

ТРЕНУВАННЯ МОДЕЛІ ДЕТЕКЦІЇ ОБ'ЄКТІВ НА ЧАСТКОВО РОЗМІЧЕНИХ ЗОБРАЖЕННЯХ/OBJECT DETECTOR TRAINING WITH PARTIALLY LABELED IMAGES

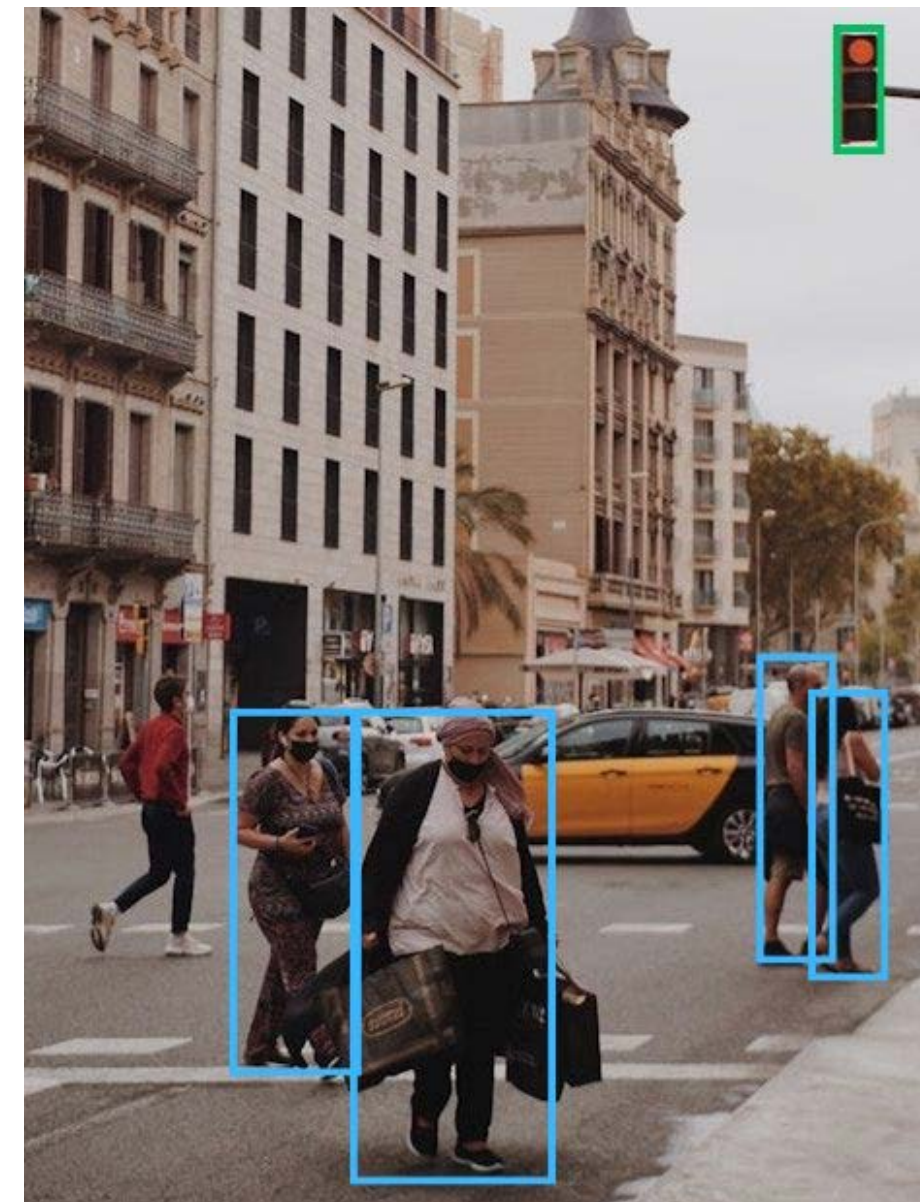
Керівник: кандидат фіз.-мат. наук, доцент Швай
Н. О.

Виконав: студент ПМ-4 Шакіров А. Т.

ПРОБЛЕМА ЧАСТКОВОЇ РОЗМІТКИ

На частково розмічених зображеннях присутні різні нерозмічені об'єкти, які модель може побачити та хибно віднести до фону.

Після декількох епох навчання це призведе до призначення цим об'єктам хибних міток або неможливості їх виявлення.



МЕТА ТА ЗАВДАННЯ

Мета - розробити ефективний метод тренування моделі детекції об'єктів на частково розмічених зображеннях.

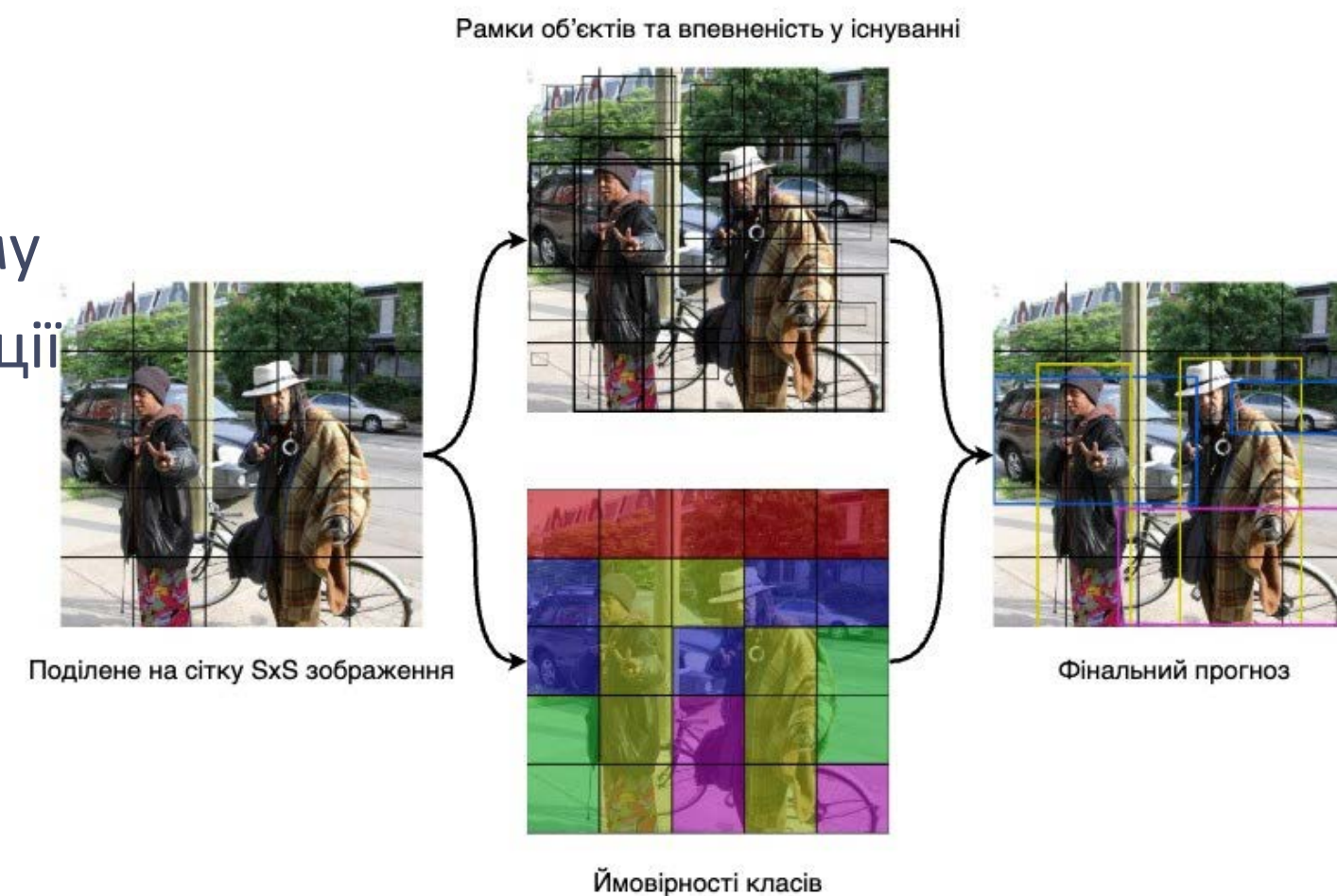
Завдання:

- оглянути методи використання часткової розмітки для навчання моделі.
- розробити модифіковану процедуру навчання із запропонованими підходами.
- експериментально оцінити їх ефективність.

МОДЕЛЬ ДЕТЕКЦІЇ YOLO

Для модифікації та експериментування обрано модель YOLO11. Вона здатна працювати в реальному часі та є однією з найпопулярніших моделей детекції об'єктів.

Функція втрат YOLO складається з 3 компонентів: втрати локалізації, об'єктності та класифікації.



ФУНКЦІЯ ВТРАТ YOLO

$$L_{loc} = \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{obj} \cdot L_{CIoU}(\hat{B}_{ij}, B_{ij})$$

$$L_{obj} = \sum_{i=0}^{S^2} \sum_{j=0}^B -\mathbb{1}_{ij}^{obj} (\hat{C}_i \log C_i + (1 - \hat{C}_i) \log(1 - C_i)) \\ - \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{noobj} (\hat{C}_i \log C_i + (1 - \hat{C}_i) \log(1 - C_i))$$

$$L_{cls} = - \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{obj} \sum_{c \in K} (\hat{p}_i(c) \log(p_i(c)) + (1 - \hat{p}_i(c)) \log(1 - p_i(c)))$$

НАБІР ДАНИХ

Експериментування проводилось із датасетом PASCAL VOC 2012. Він був поділений на 2 частини:

1. **Набір для тренування базової моделі** - використано 20% анотацій та зображень із обраним класом.
2. **Набір часткової розмітки для дотренування** - застосовано всі зображення із об'єктами обраного класу та їх розміткою. Всі інші анотації видалено.

СПОСОБИ ВИРІШЕННЯ ПРОБЛЕМИ

Для вирішення цієї проблеми існує 2 шляхи:

1. Модифікація процесу навчання.
2. Модифікація часткової розмітки.

Перший передбачає зміну самого тренування або функції втрат моделі.

Другий - зміну набору даних, який використовує модель для навчання.

МОДИФІКАЦІЯ ПРОЦЕСУ НАВЧАННЯ

Відкидання втрати для прогнозів моделі, які зроблені для неіснуючих об'єктів.

$$L_{total} = \sum_{i=0}^B M_i \cdot (L_{i_loc} + L_{i_obj} + L_{i_cls})$$

МОДИФІКАЦІЯ ПРОЦЕСУ НАВЧАННЯ

Застосування техніки дистиляції знань для імітації існуючої моделі завдяки новим компонентам у функції втрат.

$$L_{KD_cls} = D_{KL}(\hat{P}_{teacher} || \hat{P}_{student})$$

$$L_{KD_loc} = (\hat{B}_{teacher} - \hat{B}_{student})^2$$

МОДИФІКАЦІЯ ЧАСТКОВОЇ РОЗМІТКИ

Для модифікації набору даних запропоновано 3 підходи:

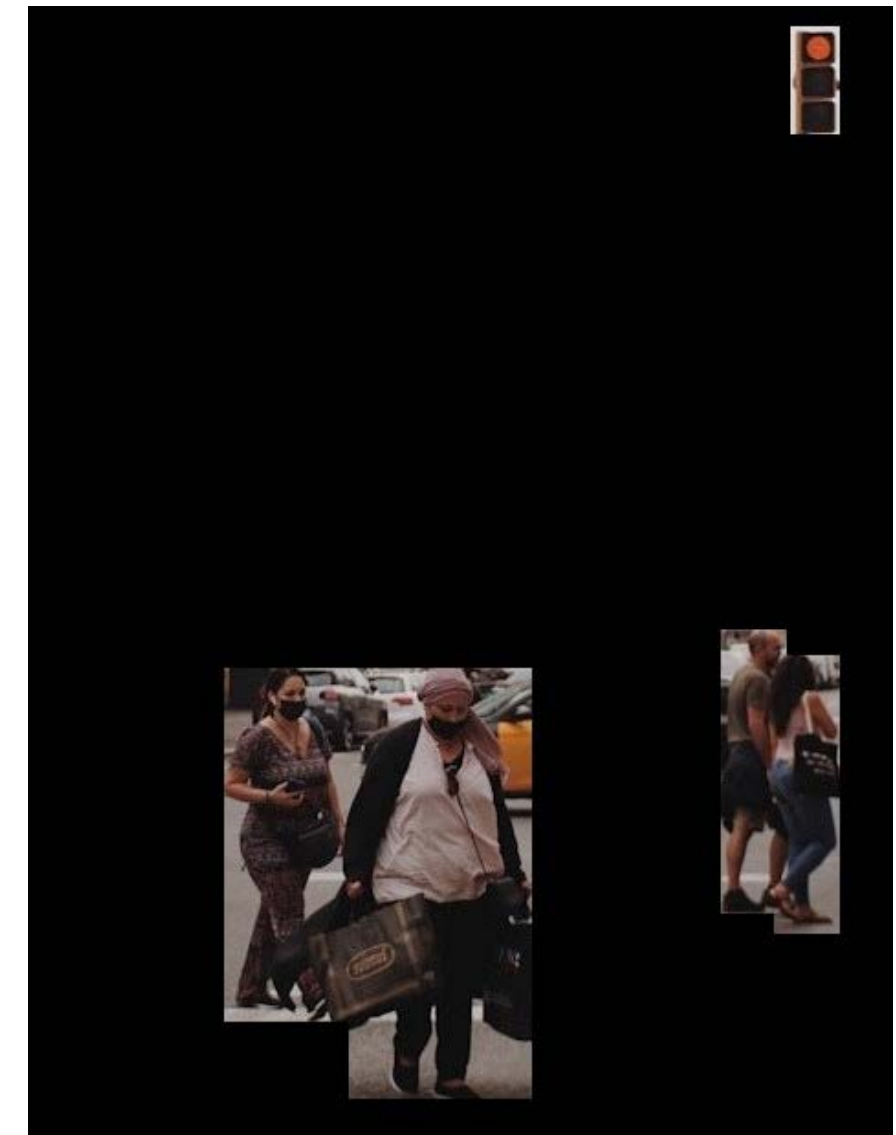
1. Маскування фону зображення

$$I_n^{masked} = I_n(x, y) \cdot \mathbb{1}_{A_n}(x, y)$$

2. Комбінація повної та часткової розмітки

3. Псевдорозмітка

$$T_{pseudo}(I_n) = (f_{\theta}(I_n) \setminus \{(c, b, s) \mid c = c_0\}) \cup T(I_n)$$



ЕКСПЕРИМЕНТУВАННЯ

Підхід	Всі класи		Обраний клас	
	mAP50	mAP50–95	mAP50	mAP50–95
Базова модель	0.785	0.599	0.291	0.136
Маскування класів	0.275	0.166	0.430	0.268
Дистиляція знань	0.644	0.453	0.298	0.177
Повна/часткова розмітка	0.720	0.524	0.439	0.271
Псевдорозмітка	0.696	0.507	0.438	0.266
Маскування фону	0.156	0.090	0.028	0.012
Дистиляція та м. класів	0.737	0.553	0.300	0.139
Дистиляція та п/ч розмітка	0.766	0.579	0.294	0.145
М. класів та п/ч розмітка	0.667	0.487	0.426	0.229
М. фону та п/ч розмітка	0.367	0.199	0.114	0.044

ВИСНОВКИ

Оцінивши проведені експерименти, можна зробити наступні висновки:

- Тренування моделей детекції об'єктів на частково розмічених даних можливе.
- Модифікація функції втрат дозволяє частково зберегти загальну точність моделі.
- Отримання найкращих результатів вимагає найбільш наближену оптимізацію набору даних до повної розмітки.

The slide features a central white rectangular area with a thin dark border. The corners of the slide are decorated with dark blue geometric patterns. The top-left corner has a circle with a dotted outline and a solid center. The top-right corner has a circle with a dotted outline and a solid center. The bottom-left corner has a circle with a dotted outline and a solid center. The bottom-right corner has a circle with a dotted outline and a solid center. The text "ДЯКУЮ ЗА УВАГУ!" is centered in the white area in a bold, dark blue, sans-serif font.

ДЯКУЮ ЗА УВАГУ!