

Національний Технічний Університет України
“КПІ ім. Ігоря Сікорського”
Інститут прикладного системного аналізу

САМОСТІЙНА РОБОТА № 4
З курсу «Розпізнавання образів»

Виконали:
Студенти групи КА-77
Котів Сергій
Буркацький Мината
Морозов Роман

Київ 2020

Завдання

Маючи детектор та класифікатор ми не викидаємо нічого а натомість тренуємо класифікатор, щоб той розпізнавав якийсь маркер (обкладинку книги, листок А4 з характерним малюнком, постер КіШа, відкритку з Мальдів, вибір за вами). Після розпізнавання, додаток має достатньо точно замінити картинку на маркері іншим зображенням також на ваш розсуд, але зображення має набувати тої ж форми і подоби що зображення в маркері (наприклад, якщо візитка Патріка Бейтмана лежить на столі, то в результаті роботи додатку, зображення, що має замінити візитку має також у дуже схожий спосіб лежати на столі під таким же кутом і таких же розмірів).

Класифікатор на розгляді: Detectron

Предмет: Пачка сірників

Виконання роботи

Створення датасету

За допомогою програми labelme, ми створили маски розмітки для сотні фото сірникової коробки. Розмітка виглядає наступним чином:



Навчання класифікатора

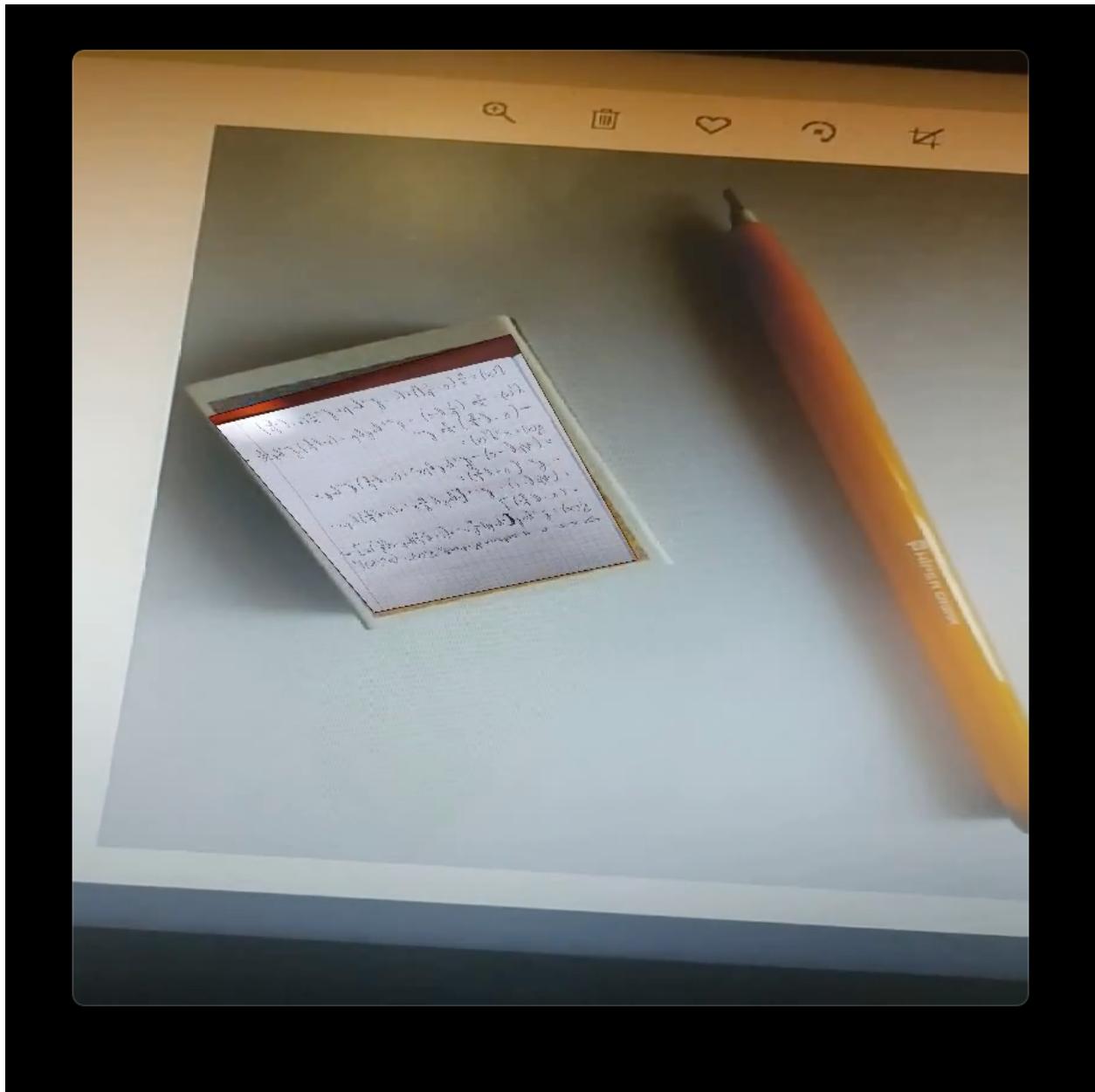
Використовуючи нейронну мережу із архітектурою R50-FPN Mask R-CNN, ми провели 300 навчальних епох. Результати розпізнавання виглядають наступним чином:



Заміна зображення

Ми читуємо відео по кадрам та замінюємо зображення, що класифікатор визначив як коробку сірників.

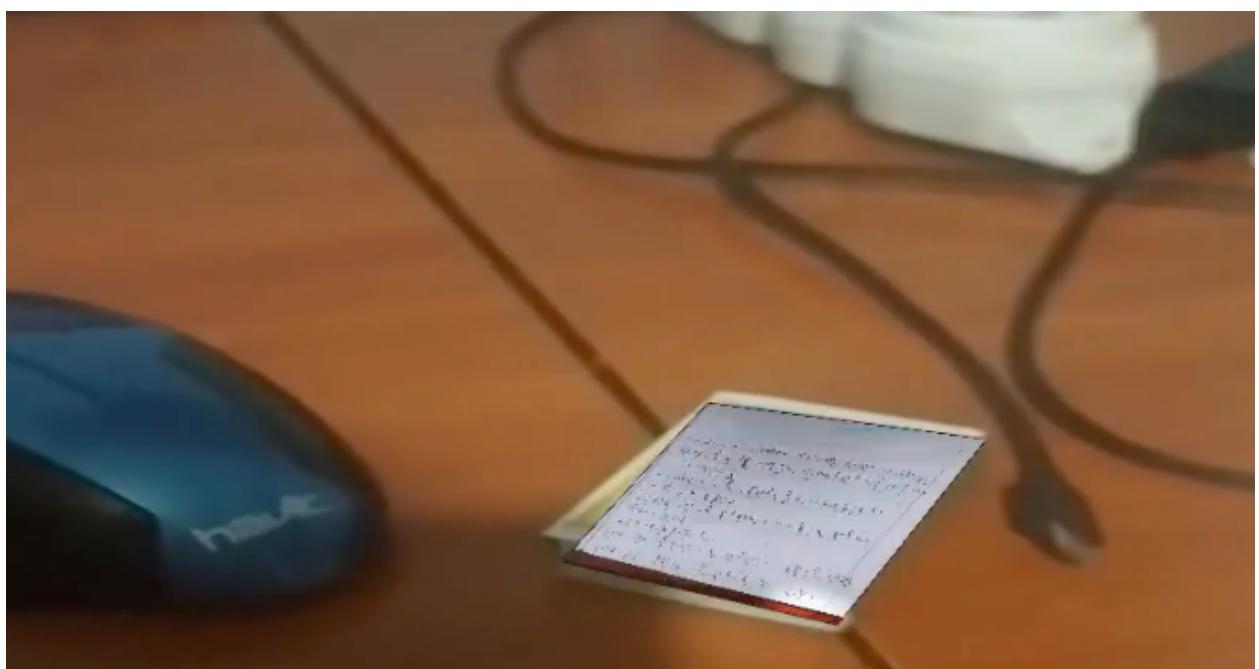
Приклад заміни:



Бачимо, що нове зображення знаходиться під тим самим кутом, що і розпізнане, досить точно його замінює та набуває тієї ж форми.



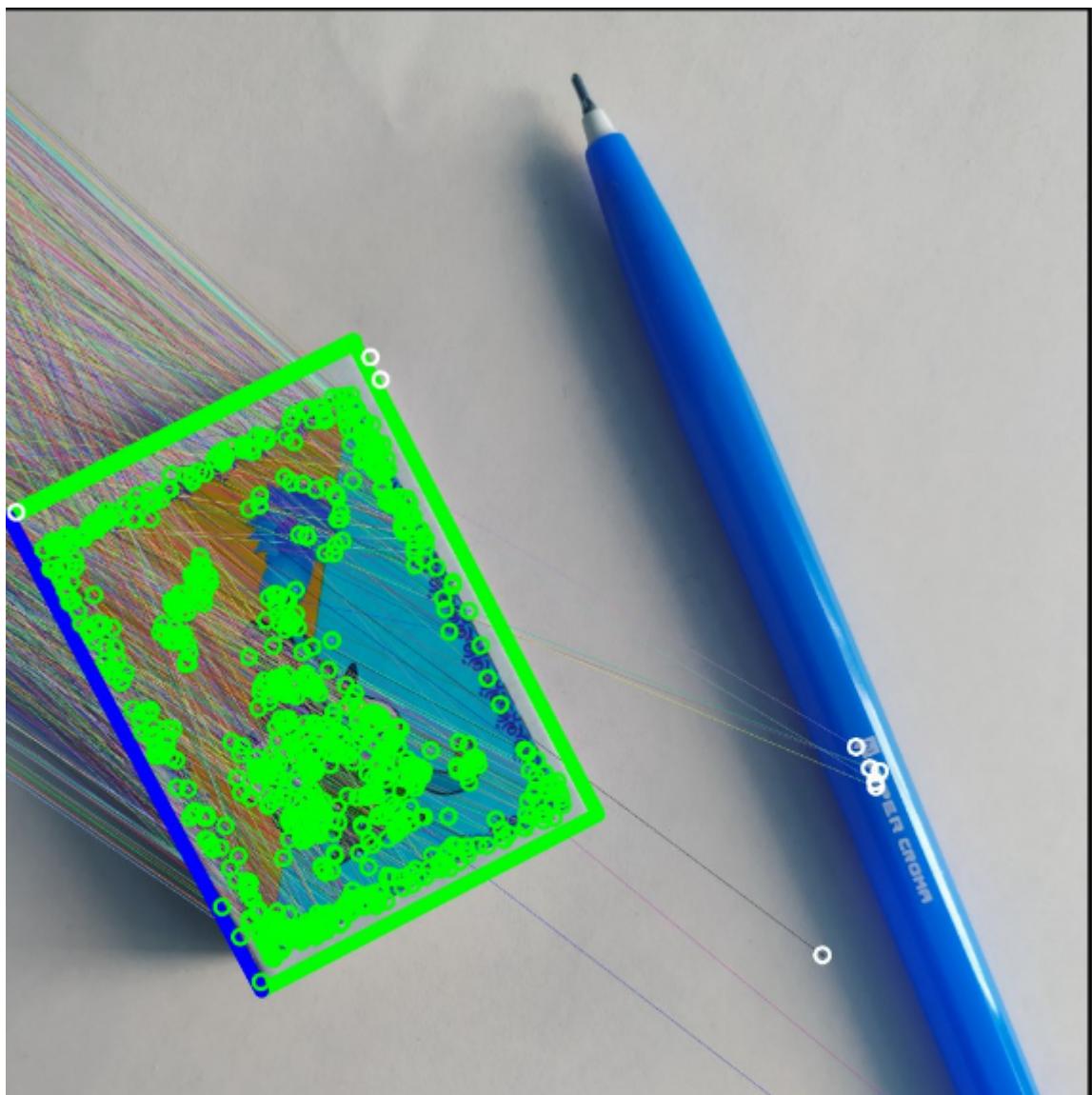
На цьому кадрі, класифікатор не визначив пачку сірників із необхідною точністю.



На цьому зображенні, класифікатор розпізнав іншу пачку сірників, тобто інший об'єкт як наш. Втім, воно досить точно замінює об'єкт, набуваючи тієї ж форми та під схожим кутом, хоч і не ідеальним.

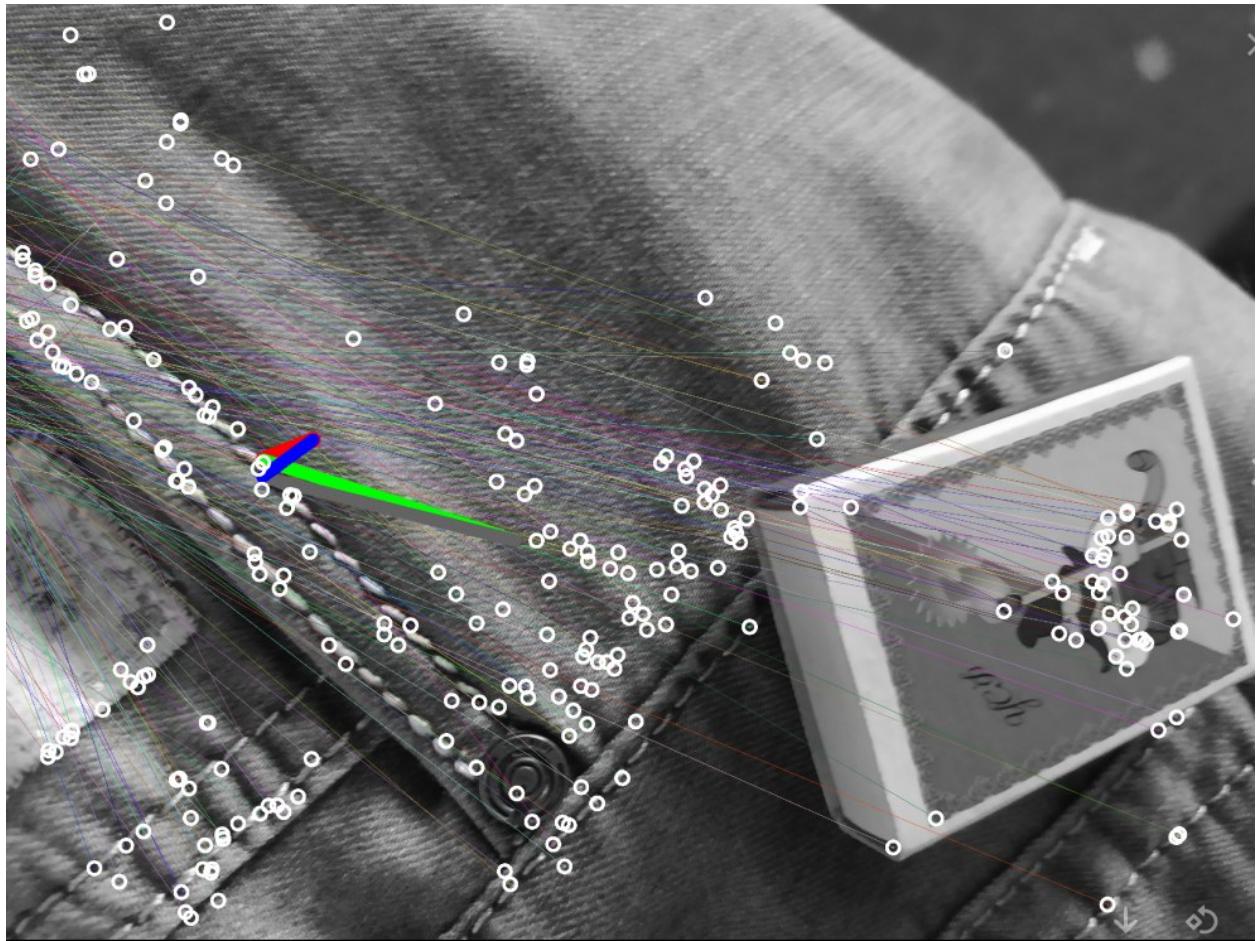
Висновки

Звісно, для якісної роботи програми із вибраним нами класифікатором, розмічений датасет має бути більшим. Ми, звісно, могли використовувати дескрипторний варіант, але результати побудови bounding box використовуючи типові методи open-cv через `perspectiveTransform` і `findHomography` змушує бажати кращого.



Ось приклад хорошої побудови. Та в більшості випадків, цей квадрат

виглядає десь так:



Використовуючи потужніший інструмент, ми змогли домогтися певних результатів.

Опис файлів для виконання роботи

lab4_cvpr.ipynb – вихідний код

trainval.json – розмічений датасет

model_final.pth – ваги натренованої моделі

my_matches_video2.mp4 – оригінал відео

target.jpg – картинка для заміни

matches_result.avi – вихідне відео

Джерела

Мануал із тренування мережі із власним датасетом

<https://www.dlogy.com/blog/how-to-train-detectron2-with-custom-coco-datasets/>

Програма для розмітки зображень - <https://github.com/Tony607/labelme2coco>