

ДОДАТОК А

Графічний матеріал атестаційної роботи

**ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
РАДІОЕЛЕКТРОНІКИ**

Кафедра ЕОМ

Атестаційна робота на тему:

Моделі семантичної сегментації об'єктів міського середовища на цифрових аерофотознімках

Виконав: студент групи КСМзм-19-1 Мазур Ф.Л.

Керівник: дтн Олізаренко С.А.

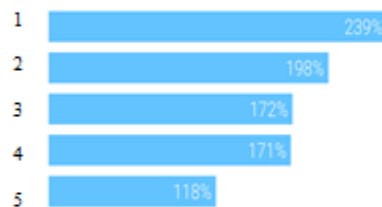
- *Об'єктом дослідження є процеси сегментації об'єктів міського середовища з використанням методології глибокого навчання.*
- *Метою атестаційної роботи є дослідження та модифікація моделей семантичної сегментації об'єктів міського середовища з використанням методології глибокого навчання.*
- *Методи дослідження: теорія штучних нейронних мереж, методи і моделі машинного навчання, методи математичної статистики.*

Задачі роботи:

- провести аналіз напрямків автоматизації процесів виявлення та класифікації простих об'єктів міського середовища на цифрових аерофотознімках;
- визначити підхід щодо сегментації об'єктів міського середовища на цифрових аерофотознімках;
- визначити підхід щодо формування навчальної вибірки зображень для навчання ЗНМ сегментації об'єктів міського середовища;
- провести моделювання процесів виявлення та класифікації об'єктів міського середовища з використанням ЗНМ для семантичної сегментації.

Актуальність

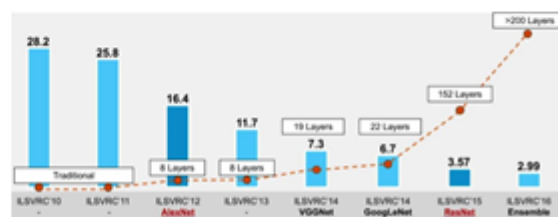
Стрімке зростання попиту на аерофотозйомку по галузям (1 – будівництво, 2 - юрисні юпаліни, 3 - сільське господарство, 4 - геодезія, 5 - нерухомість)



Застосування інтелектуальних програмних додатків для обробки цифрових зображень

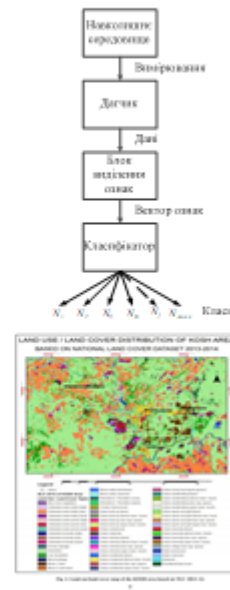
Сучасні методи та моделі CV на основі глибокого навчання

У комп'ютерному зорі, сегментація - це процес поділу цифрового зображення на декілька сегментів (безліч пікселів)



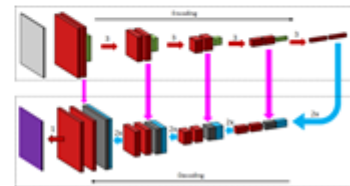
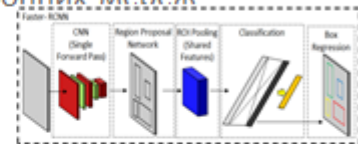
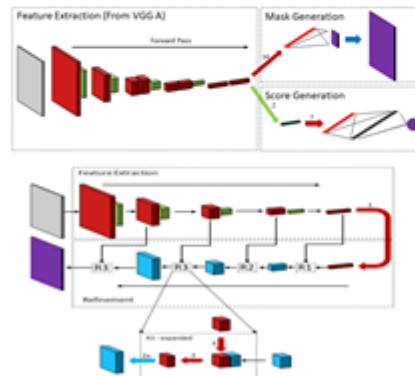
Аналіз змісту процесів виявлення об'єктів на цифрових знімках

Назва методу	Точність (кількість правильно класифікованих об'єктів вибірки)	Можливість реалізації в режимі часу, найближчого до реального	Можливість одночасного виявлення та класифікації об'єктів
BODATA	53-58%	Ні	Ні
K-means	61-66%	Ні	Ні
Метод контурного аналізу	50-56%	Так	Ні
Random forest	70-73%	Ні	Ні
Boosting	60-64%	Так	Ні
Нейронні мережі (перцептрон)	59-67%	Ні	Так
Вейклет	68-71%	Ні	Так
Support vector machine	75-81%	Ні	Так
Згорювані нейронні мережі	80-89%	Так	Так
Детектори HOG, Haar, LBP, SIFT	-	Так	Ні



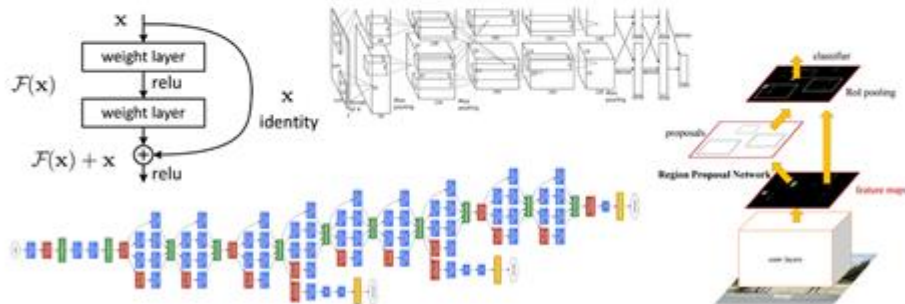
Аналіз основних моделей сегментації зображень на основі глибоких нейронних мереж

- 1) DeepMask ;
- 2) Sharpmask;
- 3) RCNN;
- 4) PSPNet;
- 5) U-Net.

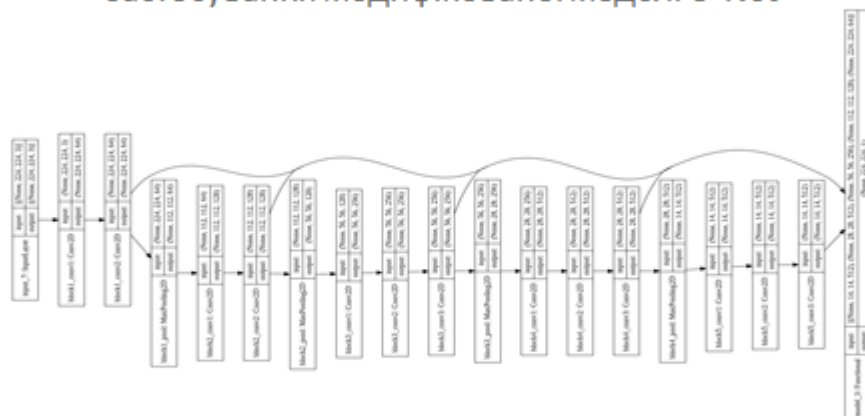


Аналіз основних попередньо навчених багатоцільових моделей CV на основі глибоких нейронних мереж

- 1) VGG16;
- 2) GoogLeNet ;
- 3) Google BERT;
- 4) ResNet;
- 5) AlexNet.



Дослідження особливостей побудови, навчання та застосування модифікованої моделі U-Net



В якості нейронної мережі використовується модифікована топологія мережі U-net зі вхідом 480x640 (кількість карт ознак однієї гілки: 32 → 64 → 128 → 256 → 256 + 256 (keras і tensorflow з використанням мови програмування python). У якості енкодера використовується VGG16.

Узагальнений алгоритм навчання модифікованої мережі (1 етап)

Для зменшення часу навчання використовується технологія «bottleneck features» згідно з якою на виході предобученної моделі знімаються ознаки (в даному випадку у енкодера в форматі VGG-16) і подаються на вхід навченою частини мережі (в даному випадку це декодер в форматі U-Net) і в подальшому навчається лише друга частина (т. є в даному випадку тільки декодер). Особливістю моделі U-Net є наявність проміжних зв'язків між енкодером і декодером. Для декодера має місце кілька входів (в даному випадку 5), що значно ускладнює використання технології «bottleneck features» при семантичній сегментації (на відміну від нейронного класифікатора у якого тільки один вхід).

В ході дослідження запропоновано підхід до вирішення даної проблеми, який дозволив скоротити одну епоху навчання на однакових даних з 1 години до 5 хвилин (при цьому використовувати всі переваги попередньо навченої моделі на базі даних ImageNet).



Узагальнений алгоритм навчання модифікованої мережі(2 етап)

Тонка настройка. На етапі розморожуються 1-2 шари енкодера, з'єднуються ваги енкодера (попередньо навченої моделі VGG-16 на бібліотеці зображень ImageNet) і отримані на першому етапі ваги декодера і виконується тонка настройка розморожених шарів енкодера і шарів декодера..

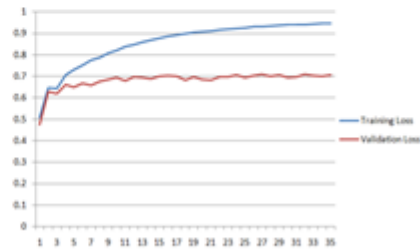


Оцінка ефективності результатів сегментації

Таблиця 1 – Результати оцінки

(сегментації) об'єктів

Тип нейронної мережі	Точність	Помилка
SegNet	82,7	81,6
FCN	72,7	75,0
U-Net	82,3	82,2
Розроблений в роботі підхід на основі модифікованого U-Net	84,8	84,1



ВИСНОВКИ

1. В результаті проведеного аналізу відомих методів для автоматизації виявлення так класифікації об'єктів на цифрових знімках. Встановлено, що методи, які реалізують піксельно-орієнтований підхід доцільно застосовувати через їх більшу точність.
2. Аналіз основних напрямків підходів з автоматизації процесів обробки даних виявлення так класифікації об'єктів на цифрових знімках показав необхідність застосування знання-орієнтованих методів для семантичної сегментації об'єктів міського середовища.
3. Досліджено можливості використання згорткової нейронної мережі для автоматизації процесів виявлення та класифікації об'єктів на цифрових аерофотознімках.
4. У модифікованій моделі U-Net була випробувана технологія "battlenack" в поєднанні з технологією «transfer learning». Валідаційні помилка при цьому зменшилася до 7-8%. Для того, щоб підвищити точність розпізнавання необхідно додатково збільшувати навчальну вибірку.

ДОДАТОК Б

Топологія модифікованої нейромережевої моделі U-Net для сегментації
об'єктів на цифрових знімках