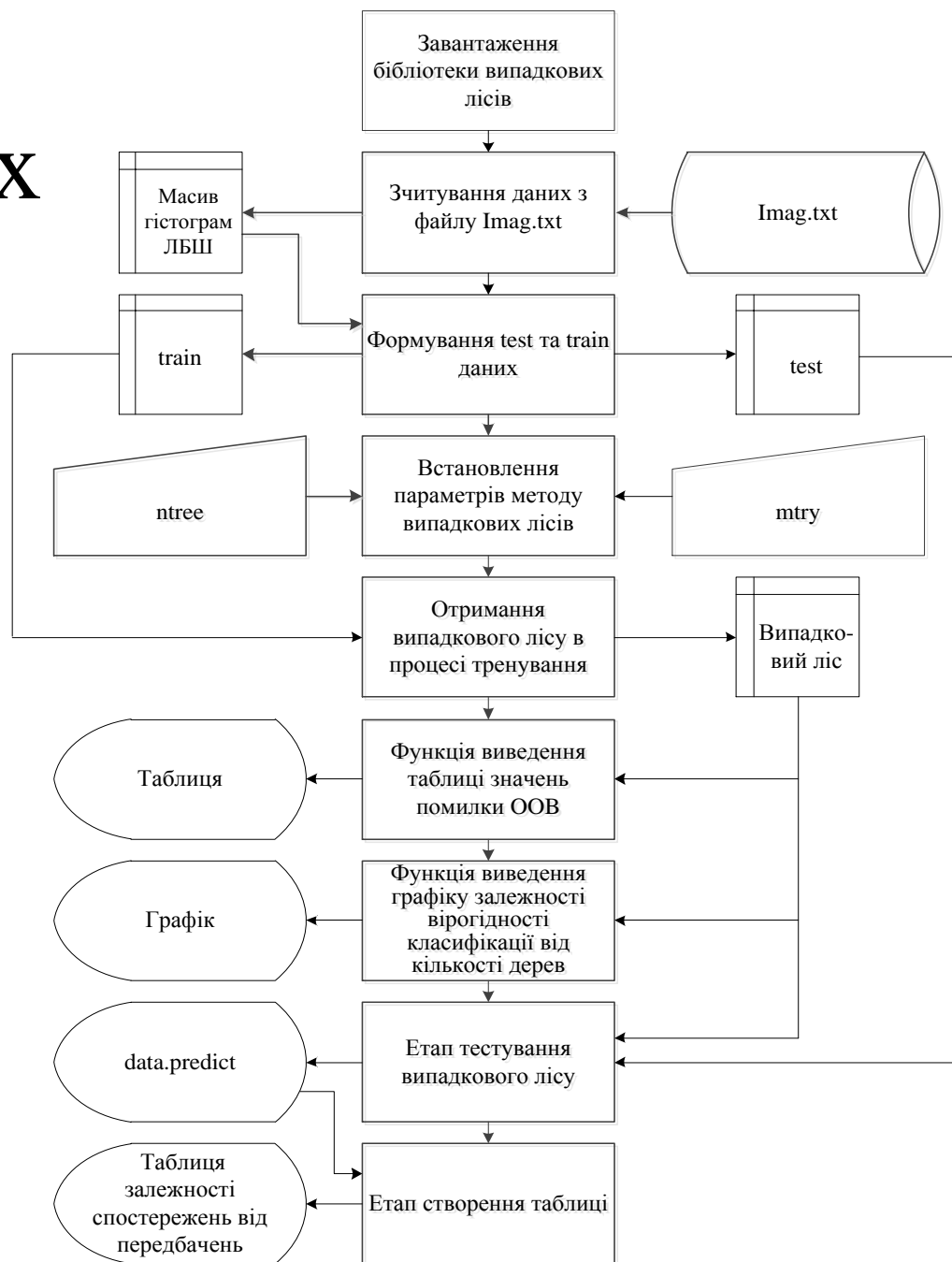
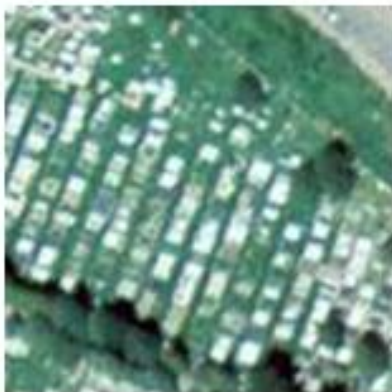


Схема даних



Приклади зображень з бази Aerial Image Classification



а)



б)



в)



г)



д)



е)

Приклад зашумлених зображення класу field з бази АІС



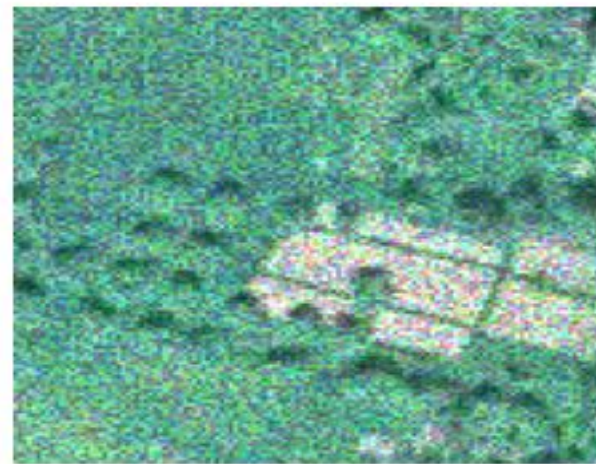
а)



б)



в)



г)

Дослідження розробленого методу класифікації зображень на тестовому наборі зображень

OOB estimate of error rate: 15.95%

Confusion matrix:

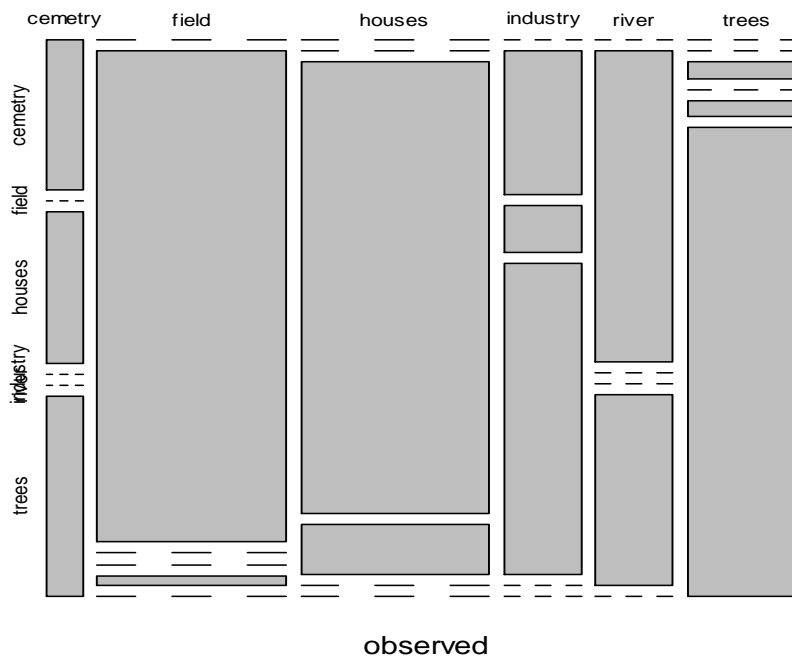
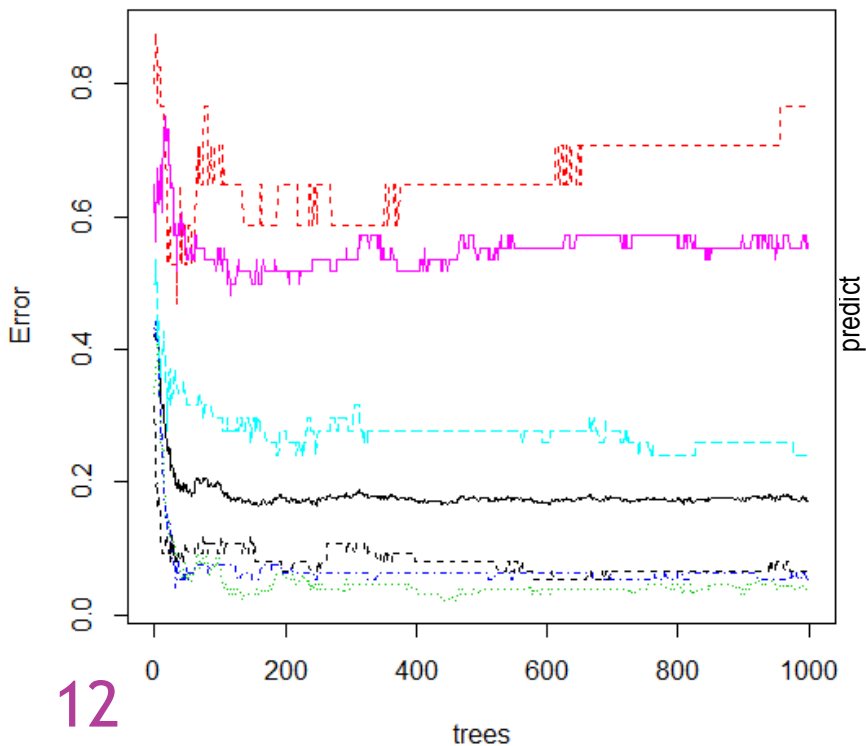
	cemetery	field	houses	industry	river	trees	class.error
cemetery	7	2	5	0	0	3	0.58823529
field	0	122	1	0	3	1	0.03937008
houses	0	0	86	3	0	3	0.06521739
industry	0	4	7	43	0	0	0.20370370
river	0	25	0	1	26	4	0.53571429
trees	0	0	1	0	4	69	0.06756757

r

predict

observed	cemetery	field	houses	industry	river	trees
cemetery	4	0	3	0	0	3
field	0	50	0	0	1	0
houses	0	0	46	5	0	0
industry	0	5	2	14	0	0
river	0	13	0	0	8	0
trees	0	0	1	0	1	29

t



Дослідження розробленого методу класифікації зашумлених зображень в порівнянні з оригінальними зображеннями на тестовому наборі АІС

OOB estimate of error rate: 9.48%

Confusion matrix:

	field	houses	industry	trees	class.error
field	122	2	2	2	0.04687500
houses	1	87	2	2	0.05434783
industry	8	7	38	1	0.29629630
trees	1	5	0	68	0.08108108

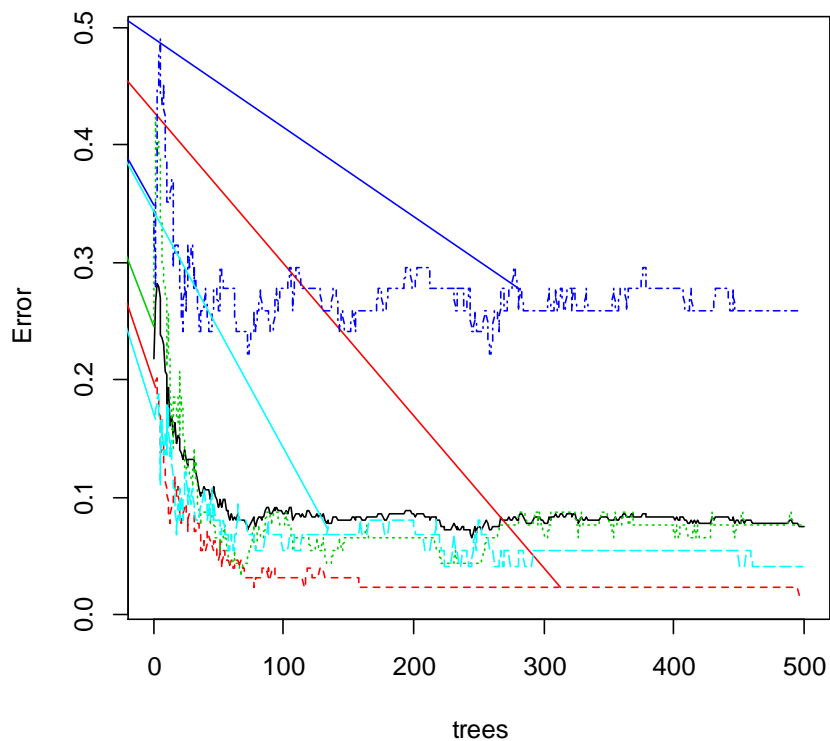
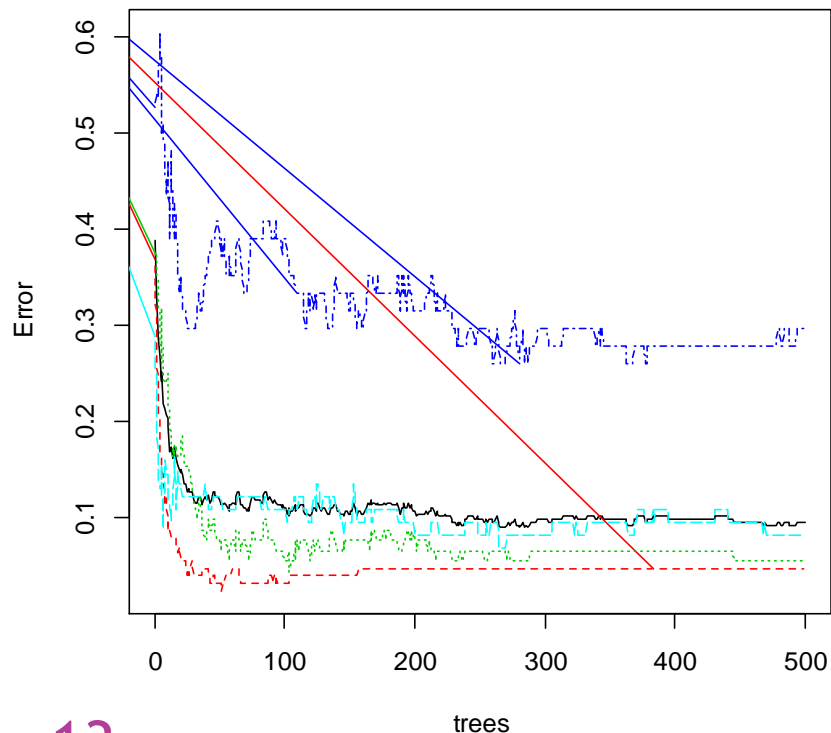
r

OOB estimate of error rate: 7.47%

Confusion matrix:

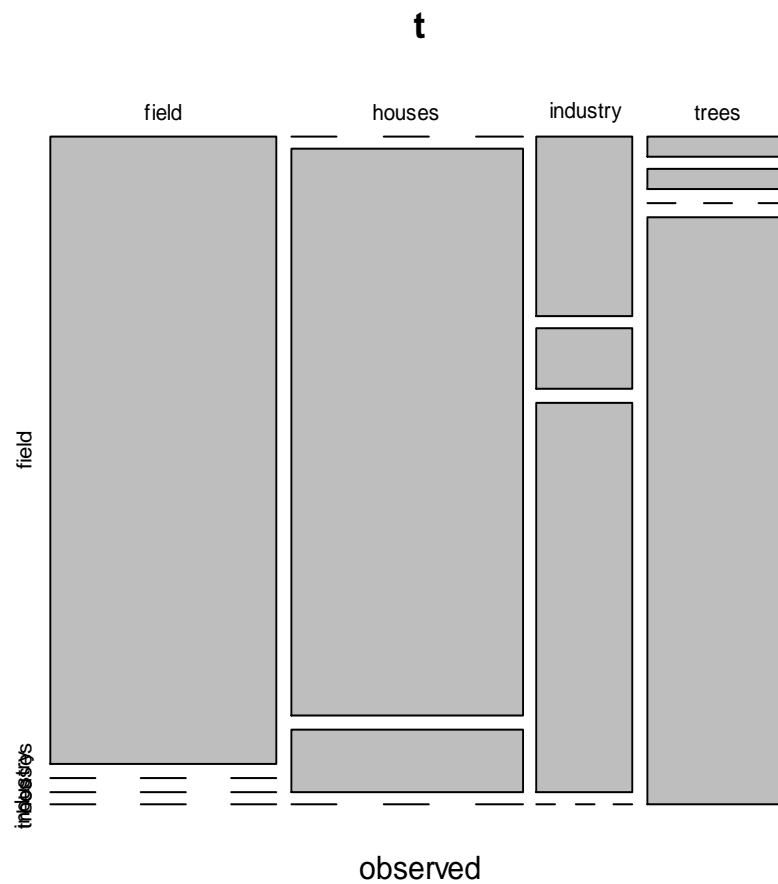
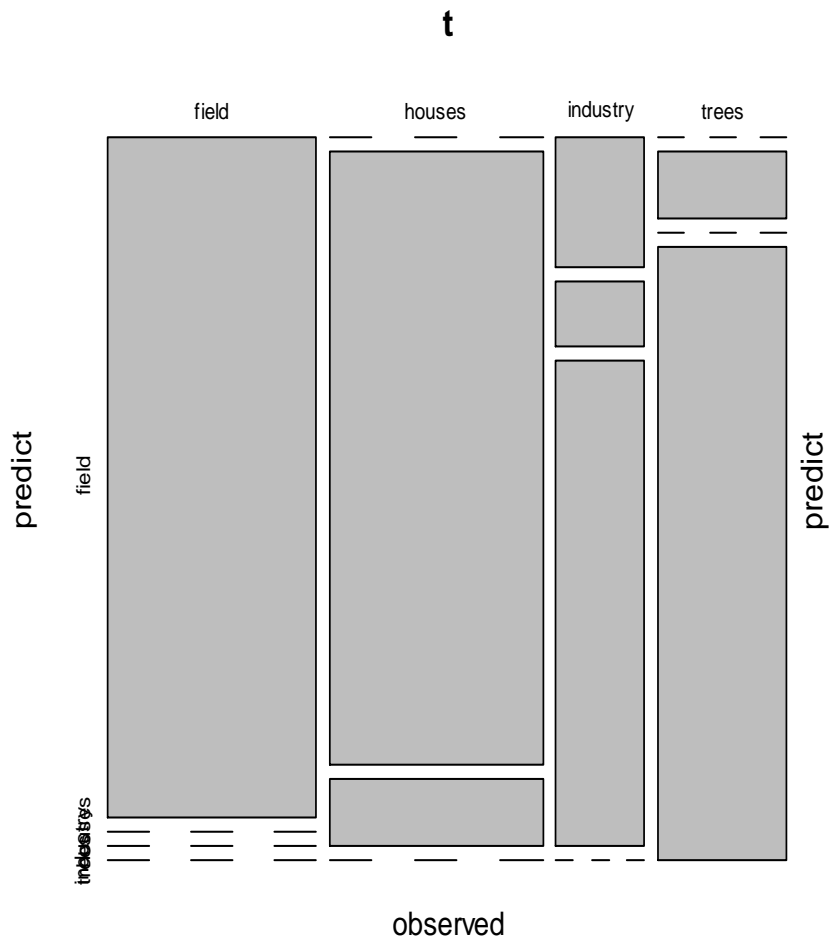
	field	houses	industry	trees	class.error
field	126	0	0	2	0.01562500
houses	1	85	3	3	0.07608696
industry	7	7	40	0	0.25925926
trees	1	2	0	71	0.04054054

r



	predict			
observed	field	houses	industry	trees
field	50	0	0	0
houses	0	46	5	0
industry	4	2	15	0
trees	0	3	0	28

	predict			
observed	field	houses	industry	trees
field	50	0	0	0
houses	0	46	5	0
industry	6	2	13	0
trees	1	1	0	29



ВИСНОВКИ: новий метод класифікації зображень, який на відміну від існуючих використовує випадкові ліси та ЛБШ у якості дескрипторів зображення, дозволив поліпшити класифікацію зображення для гаусівського білого шуму у діапазоні $[0,01; 0,03]$, для імпульсного шуму у діапазоні $[0,04; 0,06]$ та для мультиплікативного шуму - $[0,03; 0,05]$.

Було проведено експериментальне дослідження розробленого методу класифікації зображення на тестовому наборі зображень, з використанням різних параметрів та його порівняння з дослідженням класифікації зашумлених даних. Та після того як були розглянуті дослідження на 3 видах шумів: сіль та перець, спекл-шум та гаусівський шум, було встановлено, що найкраща помилка класифікації була отримана при накладанні шуму сіль та перець: для класу *field* і *houses* помилка класифікації зовсім не змінилася, і дорівнювала 0 та 0.098 відповідно; для *industry* на зашумлених зображеннях помилка дорівнювала 0.286 та на оригінальних – 0.333; відповідно для *trees* 0.097 та 0.064. З цих результатів можна побачити, що в порівнянні з результатом роботи випадкового лісу на оригінальних зображеннях, було отримано не значну різницю помилки класифікації, в порівнянні з іншими, при накладанні спекл-шум та гаусівського шуму.

Дякую за увагу!