

Виявлення об'єктів у зображеннях за допомогою JETSON TX2

Виконав: студент гр. 1АКІТ-17м

Тавдарханов Руслан Резванович

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

Маслій Роман Васильович

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.

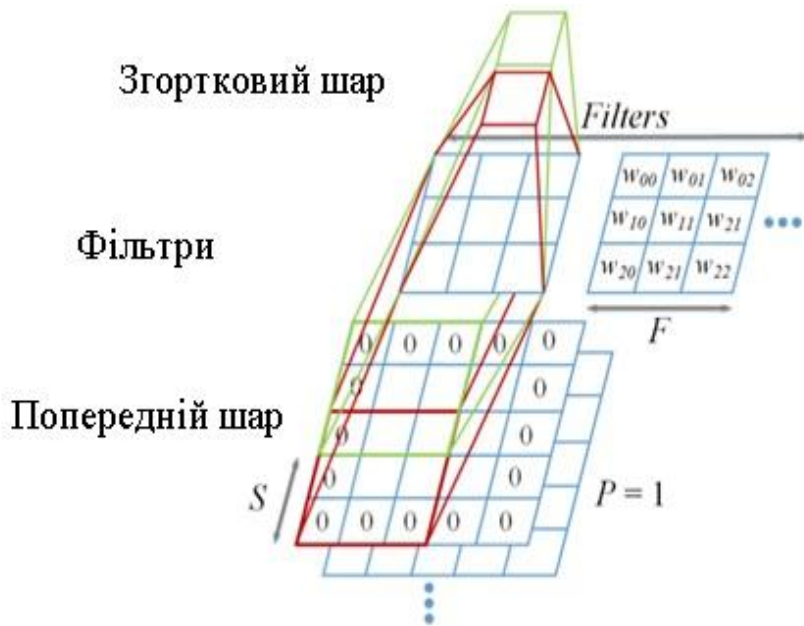
Наукова новизна, об'єкт та предмет дослідження

- v **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.
- v **Об'єкт дослідження** - процеси глибокого навчання нейронних мереж та виявлення об'єктів у зображеннях.
- v **Предмет дослідження** - методи виявлення об'єктів у зображеннях, основані на нейромережевому підході.

Мета та задачі дослідження

- υ **Мета та задачі дослідження.** Метою роботи є підвищення якості виявлення об'єктів у зображеннях при використанні методів глибокого навчання, засобу глибокого навчання NVIDIA DIGITS та суперкомп'ютера Jetson TX2.
 - υ Для досягнення поставленої мети потрібно вирішити такі задачі:
 - υ проаналізувати існуючі методи виявлення об'єктів у зображеннях;
 - υ дослідити роботу алгоритмів виявлення, основаних на нейромережевому підході;
 - υ визначити недоліки існуючих методів та знайти шляхи підвищення швидкодії та покращення вірогідності виявлення об'єктів у зображеннях;
 - υ реалізувати програмно модифікований алгоритм та порівняти результати.

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

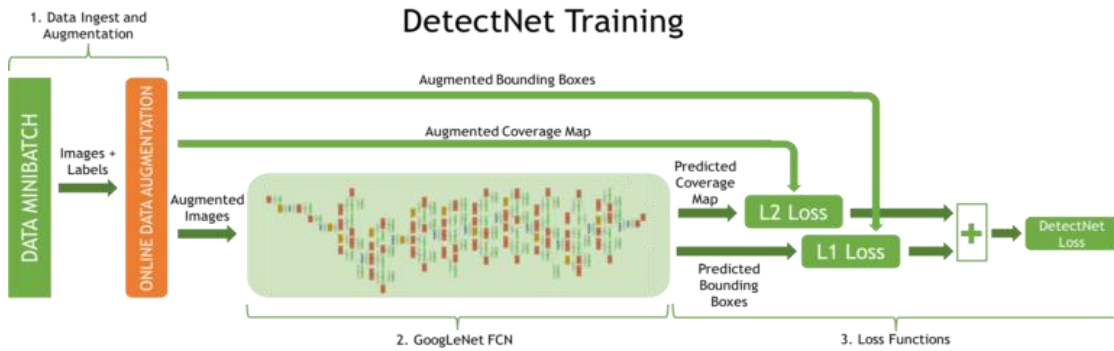


Рисунок 6.1 – Структура мережі DetectNet для навчання

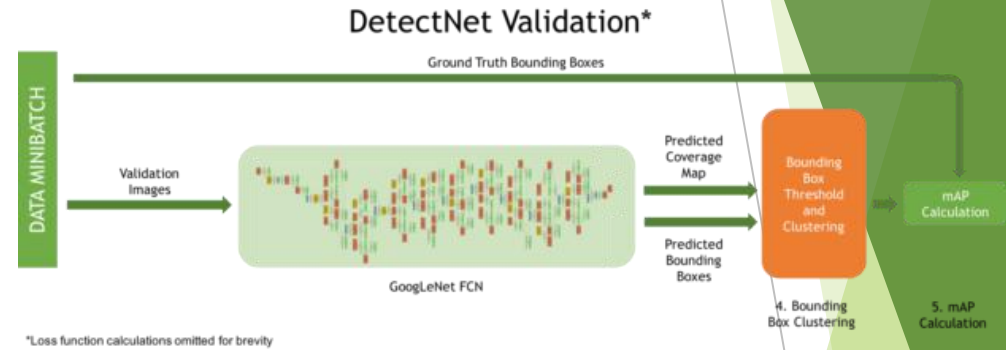
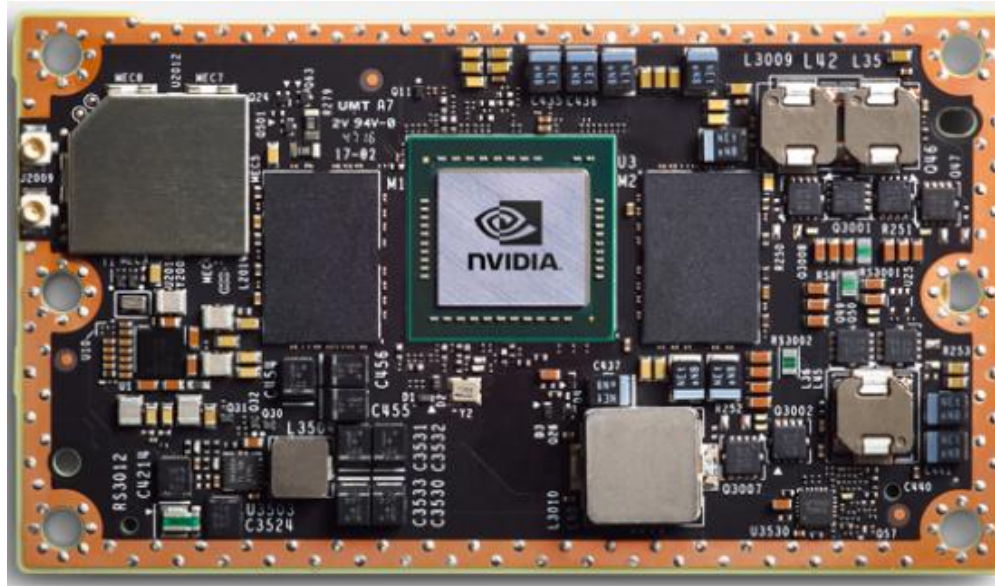


Рисунок 6.2 – Структура мережі DetectNet для перевірки ефективності моделі на тестових даних

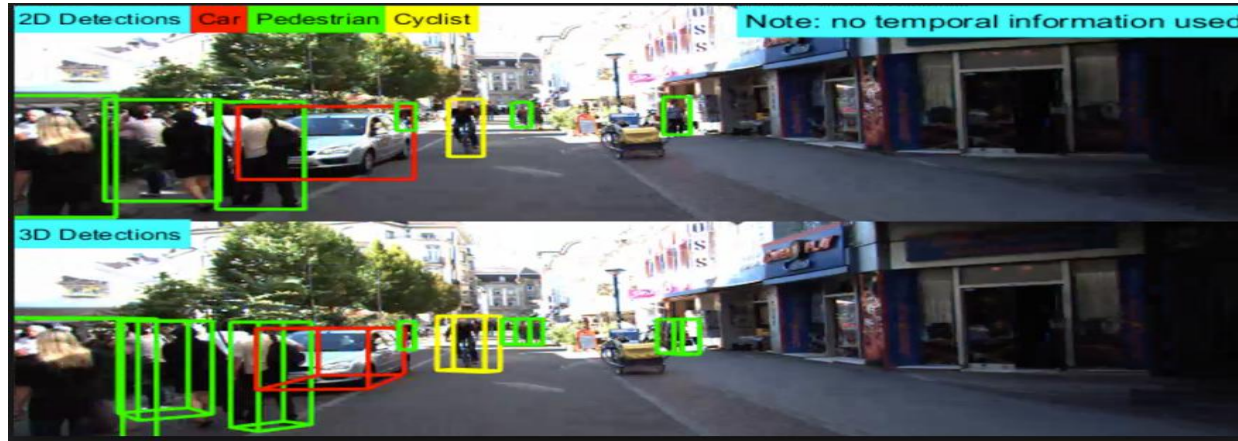
- DetectNet — це нейромережева архітектура, яка є розширенням популярної нейромережі GoogLeNet, що одночасно здійснює виявлення та класифікацію об'єктів.

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 .

Багатопрофільна база зображень «Kitti»



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

Результати досліджень

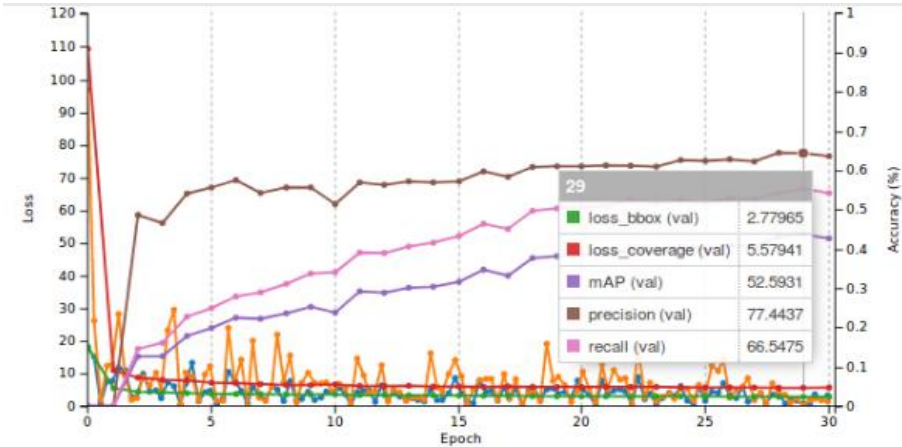


Рисунок 9.1 – Результати тестування моделі Kitti-4

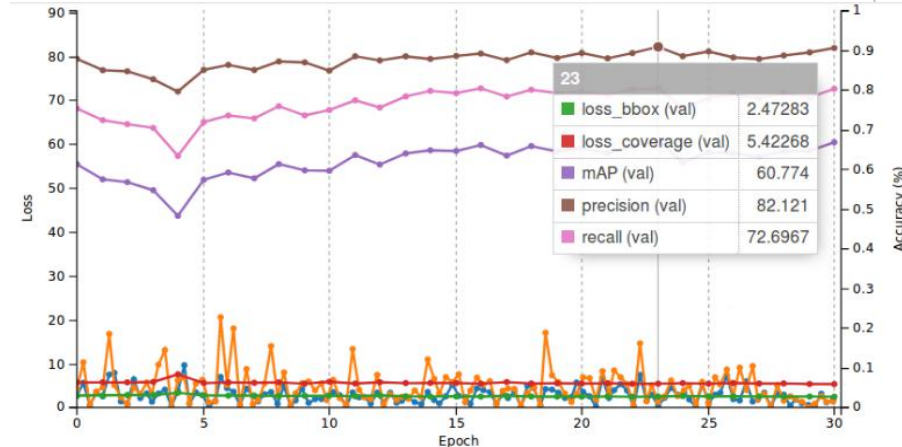


Рисунок 9.2 – Результати тестування моделі Kitti-7

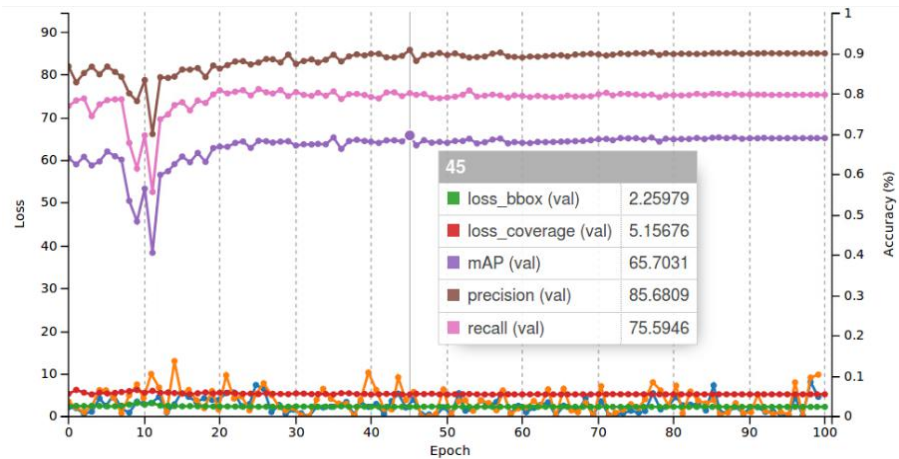


Рисунок 9.3 – Результати тестування моделі Kitti-9

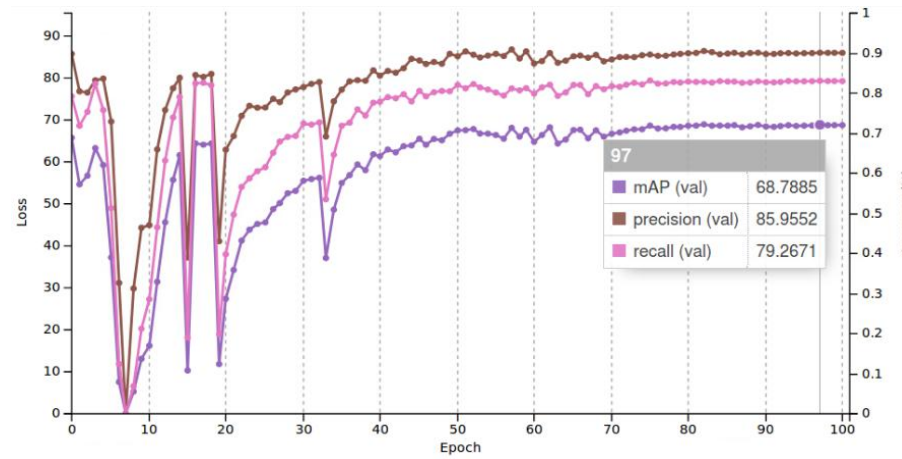
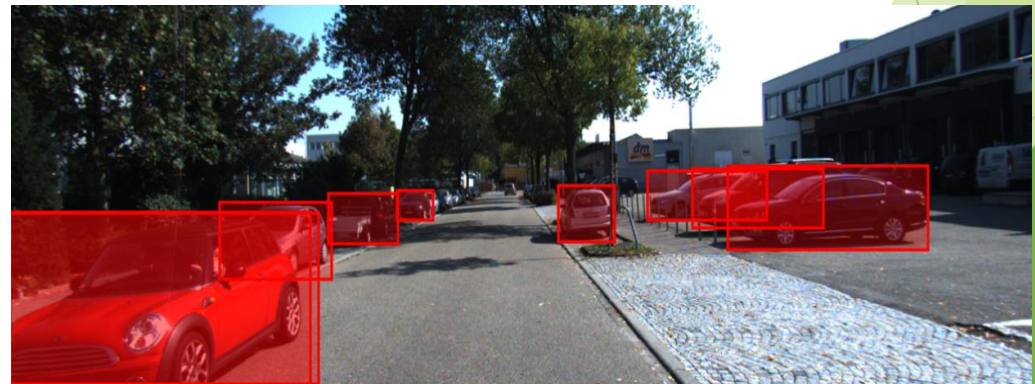
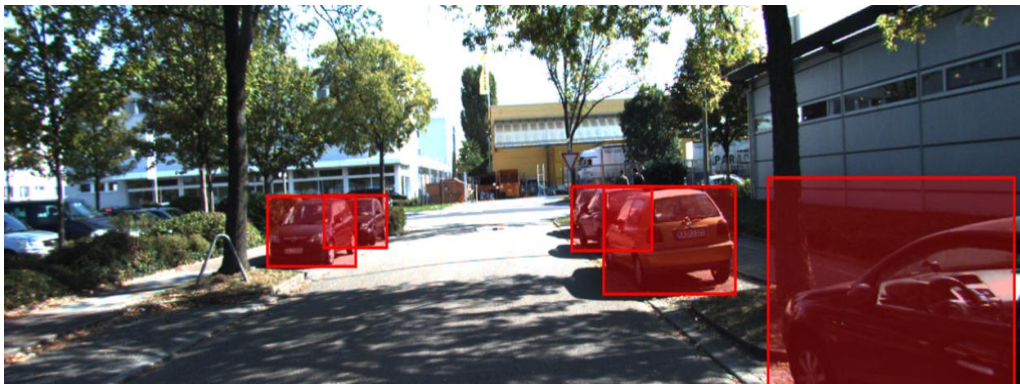
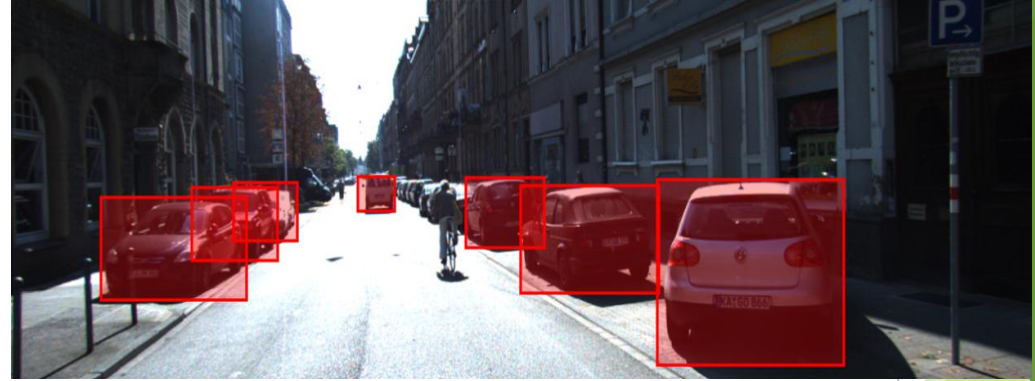
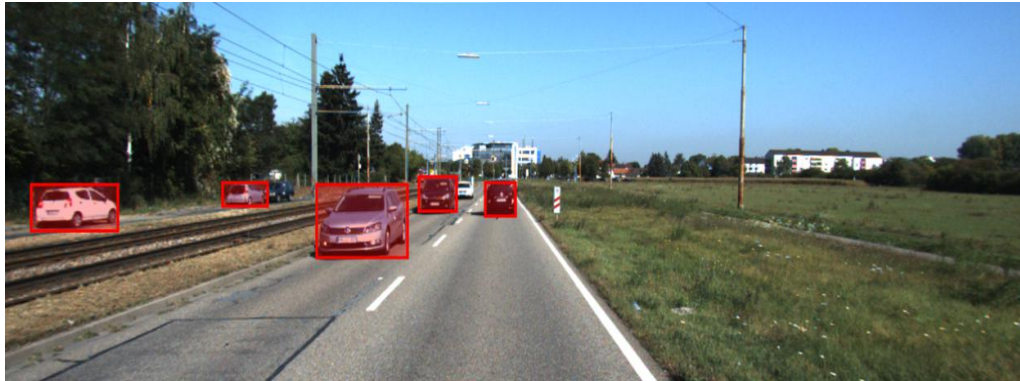


Рисунок 9.4 – Результати тестування моделі Kitti-10

Результати досліджень

Назва моделі	Кількість епох / найкраща епоха	Функція зміни швидкості навчання	Швидкість навчання	Наперед тренована модель	mAP
Kitti-1	30/16	Exponential decay	0.0001	GoogleNet	37.79
Kitti-2	30/22	Exponential decay	0.000075	GoogleNet	57.80
Kitti-3	30/30	Exponential decay	0.00005	GoogleNet	55.41
Kitti-4	30/30	Exponential decay	0.00025	GoogleNet	52.59
Kitti-5	100/94	Exponential decay	0.00001	GoogleNet	50.23
Kitti-6	100/22	Exponential decay	0.000075	GoogleNet	57.8
Kitti-7	30/23	Exponential decay	0.000025	Kitti-3	60.77
Kitti-8	100/97	Exponential decay	0.00001	Kitti-7	63.22
Kitti-9	100/45	Polynomial decay	0.00005	Kitti-7	65.7
Kitti-10	100/97	Polynomial decay	0.000025	Kitti-9	68.78

Приклади виявлення автомобілів



Висновки

- В цілому в дослідженнях було сформовано 10 наборів параметрів і навчено 10 моделей. Частина цих моделей використовувало у якості наперед натренованої моделі – GoogleNet. Інші моделі навчалися на власних моделях. Найкраща модель показала значення mAP=68.78 на тестовому наборі KITTI.