

МОДЕЛІ ТА ПРОГРАМНІ ЗАСОБИ ОЦІНЮВАННЯ ПРОСТОРОВОЇ ГЛИБИНИ СЦЕНИ НА ОСНОВІ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ

Роботу виконав

Приходченко Антон Сергійович
11-А клас, Ліцей №67 Джерело

Під керівництвом

Приходченка Сергія Дмитровича
Кандидата технічних наук,
Доцента НТУ «Дніпровська політехніка»



Мета роботи та об'єкт досліджень

Об'єктом досліджень є гібридний підхід до аналізу просторової глибини зображення із застосуванням розпізнавання об'єктів і нейронної мережі що будує відносну карту глибини.

Мета роботи: Метою роботи є розробка та дослідження моделей і програмних засобів оцінювання просторової глибини сцени за монокулярними зображеннями на основі поєднання методів детекції об'єктів і глибокого навчання.

Монокулярний зір

це коли сприйняття об'єктів відбувається за допомогою одного об'єктива.
Більшість камер працює за таким принципом.

Предмет досліджень та практична цінність

Предметом досліджень виступають моделі, методи та програмні засоби оцінювання просторової глибини сцени на основі нейромережевих алгоритмів і детекції об'єктів.

Практичне значення полягає у створенні програмного прототипу системи оцінювання глибини сцени з веб-орієнтованим інтерфейсом, який може бути використаний у навчальних цілях, експериментальних дослідженнях та прикладних задачах комп'ютерного зору. Результати роботи можуть бути використані для подальшого розвитку систем монокулярної оцінки глибини та інтеграції в реальні програмні застосунки.

Передумови дослідження

Оцінювання просторової глибини сцени є однією з ключових задач комп'ютерного зору, що має важливе значення для багатьох прикладних галузей, зокрема робототехніки, систем відеоспостереження, обробки зображень та інтелектуальних застосунків. Традиційні підходи до визначення відстані до об'єктів ґрунтуються на використанні стереозору або спеціалізованих сенсорів, що ускладнює апаратну реалізацію та підвищує вартість систем.

Сутність досліджуваної проблеми полягає в тому, що монокулярне зображення не містить прямої інформації про реальні відстані до об'єктів, а сучасні нейронні методи зазвичай забезпечують лише відносну оцінку глибини. Тому доцільним є поєднання методів глибокого навчання з детекцією об'єктів, що дозволяє відновлювати масштаб сцени на основі геометричних характеристик відомих об'єктів.

Глибина зображення

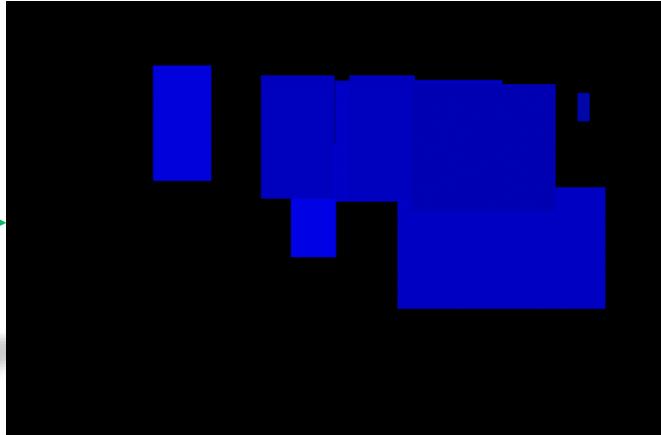
Канал пікселів зображення, що виражає їх відстань від точки зору.
Можна уявити це як туман, який рівномірно огортає сцену.
Чим далі від вас, тим більш густим він буде

Ідея запропонованої системи

Оригінальне зображення



(обробка за знайденими об'єктами)



(Відносна карти глибини)



Точний результат



Корегування

Як працює модель 1

Трикутник між фокусною відстанню і висотою у пікселях є подібним до трикутника між реальною відстанню і реальною висотою

Фокусна відстань камери і висота об'єкта на фотографії нам відома, і ми вже знаємо висоту багатьох об'єктів завдяки стандартизації.

$$d = \frac{of}{b}$$

Трикутники є подібними

Відстань до об'єкта

Фокусна відстань

Висота у пікселях

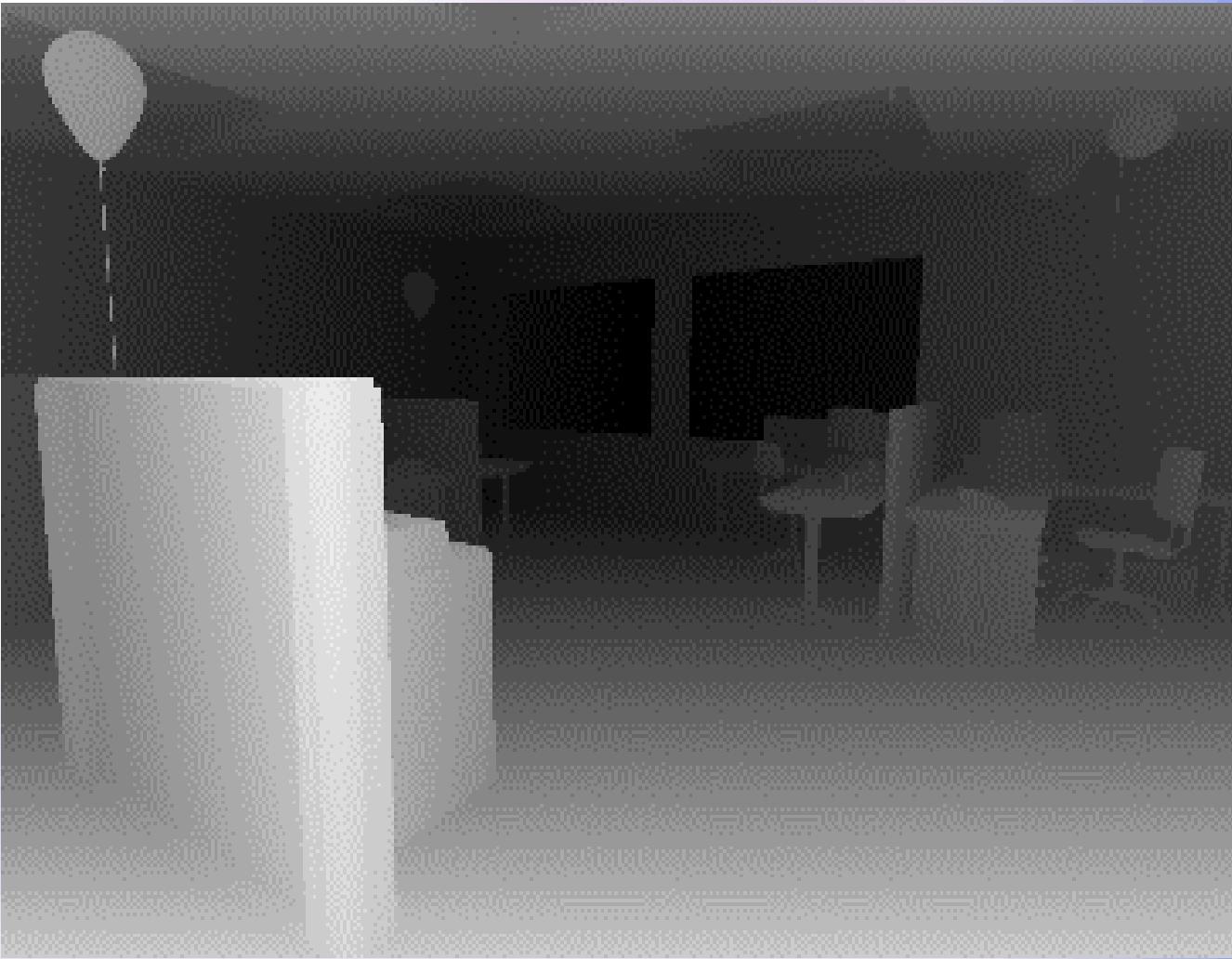
Прямий кут

Прямий кут

```
_, box.y1, _, box.y2 = b.xyxy[0] # визначаємо рамки  
box.height = abs(box.y1 - box.y2) # визначаємо висоту  
if box.height == 0 # запобігаємо ділення на 0  
    box.distance = 0  
else # Наше рівняння  
    box.distance = HeightByFocal / box.height
```

Як працює модель 2

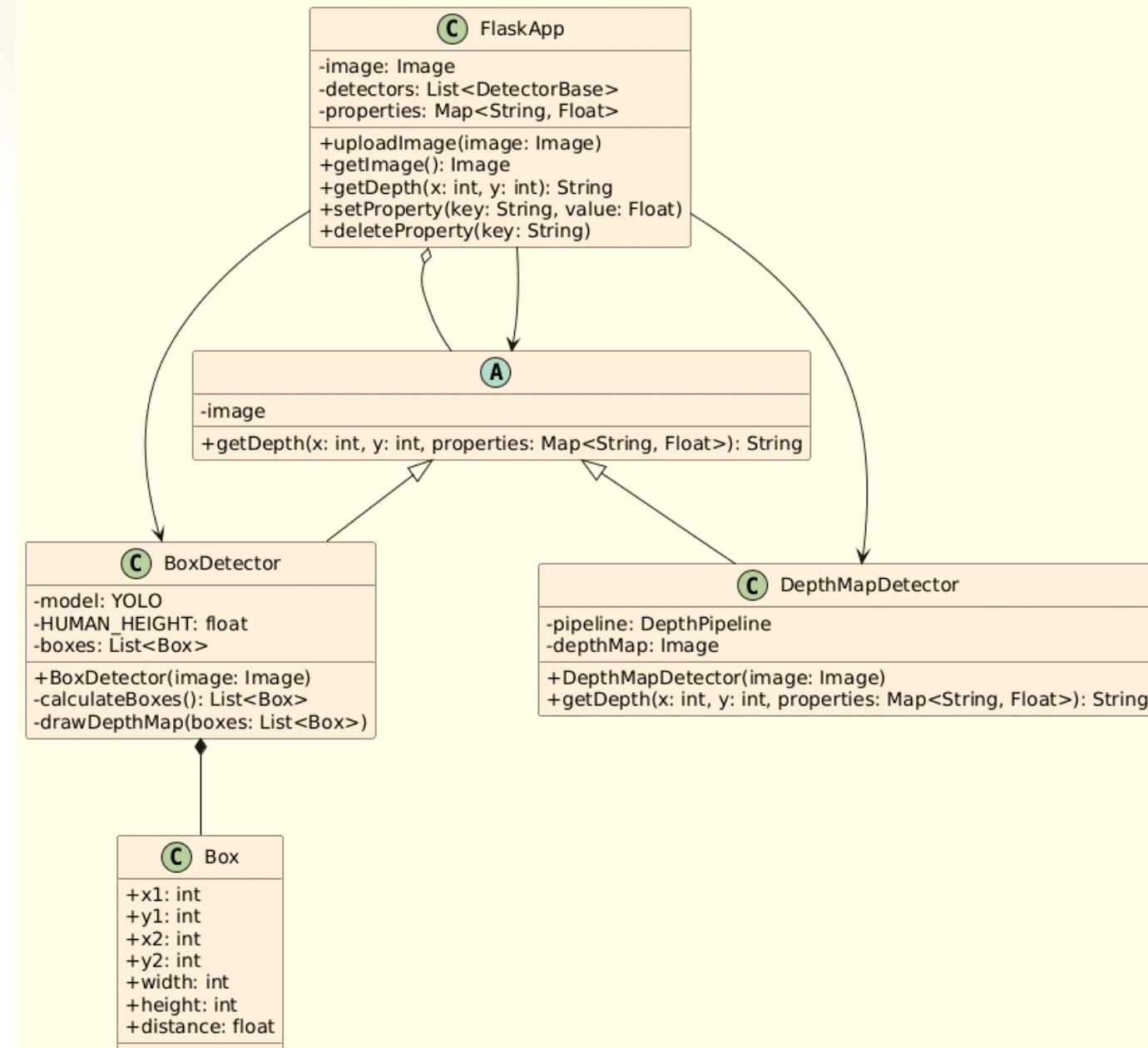
На наступному кроці, ми можемо застосувати модель Depth-Anything-V2-Small-hf для розрахунку відносної карти глибини. Вона дає чорно-біле зображення по принципу де більше, там яскравіше. Через обмеження 8 біт на піксель вона може давати лише обмежений діапазон значень, і тому ця карта не дає реальну відстань, а лише коефіцієнти порівняння.



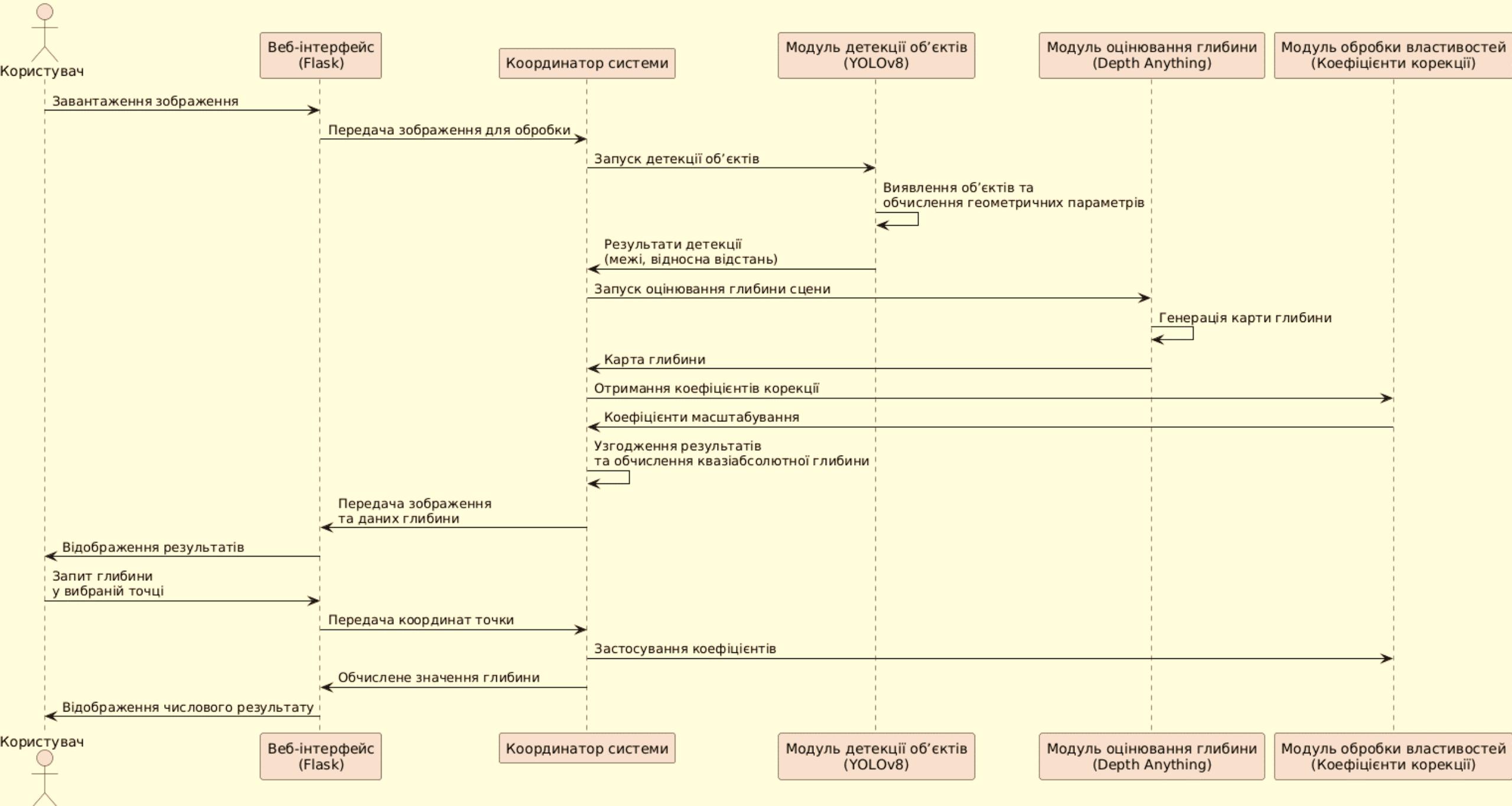
На основі запропонованого алгоритму, задля взаємодії з користувачем був побудован веб-клієнт на фреймворку Flask мови Python.

Користувач може завантажити зображення, замовляти дані для конкретної точки, замовити цілу карту глибини і змінювати коефіцієнти

Class Diagram of the Depth Estimation System



Послідовність взаємодії компонентів системи оцінювання глибини сцени

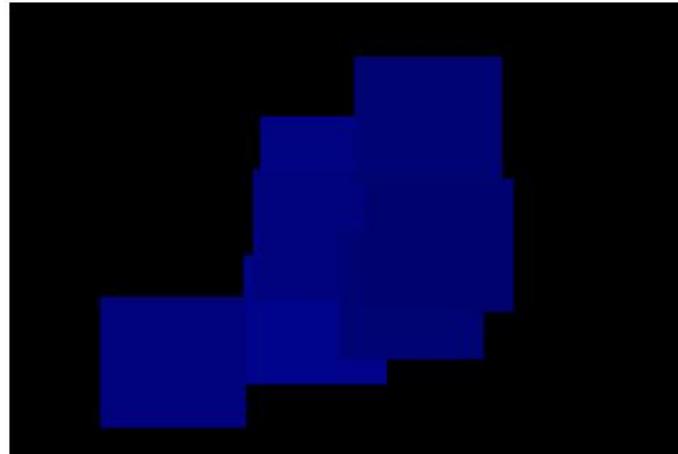


Опис інтерфейсу користувача

За необхідності передбачена можливість відстеження кроків алгоритму з їхньою візуалізацією, що допомагаючи при аналізі роботи, аналізі зображення та відповідних необхідних коригуваннях в процесі налаштування



a



б



в

а – оригінал, б – детекція об'єктів, в – відносна відстань

Порівняння методів

Метод	Додаткова апаратура	Точність	Калібрування	Глибина
Стереозір	Бажана	Висока	Потрібне	Абсолютна
LiDAR	Потрібна	Дуже висока	Нема	Абсолютна
Карта глибини	Нема	Середня	Нема	Відносна
Детекція Об'єктів	Нема	Висока, але не покриває все зображення	Потрібне	Абсолютна
Запропонований в роботі метод	Нема	Висока	Потрібне	Квазі-абсолютна

Проведене порівняння з традиційними методами показало, що запропонований підхід є гнучким, не потребує спеціалізованої апаратури та забезпечує достатню точність для багатьох прикладних задач.

Висновки

У роботі запропоновано гібридний метод визначення відстані, який базується на використанні відносної карти глибини та геометричних співвідношень для об'єктів із відомими або приблизними розмірами.

Реалізований програмний прототип підтверджив можливість практичного застосування такого підходу без використання додаткових сенсорів, що зменшує вартість і складність системи. Розроблений веб-орієнтований інтерфейс дозволяє користувачу завантажувати зображення, отримувати карту глибини та аналізувати відстань до окремих об'єктів.

Проведене порівняння з традиційними методами показало, що запропонований підхід є гнучким, не потребує спеціалізованої апаратури та забезпечує достатню точність для багатьох прикладних задач.

Роботу виконав

*Приходченко Антон Сергійович
11-А клас, Ліцей №67 Джерело*

Швидкі відповіді

Хто зробив роботу? – Приходченко Антон Сергійович

Чому така тема – Я побачив що аналоги дуже погані і легко щось зробити

Скільки часу я працював – 2 місяці, включаючи написання роботи

В чому наукова новизна? – Гібридний підхід, що робить результати точнішими

Як подивитися код? – На Гітхабі <https://github.com/AntonXCM/DepthAnalizingAlgorythms>

Як подивитися текст роботи? – На сайті <https://antonxcm.neocities.org/Pages/MAN/FullText.pdf>

Ціна ШІ? – ШІ запускається приватно на вашому сервері

Чому Пітон? – Тому що PYTorch, transformers, huggingface. Я би дуже хотів зробити на Node JS, але не зараз

Чому веб? – Тому що можна зробити багато користувачів

Що якщо об'єкт буде повернутий? – Є можливість розширити, застосувавши синус кута поворота

Що якщо об'єкт перекритий? – Навчання моделі треба провести таким чином, щоб такі об'єкти виключались

Хто тестував програму? – Я, мій науковий керівник і Колосов Дмитро Леонідович

Можна для дронів? – Да, можна

Можна для OSINT? – Да, можна

Можна для зд реконструкції? – Для цього воно і зроблено

