

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”  
ІНСТИТУТ КОМП’ЮТЕРНИХ НАУК ТА ІНФОРМАЦІЙНИХ  
ТЕХНОЛОГІЙ  
КАФЕДРА СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ



**ЛАБОРАТОРНА РОБОТА №1**

з дисципліни

**«Машинне Навчання»**

**Виконав:**

ст. групи КНСШ-12

Коваленко А. С.

**Перевірив:**

Телішевський П.А.

ЛЬВІВ – 2023

# Лабораторна робота №1

**Тема вибраної статті:** Recognizing Similar Musical Instruments with YOLO Models. <https://www.mdpi.com/2504-2289/7/2/94>

## Виконання роботи

### 1. Вибір моделі і даних.

У статті автор використовував модель YOLO для розпізнавання музичних інструментів по зображеннях. Для порівняння використовувались версії v7 та v5. Для початку я вибрав версію v7 з офіційного гіт-репозиторія. <https://github.com/WongKinYiu/yolov7>. Використаний датасет в статті був без визначених координат labels, тож я знайшов 2 потрібних мені датасети на ту ж тематику.

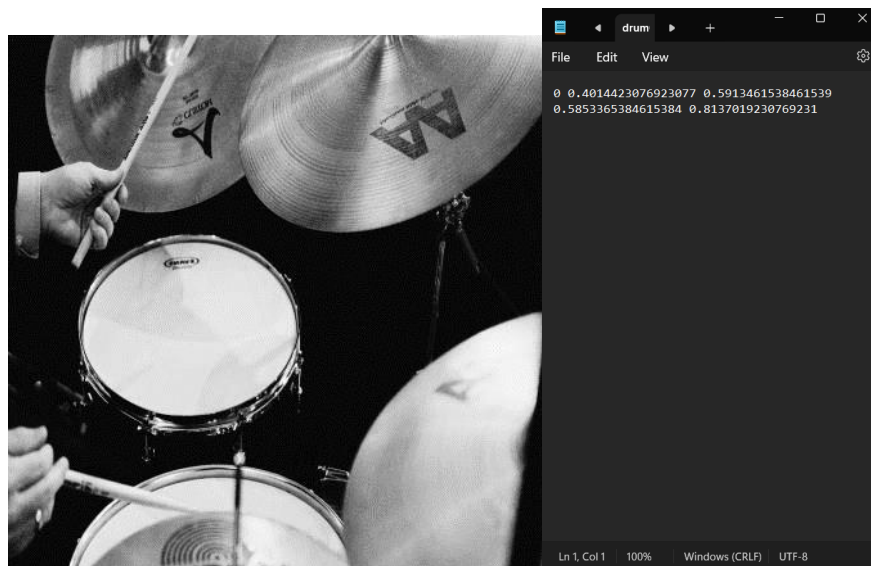


Рис.1 Приклад використаних даних

Для роботи моделі потрібно щоб назви файлів фото і labels співпадали.

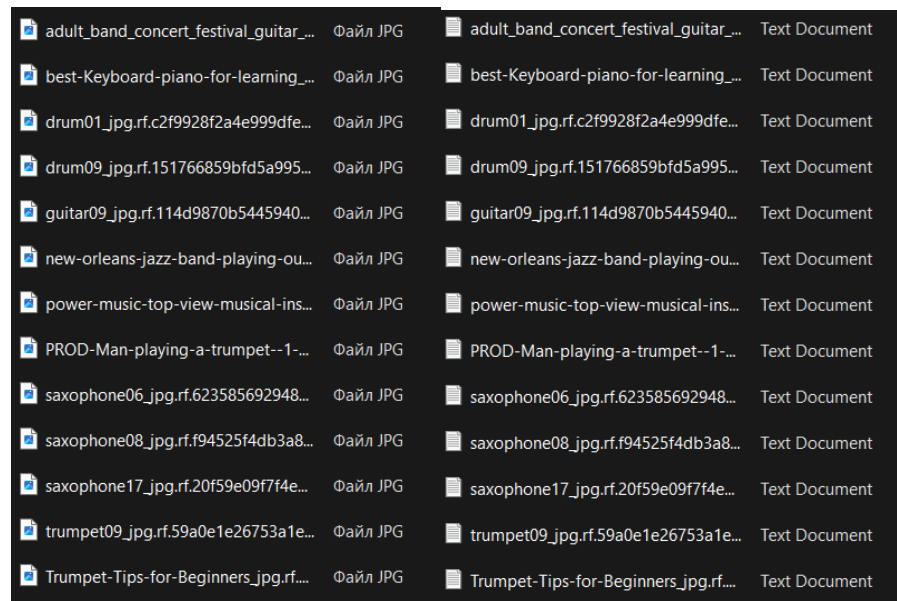


Рис.2 Вигляд файлів даних

Також для роботи YOLO моделі потрібний yaml файл в якому вказується кількість і назви класів, а також шляхи до даних.

```
train: ../train/images
val: ../valid/images

nc: 5
names: ['drums', 'keyboards', 'saxophones', 'strings', 'trumpets']
```

В першому датасеті відсутній yaml файл, тож створюю вручну.

## 2. Тренування YOLOv7.

Для початку експериментую з v7. Буду тренувати модель за своїм датасетом, на вже натренованих вагах, які завантажую з офіційного гіту.

```
!git clone https://github.com/WongKinYiu/yolov7
%cd yolov7
!pip install -r requirements.txt
%cd /content/yolov7
!wget
https://github.com/WongKinYiu/yolov7/releases/download/v0.1/yolov7_trainin
g.pt
```

Починаю тренування моделі. Налаштовую кількість епох, в даних, посилаюсь на файл custom.yaml, який я змінив для свого датасету.

nc: 34

names:- Bango- Breton Bombarde- Drum- Flute- Guitar-Vestern-- Guitar-acoustic-- Guitar-bass-- Guitar-classic-- Guitar-electric-- Guitar-folk-- Guitarron- Guzheng- Loudspeakers- Mixer- Psaltery- Trumpet- Vibraphone- Vuvuzela- accordion- balalaika- dan bau- drum bass- drum ride- gusli- lahori gogona- lute- maraca- pana- piano- sitar- synthesizer- tabla- tambourine- violin

```
!python train.py --batch 16 --cfg cfg/training/custom yolov7.yaml --epochs 50 --data /content/yolov7/data/custom.yaml --weights 'yolov7.pt' --device 0
```

### Результати тренування:

Epoch	gpu_mem	box	obj	cls	total	labels	img_size
49/49	14.2G	0.04666	0.02356	0.03988	0.1101	56	640: 100% 5/5 [00:04<00:00, 1.08it/s]
	Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95: 100% 1/1 [00:00<00:00, 1.25it/s]
	all	20	71	0.187	0.272	0.0769	0.0346
	Drum	20	11	0	0	0.0183	0.0039
	Flute	20	2	0	0	0.00646	0.00259
	Guitar-acoustic-	20	5	0	0	0.00853	0.00129
	Guitar-bass-	20	4	0.0437	0.5	0.267	0.138
	Guitar-classic-	20	15	0.102	0.667	0.338	0.215
	Guitar-electric-	20	4	0.0163	0.5	0.103	0.0233
	Guzheng	20	1	1	0	0.0128	0.00128
	Psaltery	20	5	1	0	0	0
	Trumpet	20	11	0	0	0.0276	0.00952
	balalaika	20	1	0	0	0	0
	dan bau	20	1	0	0	0	0
	drum bass	20	2	0.0366	0.5	0.0216	0.00862
	drum ride	20	1	0.0999	1	0.249	0.0498
	gusli	20	1	0.237	1	0.249	0.124
	lute	20	1	0	0	0.0151	0.00302
	maraca	20	2	0	0	0.00978	0.00293
	piano	20	2	0	0	0.0338	0.0236
	synthesizer	20	1	0.0163	1	0.0995	0.0498
	violin	20	1	1	0	0	0
50 epochs completed in 0.175 hours.							
Optimizer stripped from runs/train/exp/weights/last.pt, 75.2MB							
Optimizer stripped from runs/train/exp/weights/best.pt, 75.2MB							

Рис.3 Кінець тренування моделі

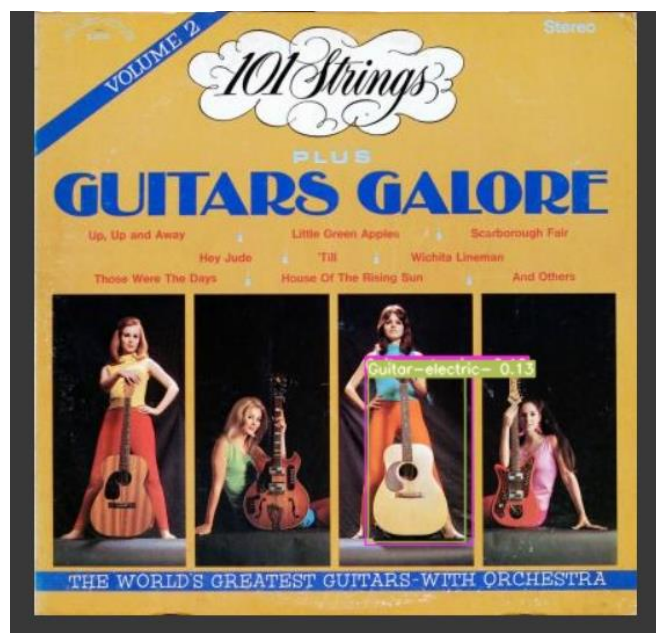


Рис.4 Приклад виводу

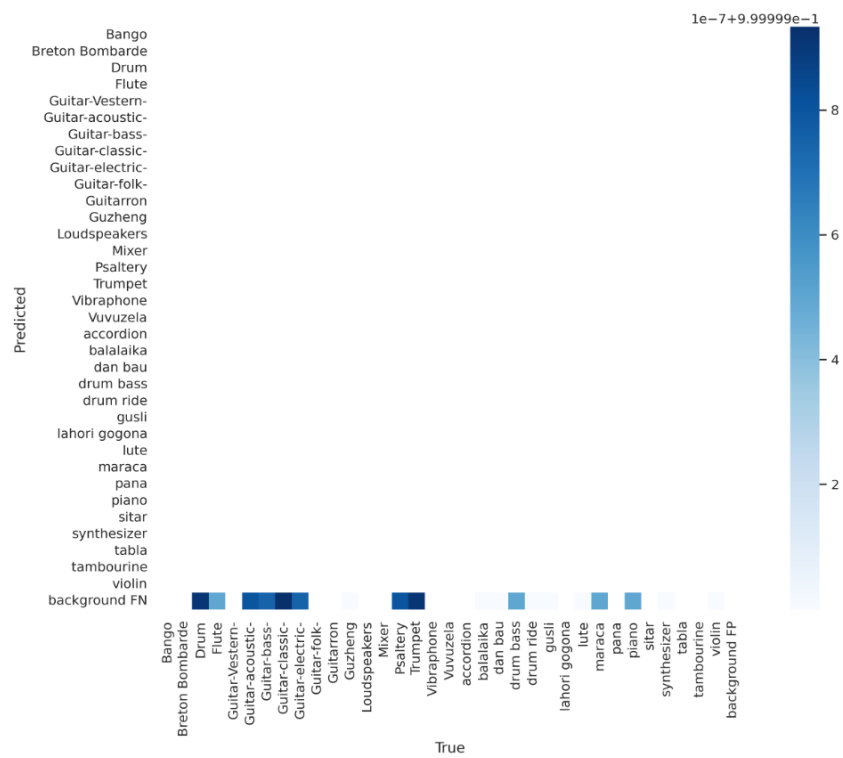


Рис.5 confusion\_matrix

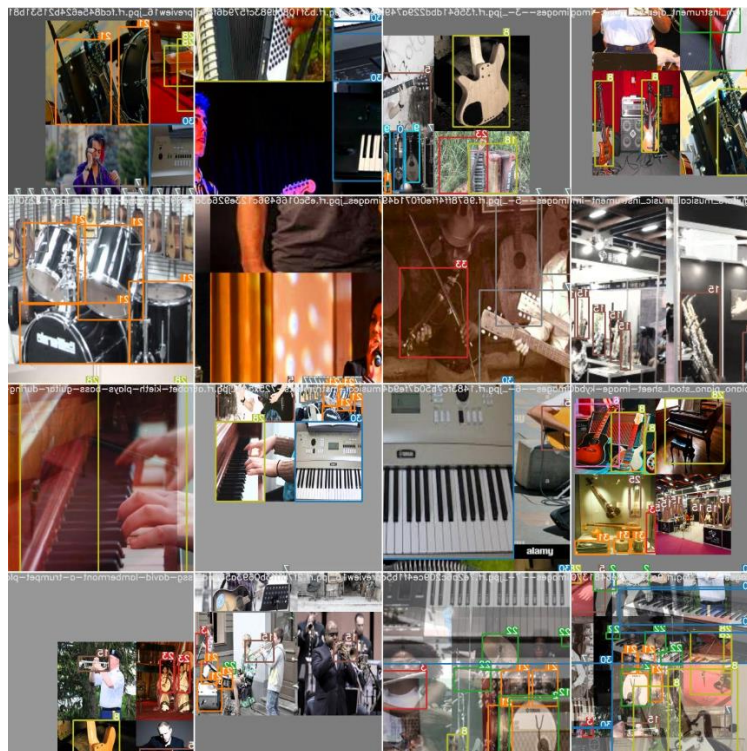


Рис.5 Процес тренування

Epoch	gpu_mem	box	obj	cls	total	labels	img_size		
99/99	12.7G	0.05295	0.01026	0.02598	0.08919	50	640:	100%	6/6 [00:06<00:00, 1.08s/it]
	Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:	100%	1/1 [00:01<00:00, 1.49s/it]
	all	28	72	0.256	0.322	0.199	0.114		
	drums	28	13	0.0409	0.0769	0.0319	0.0111		
	keyboards	28	18	0.238	0.444	0.268	0.145		
	saxophones	28	16	0.417	0.438	0.272	0.166		
	strings	28	14	0.085	0.286	0.14	0.0724		
	trumpets	28	11	0.501	0.364	0.285	0.179		

100 epochs completed in 0.374 hours.

Optimizer stripped from runs/train/exp2/weights/last.pt, 74.9MB  
Optimizer stripped from runs/train/exp2/weights/best.pt, 74.9MB

Рис.6 Тренування другого датасету зі збільшенням епох

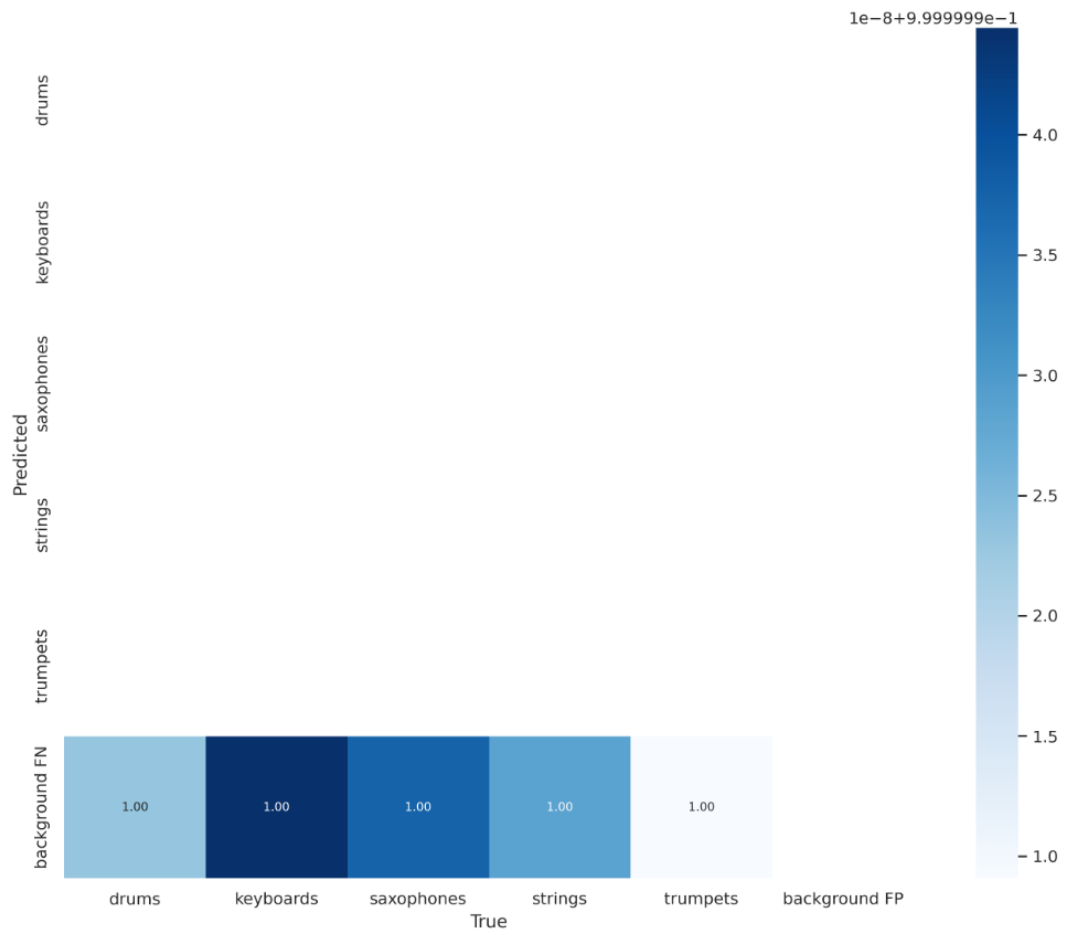


Рис.7 confusion\_matrix для моделі натренованої на другому датасеті.

### 3. Тренування YOLOv5.

Після невдалого експерименту з v7, пробує YOLO v5

<https://github.com/ultralytics/yolov5>.

```
!python train.py --img 415 --batch 16 --epochs 150 --data
/content/yolov5/musicBandInstruments00-2/data.yaml --weights yolov5x.pt --
cache
```

```
!python detect.py --weights runs/train/exp2/weights/best.pt --conf 0.1 --
source /content/yolov5/musicBandInstruments00-2/test/images
```



```
150 epochs completed in 0.344 hours.  
Optimizer stripped from runs/train/exp2/weights/last.pt, 173.1MB  
Optimizer stripped from runs/train/exp2/weights/best.pt, 173.1MB  
  
Validating runs/train/exp2/weights/best.pt...  
Fusing layers...  
Model summary: 322 layers, 86200330 parameters, 0 gradients, 203.8 GFLOPs  
Class      Images  Instances  P      R      mAP50  mAP50-95: 100% 1/1 [00:00<00:00, 1.85it/s]  
all        28      72        0.716  0.505  0.529   0.295  
drums      28      13        0.649  0.429  0.453   0.18  
keyboards  28      18        0.906  0.611  0.657   0.298  
saxophones 28      16        0.62   0.562  0.526   0.282  
strings    28      14        0.527  0.286  0.308   0.193  
trumpets   28      11        0.878  0.636  0.7     0.523  
Results saved to runs/train/exp2
```

Рис.8 Тренування YOLOv5.



Рис.9 Порівняння розпізнавання на 50 і 150 епохах.

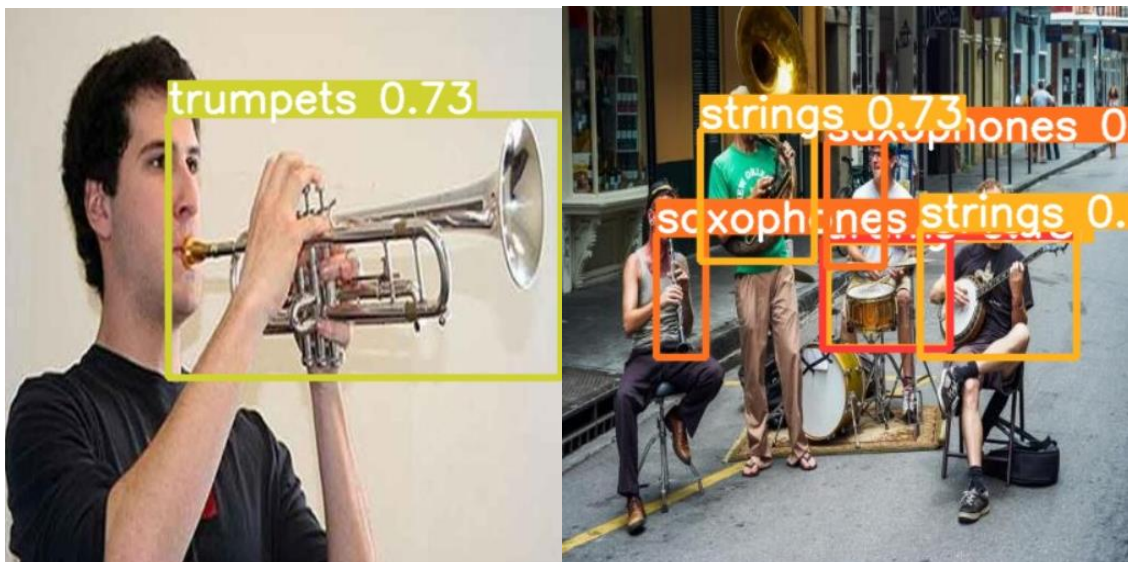


Рис.10 Розпізнавання на 150 епохах.



Рис.11 Розпізнавання

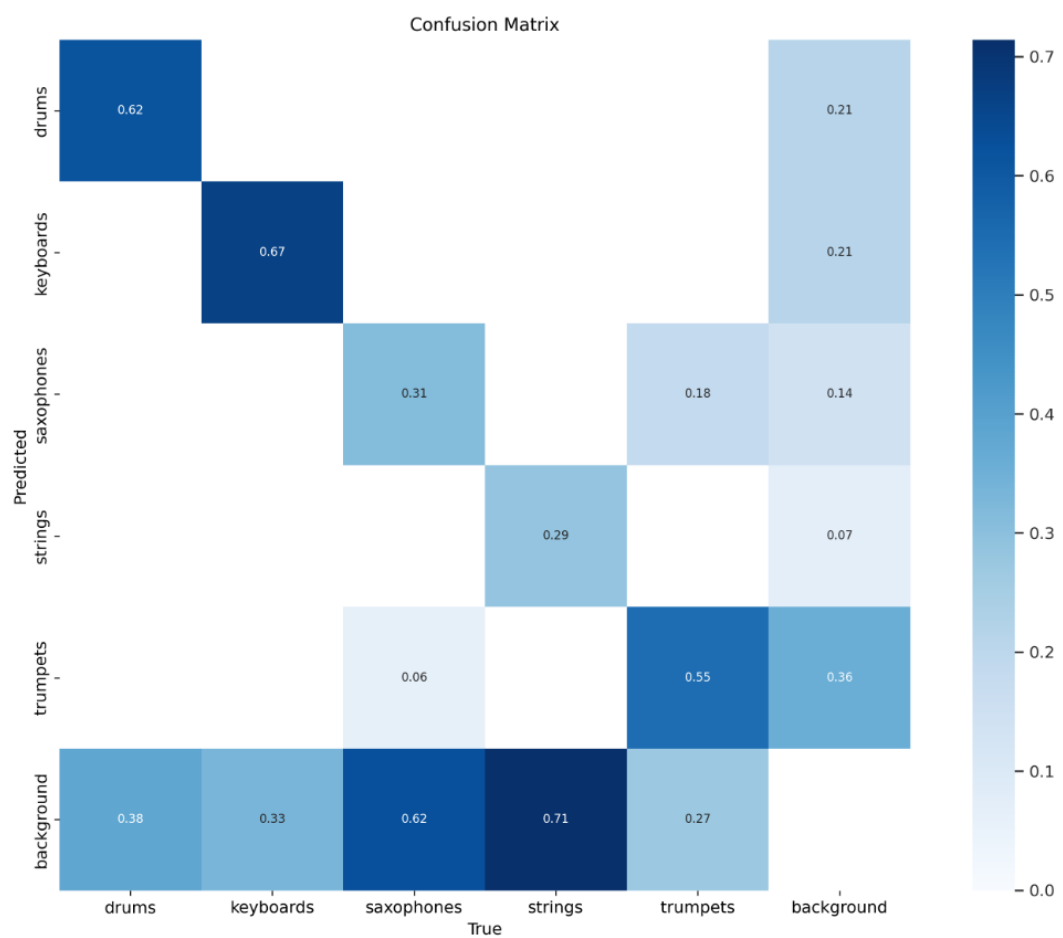


Рис.12 confusion\_matrix

Зберігаю отриману модель і результати

```
!zip -r export.zip runs/detect
!zip -r export.zip runs/train/exp/weights/best.pt
!zip export.zip runs/train/exp/*
```



## Тест моделі на іншому датасеті

```
!python detect.py --weights runs/train/exp2/weights/best.pt --conf 0.1 --source /content/yolov7_data/images/train
```

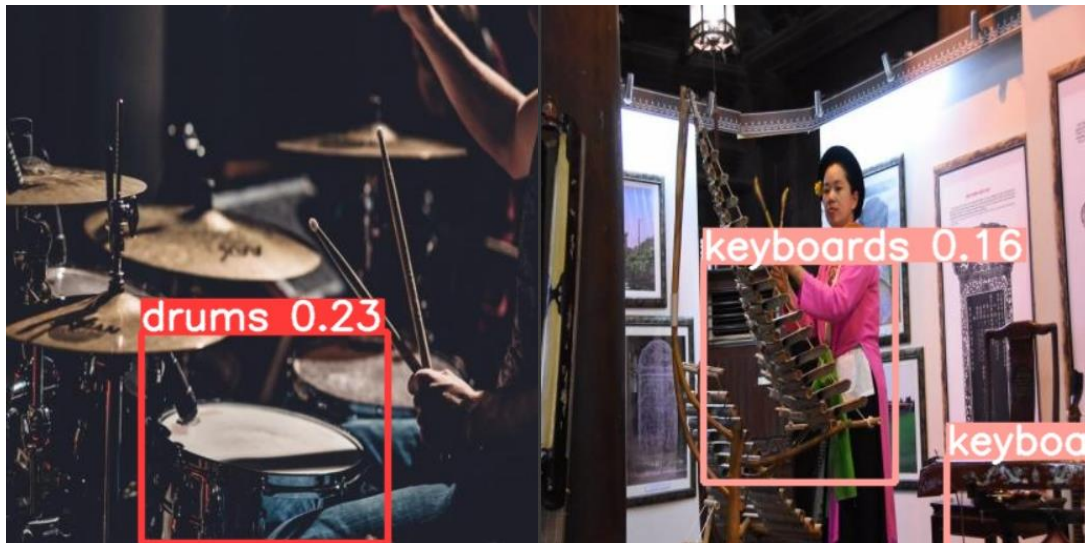
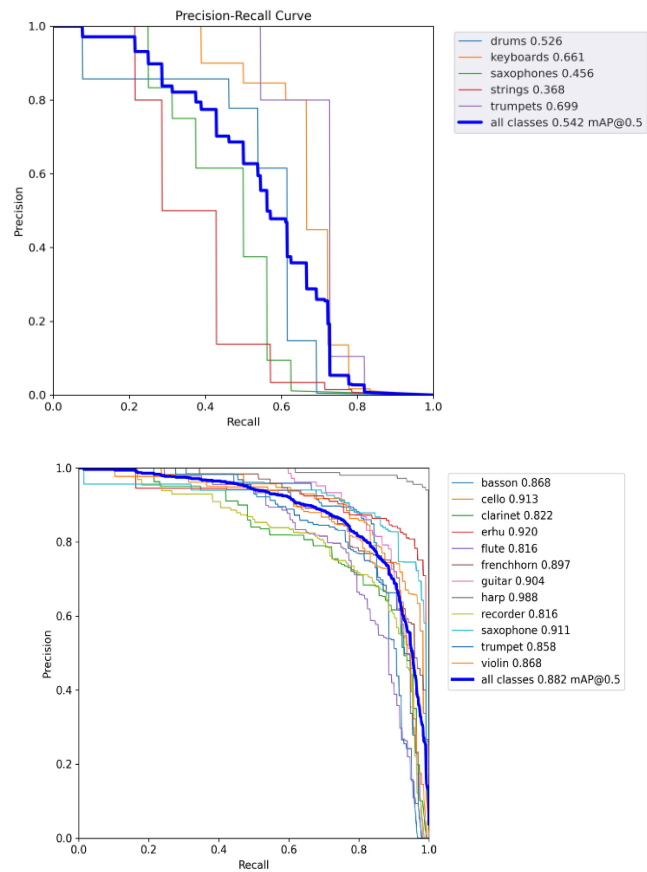


Рис.13 Розпізнавання на другому датасеті

## 4. Порівняння результатів.

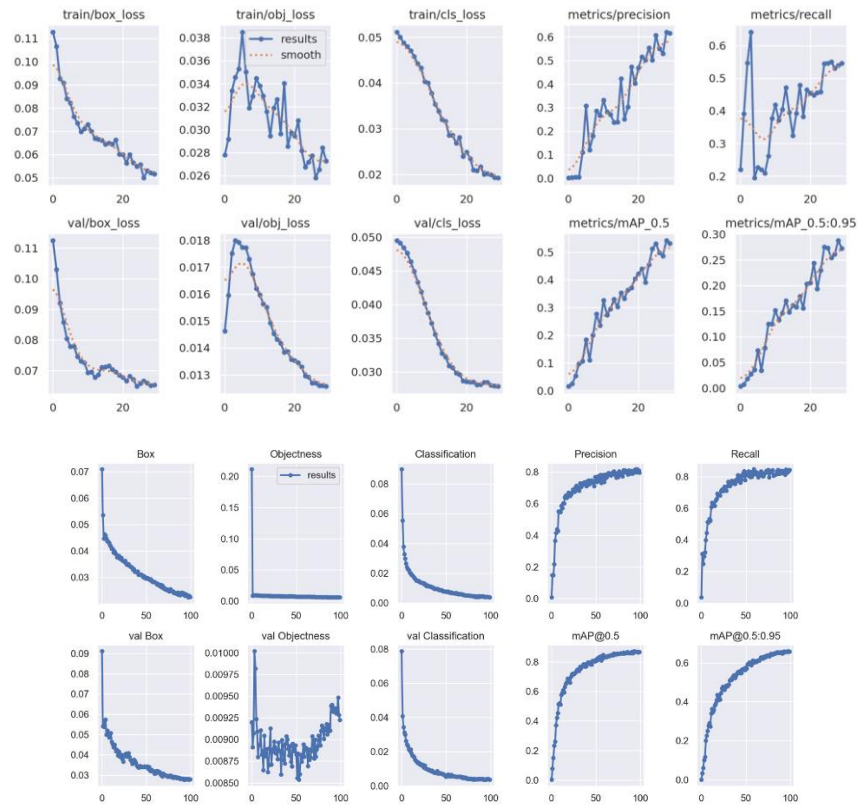
Порівнюю власні результати із автором статті. Можна зазначити що автор використовував

‘ Nvidia RTX3060Ti GPU accelerator with 11 GB of RAM, an i7 central processing unit (CPU), and 16 GBDDR2’ для тренування. Я використовував GPU від Google colab.



(a)

Рис.14 Precision and recall криві



(b)

Рис.15 training and validation curves

## **Висновок.**

На лабораторній роботі досліджував YOLO моделі для розпізнавання музичних інструментів. Модель v5 показала результат, коли v7 результату не дала. В порівнянні з статтею результат набагато гірший, в наступних лабораторних буду намагатись покращити результат. Щодо моделі v7 доцільно спробувати дослідити інший, більший датасет, для v5 збільшити кількість епох.