

Виявлення об'єктів у зображеннях при використанні нейромережевого підходу

Виконав: студент гр. 1АКІТ-17м

Рамарчук Євгеній Олександрович

Науковий керівник: к.т.н., доцент кафедри АІТ

Гармаш Володимир Володимирович

Актуальність роботи

- ▶ Виявлення об'єктів у зображеннях є однією з важливих задач в системах технічного зору. Область застосування – це виробництва, в яких задіяні системи технічного зору, робототехнічні комплекси, веб-сервіси, які здійснюють аналіз та пошук зображень, мобільні застосування для ідентифікації та пошуку, системи відео спостереження. Виявлення об'єктів є однією з перших задач в системах розпізнавання об'єктів, в яких об'єкт спочатку потрібно локалізувати на зображенні а потім його ідентифікувати.

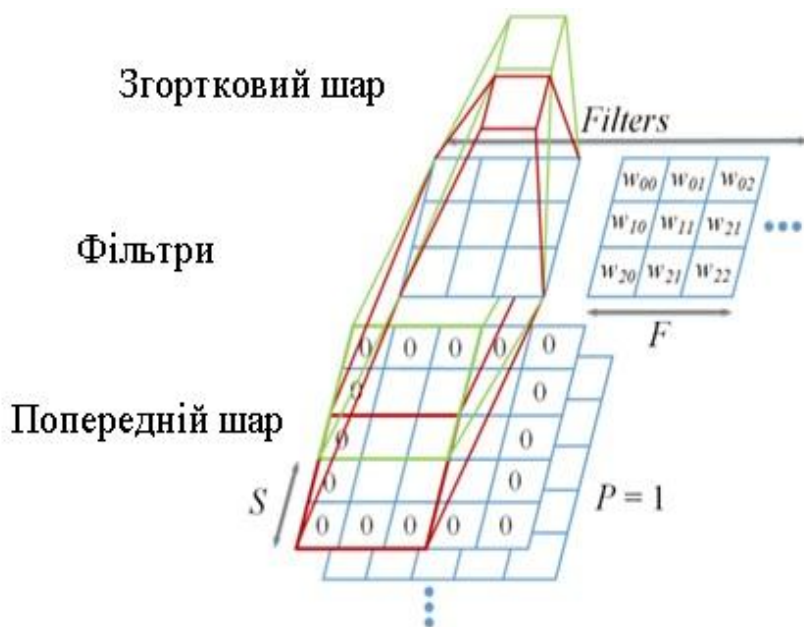
Наукова новизна, об'єкт та предмет дослідження

- ▶ **Наукова новизна одержаних результатів.** Удосконалено метод виявлення об'єктів у зображеннях, що використовує нейромережеву модель DetectNet, шляхом створення власної caffe моделі, що дозволило підвищити вірогідність виявлення.
- ▶ **Об'єкт дослідження** - процеси глибокого навчання нейронних мереж та виявлення об'єктів у зображеннях.
- ▶ **Предмет дослідження** - методи виявлення об'єктів у зображеннях, основані на нейромережевому підході.

Мета та задачі дослідження

- ▶ **Мета та задачі дослідження.** Метою роботи є підвищення якості виявлення об'єктів у зображеннях та пришвидшення навчання моделей при використанні методів глибокого навчання та суперкомп'ютера Jetson TX2
- ▶ Для досягнення поставленої мети потрібно вирішити такі задачі:
 - ▶ проаналізувати теоритичне підгрунтя розпізнавання об'єктів у зображеннях;
 - ▶ розглянути сучасні архітектури згорткових нейронних мереж;
 - ▶ визначити недоліки існуючих методів та знайти шляхи підвищення швидкодії та покращення вірогідності виявлення об'єктів у зображеннях;
 - ▶ розглянути програмні та апаратні засоби системи виявлення об'єктів у зображеннях;
 - ▶ вибрати тренувальний набір зображень, що найкраще підходить для виявлення об'єктів інтересу;
 - ▶ здійснити навчання власних нейромережових моделей та дослідити їх ефективність на тестових наборах зображень.

Згортковий шар



- Згортковий шар нейронної мережі є одним з основних і призначений для виділення ознак зображення і їх перетворення, які, в свою чергу, в подальшому на більш глибоких шарах, використовуються для отримання більш складних ознак і, в кінцевому підсумку, визначають клас об'єкта, що розпізнається. Основною характеристикою даного шару є так звані фільтри - багатовимірні (зазвичай двовірні або тривірні) матриці ваг зв'язку нейронів попереднього шару з нейронами Згорткового шару. Фільтрами вони називаються тому, що операція згортки, тобто отримання вихідного сигналу нейрона згорткового шару, дуже схожа на операцію фільтрації зображення.

Нейромережа DetectNet

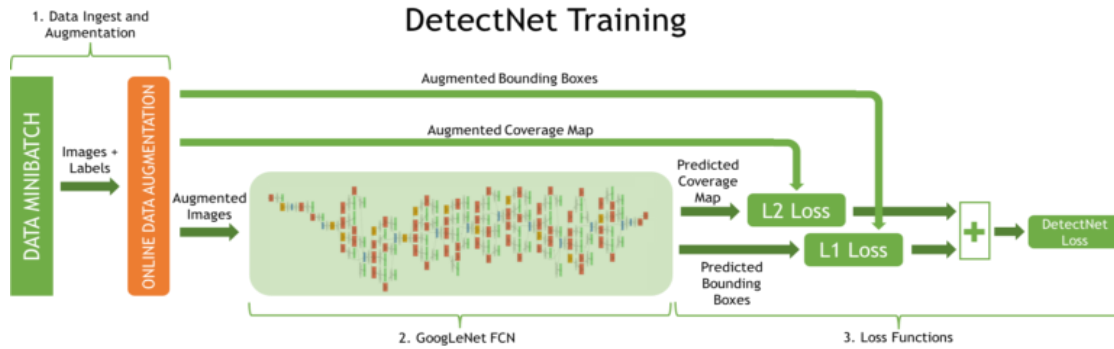


Рисунок 6.1 – Структура мережі DetectNet для навчання

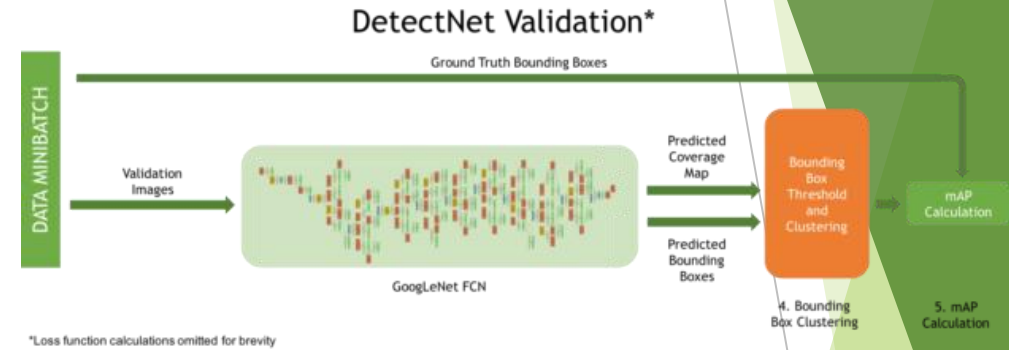
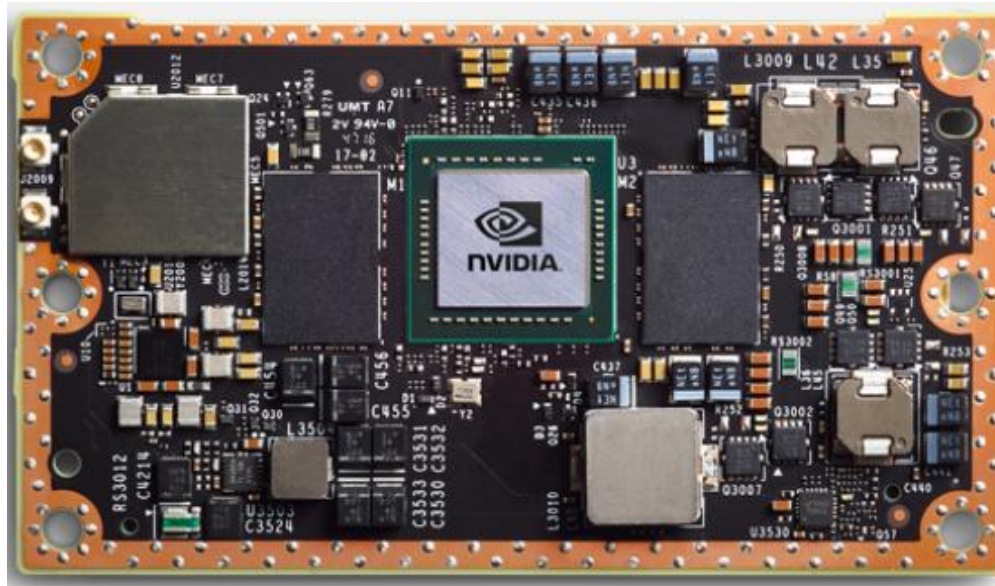


Рисунок 6.2 – Структура мережі DetectNet для перевірки ефективності моделі на тестових даних

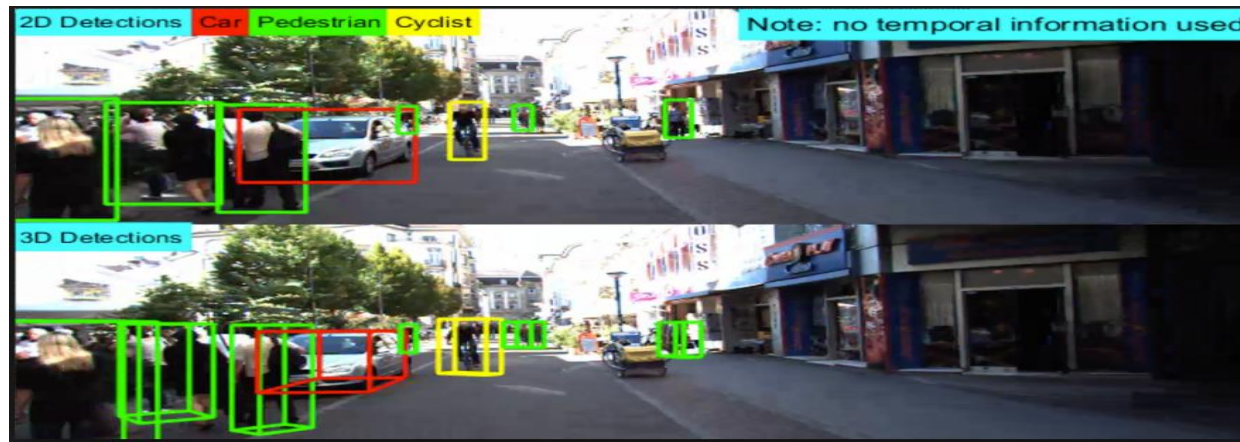
- DetectNet — це нейромережева архітектура, яка є розширенням популярної нейромережі GoogLeNet, що одночасно здійснює виявлення та класифікацію об'єктів.

Огляд особливостей Jetson TX2



- ▶ Jetson TX2 є одним з найшвидших, найефективніших вбудованих обчислювальних пристроїв штучного інтелекту. Цей 7,5-ватний суперкомп'ютер розміщується на одній платі. Він побудований на сімействі процесорів GPU NVIDIA Pascal і завантажується з 8 Гб оперативної пам'яті і 59,7 Гб/с пропускної здатності пам'яті. Він має різноманітні стандартні апаратні інтерфейси, які дозволяють легко інтегрувати його в широкий спектр продуктів і форм-факторів .

Багатопрфільна база зображень «Kitti»



- ▶ Kitti – багатопрфільна база зображень для систем автономного водіння. Використовується для одометрії, оптичного потоку, задач виявлення та відстеження 2D та 3D об'єктів.
 - ▶ Основні характеристики наборів для виявлення об'єктів:
 - ▶ 7481 тренувальних зображень та 7518 тестових зображень;
 - ▶ 80 тисяч об'єктів у зображеннях;
 - ▶ 3 категорії об'єктів (автомобілі, мотоцикли, пішоходи)
 - ▶ займає 12 Гігабайт

Результати досліджень

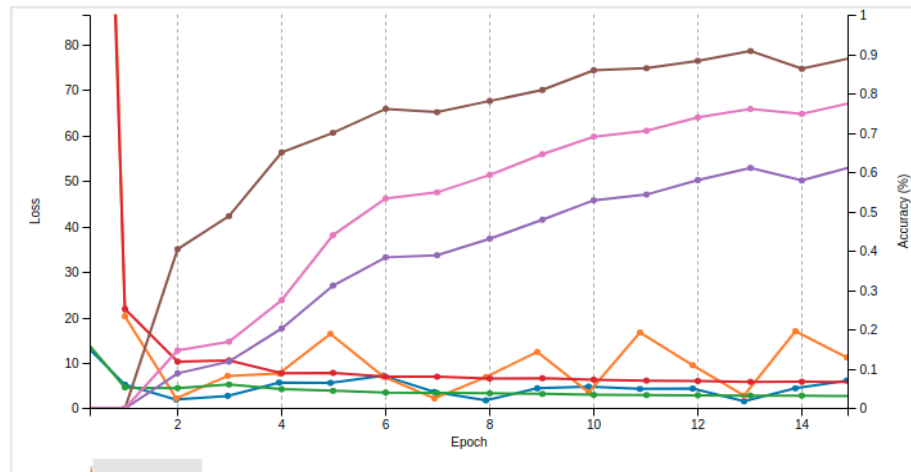


Рисунок 9.1 – Результати тестування моделі 2

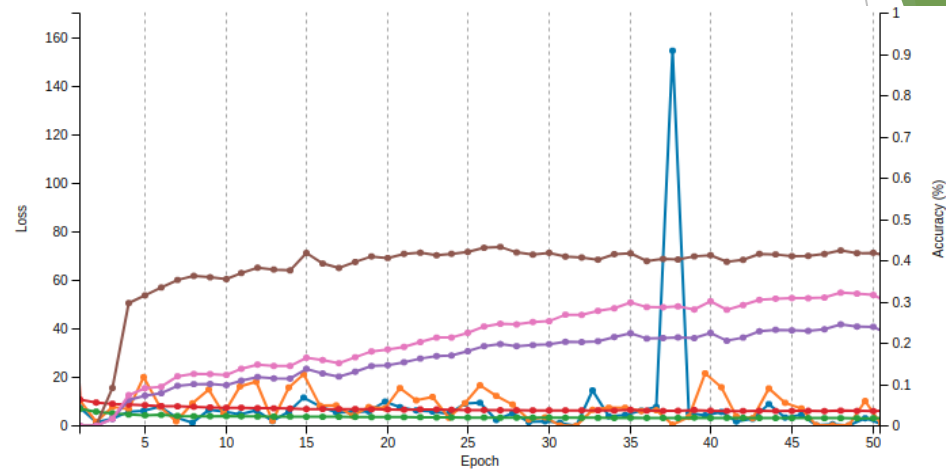


Рисунок 9.2 – Результати тестування моделі 5

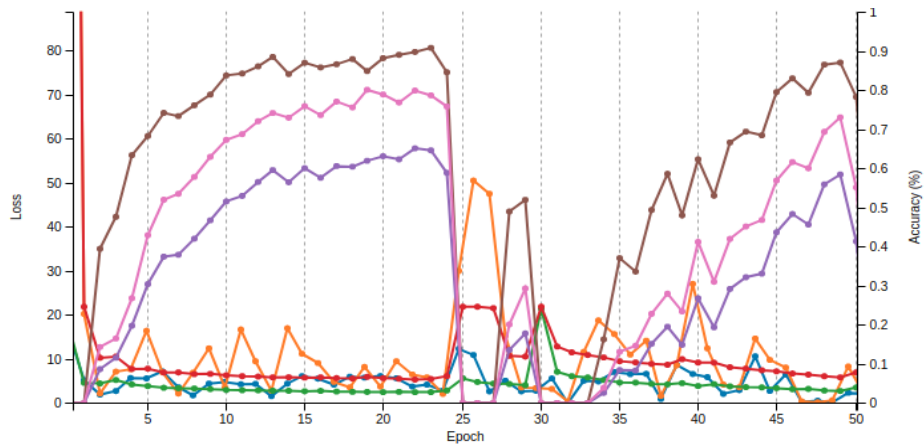


Рисунок 9.3 – Результати тестування моделі 6

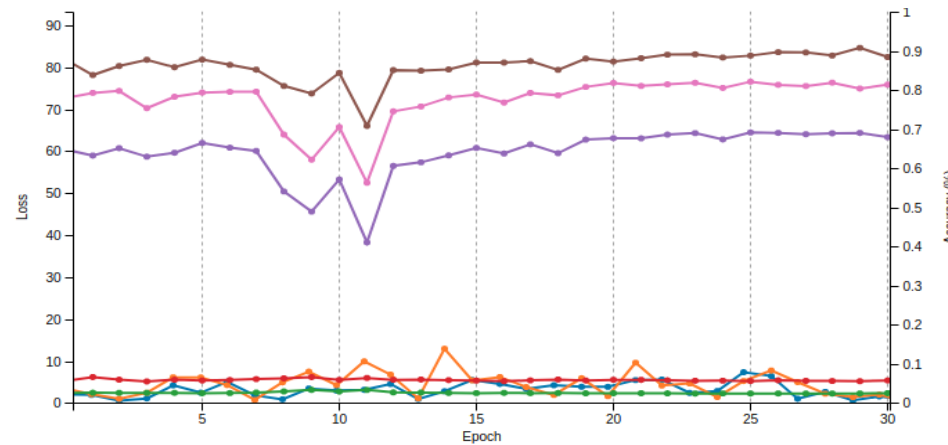
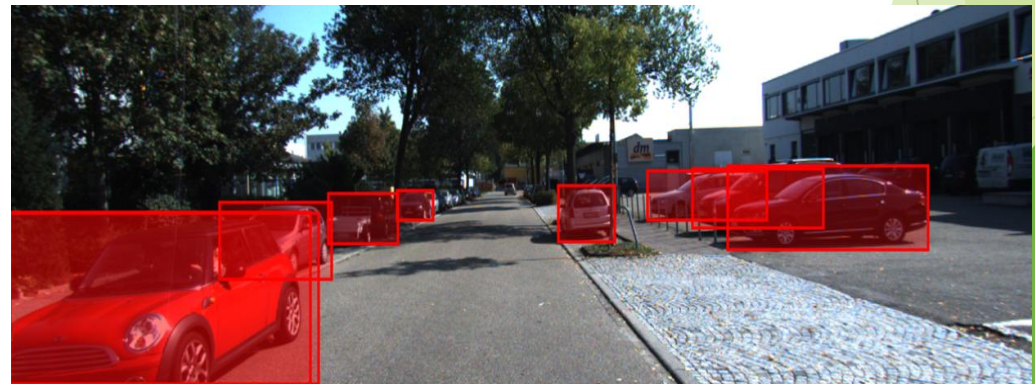
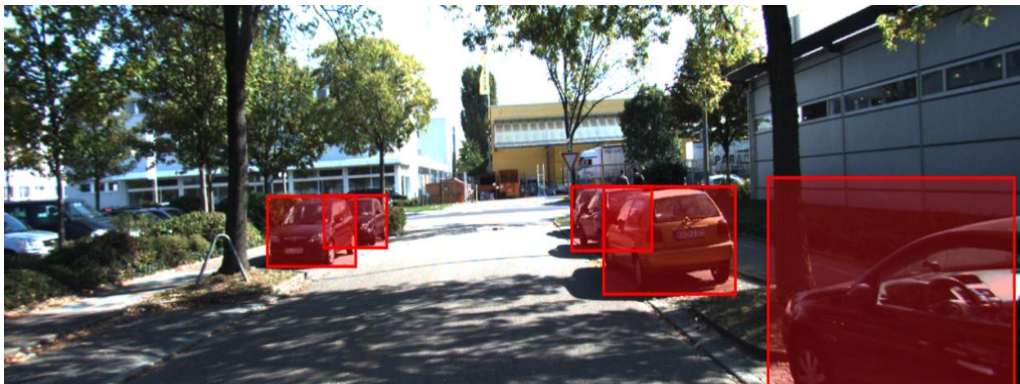
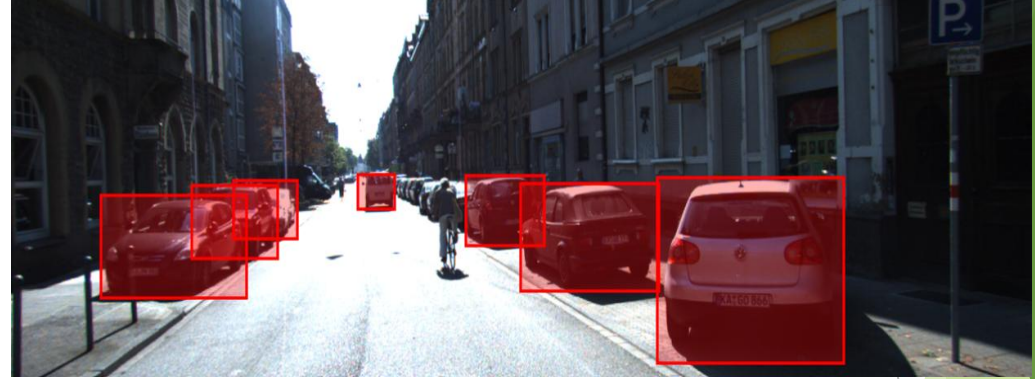
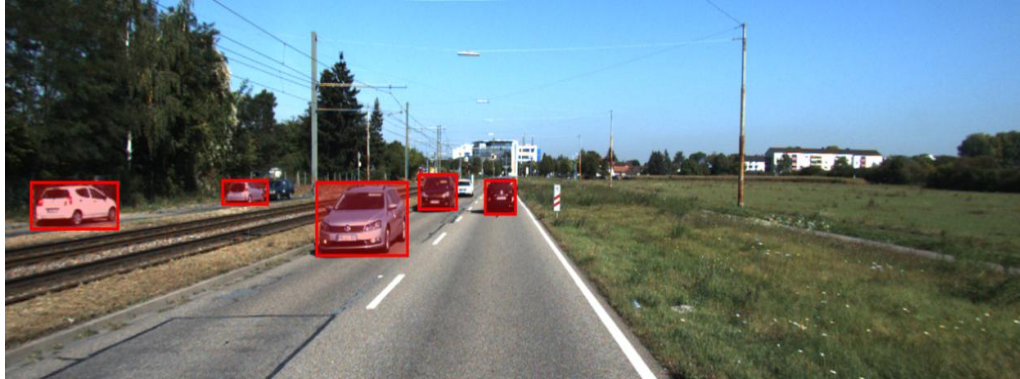


Рисунок 9.4 – Результати тестування моделі 8

Результати досліджень

| Назва моделі | Кількість епох / найкраща епоха | Функція зміни швидкості навчання | Швидкість навчання | Наперед тренована модель | mAP |
|--------------|---------------------------------|----------------------------------|--------------------|--------------------------|-------|
| 1 | 15/15 | Exponential decay | 0.0001 | GoogleNet | 34.8 |
| 2 | 15/15 | Exponential decay | 0.000075 | GoogleNet | 53.29 |
| 3 | 15/15 | Exponential decay | 0.00005 | GoogleNet | 45.67 |
| 4 | 15/30 | Exponential decay | 0.00025 | GoogleNet | 38.04 |
| 5 | 50/48 | Exponential decay | 0.00001 | GoogleNet | 41.81 |
| 6 | 50/23 | Exponential decay | 0.000075 | GoogleNet | 57.33 |
| 7 | 20/16 | Exponential decay | 0.000025 | 2 | 60.77 |
| 8 | 30/29 | Polynomial decay | 0.00005 | 7 | 64.33 |

Приклади виявлення автомобілів



Висновки

- ▶ В цілому в дослідженнях було сформовано 8 наборів параметрів і навчено 8 моделей. Частина цих моделей використовувало у якості наперед натренованої моделі – GoogleNet. Інші моделі навчалися на власних моделях. Найкраща модель показала значення $mAP=64.33$ на тестовому наборі KITTI.