AI-based Pilgrim Detection using Convolutional Neural Networks

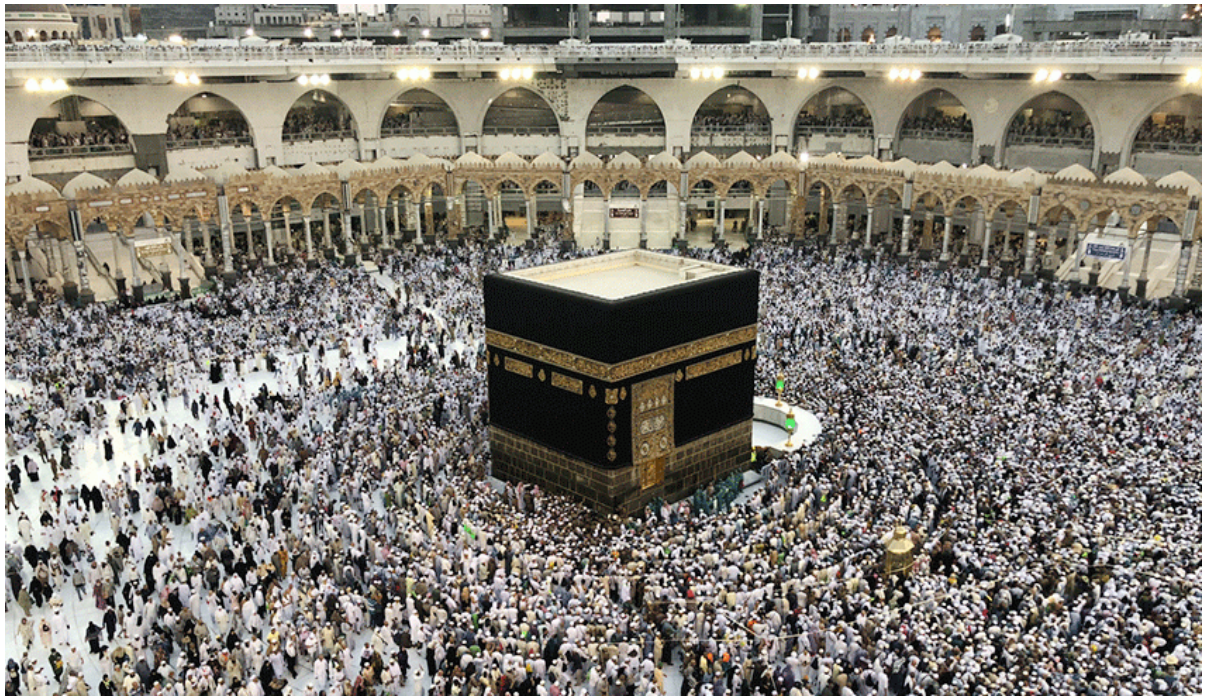
MARWA BEN JABRA , ADEL AMMAR , ANIS KOUBAA , OMAR CHEIKHROUHOU , HABIB HAMAM

Plan

- INTRODUCTION
- ALGORITHMS BACKGROUND
- THE PILGRIMS DATASET
- EXPERIMENTAL EVALUATION
- CONCLUSION
- PERSPECTIVES

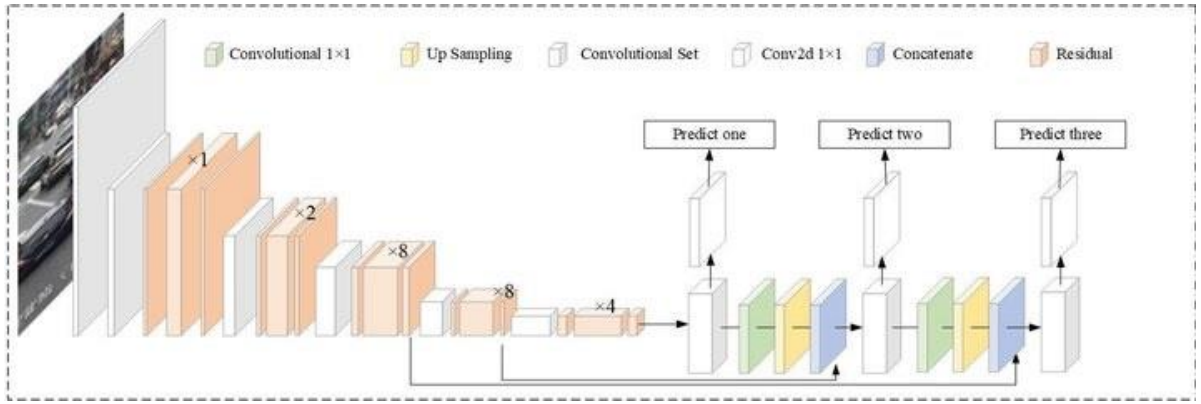
INTRODUCTION

- The safety and security of pilgrims is the highest priority for the authorities.
- CNN to detect and identify Pilgrims and their features.
- Dataset for the detection of pilgrims and their genders.
- Two CNN models based on YOLOv3 and Faster-RCNN.

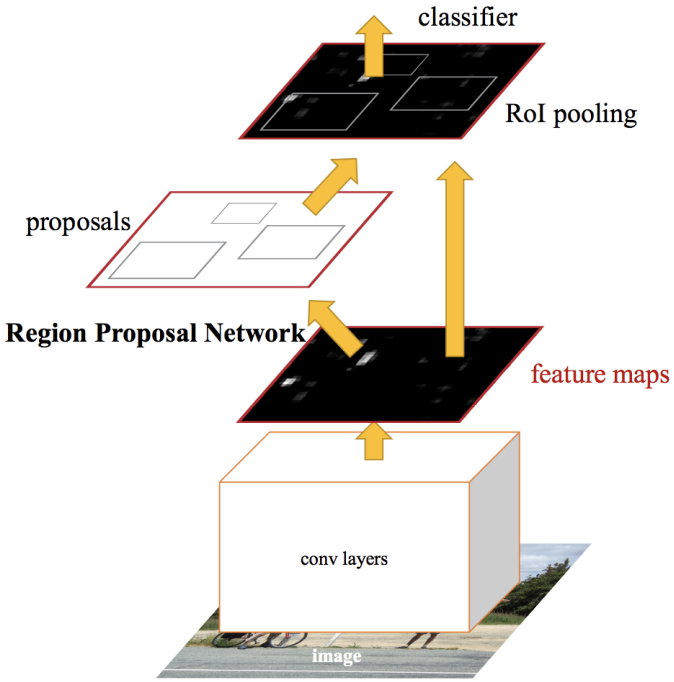


ALGORITHMS BACKGROUND

YOLOv3 is a one-stage detector that is known to be the fastest detection algorithm.



Faster R-CNN is an improvement of R-CNN that represents the most efficient region-based CNN algorithm for image detection.



Faster R-CNN

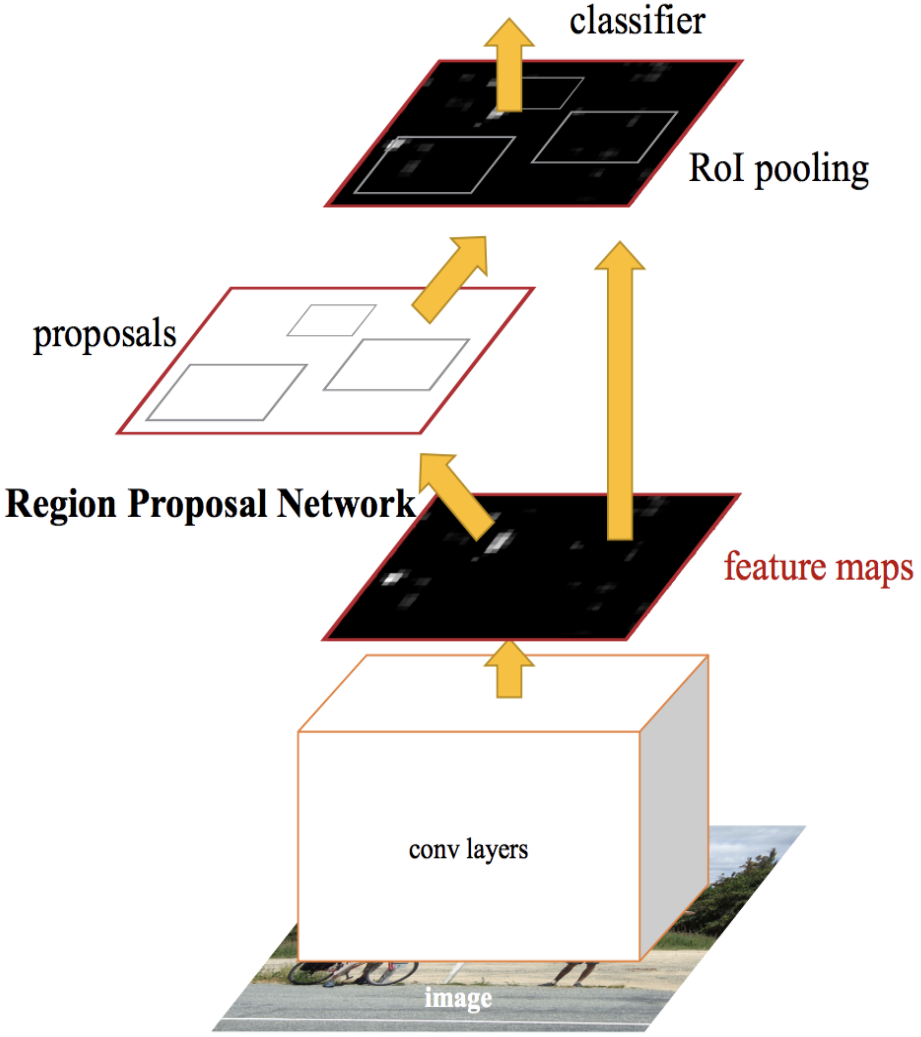
❖The RPN module generate the region proposals.

❖The Fast R-CNN detector :

- The extraction of feature vectors from the region of interest.
- The feature vector obtained is the input of the classifier

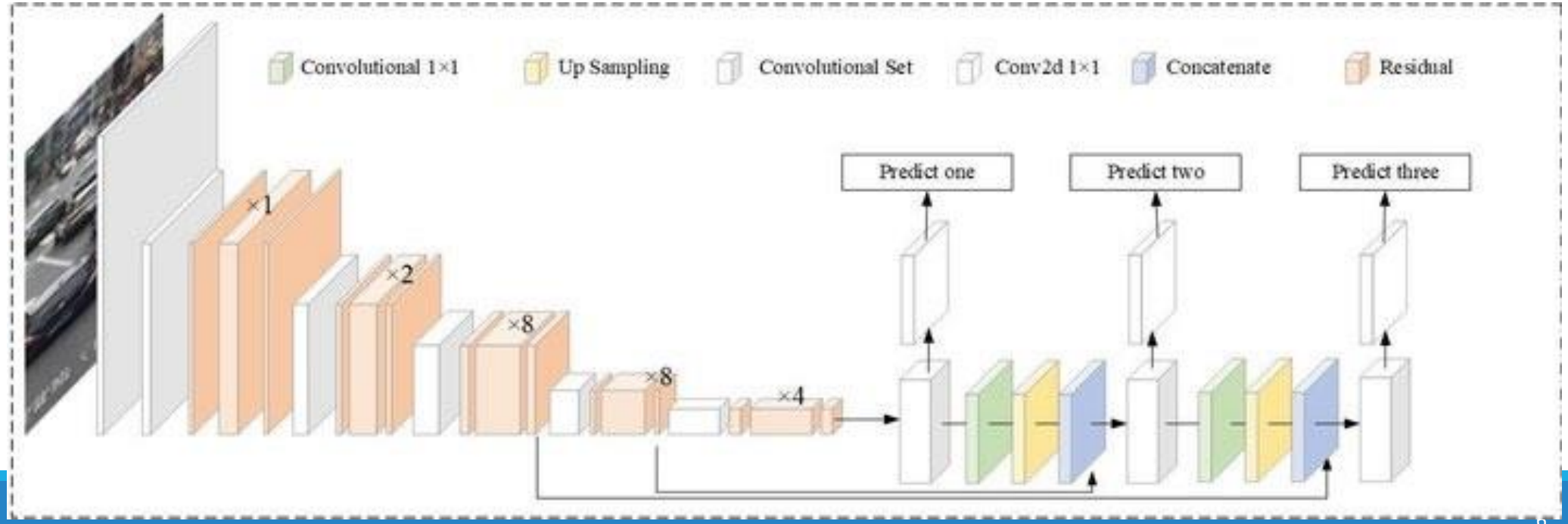
The classification output are :

- A sequence of probabilities estimated of the different objects considered.
- The coordinates of the regions proposals.



YOLOv3

- Multi-label classification based on logistic regression.
- Cross-entropy loss function.
- The prediction of bounding boxes .
- The concept of Feature Pyramid Network for the prediction
- Darknet-53 CNN features extractor.



THE PILGRIMS DATASET

Pilgrim, Not Pilgrim classes designate a male

(a)

(b)

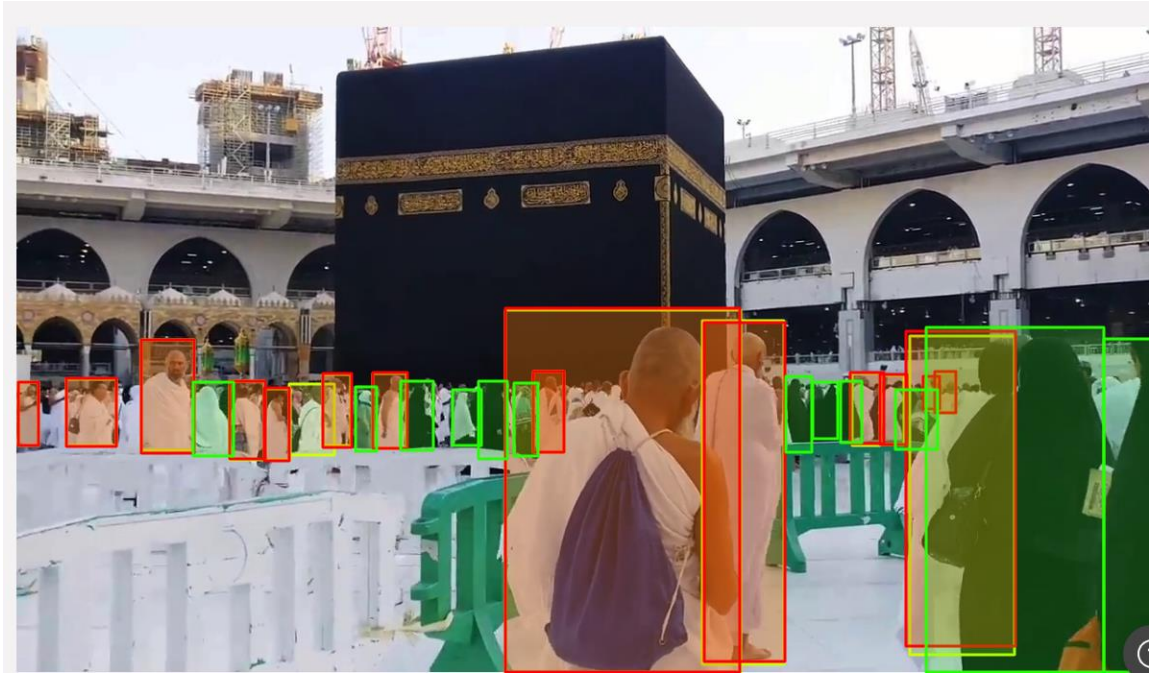
Woman classes with no additional feature

THE PILGRIMS DATASET

- 622 images of people in the holy places of Makkah and Madinah.

- Using the LabelImg software, we labeled the collected dataset into three chosen labels.

		Training	Testing	Total
Number of images		**560**	**62**	**622**
Number of instances	**Pilgrim men**	**1228**	**111**	**1339**
	Non-pilgrim men	**859**	**111**	**970**
	Women	**1016**	**162**	**1178**



EXPERIMENTAL SETUP

- For Faster R-CNN: Inception-v2 and ResNet50
- For YOLOv3: (320x320), (416x416), and (608x608) pixels.
- For the learning rate:
  - YOLOv3: an initial rate of 0.001,
  - Faster R-CNN: an initial rate of 0.0002 with Inception-v2 and 0.0003 with ResNet50.
- The weight decay value of 0.0005.
- Stochastic Gradient Descent (SGD) of momentum (0.9).

	Machine 1	Machine 2
CPU	**Intel Core i7-8700K (3.7 GHz)**	**Intel Core i9-9900K (Octa-core)**
Graphics card	**NVIDIA GeForce 1080 (8 GB) GPU**	**NVIDIA GeForce RTX 2080T (11 GB) GPU**
RAM	**32GB**	**64GB**
Operating system	**Linux (Ubuntu 16.04 TLS)**	**Linux (Ubuntu 16.04 TLS)**

EXPERIMENTAL METRICS

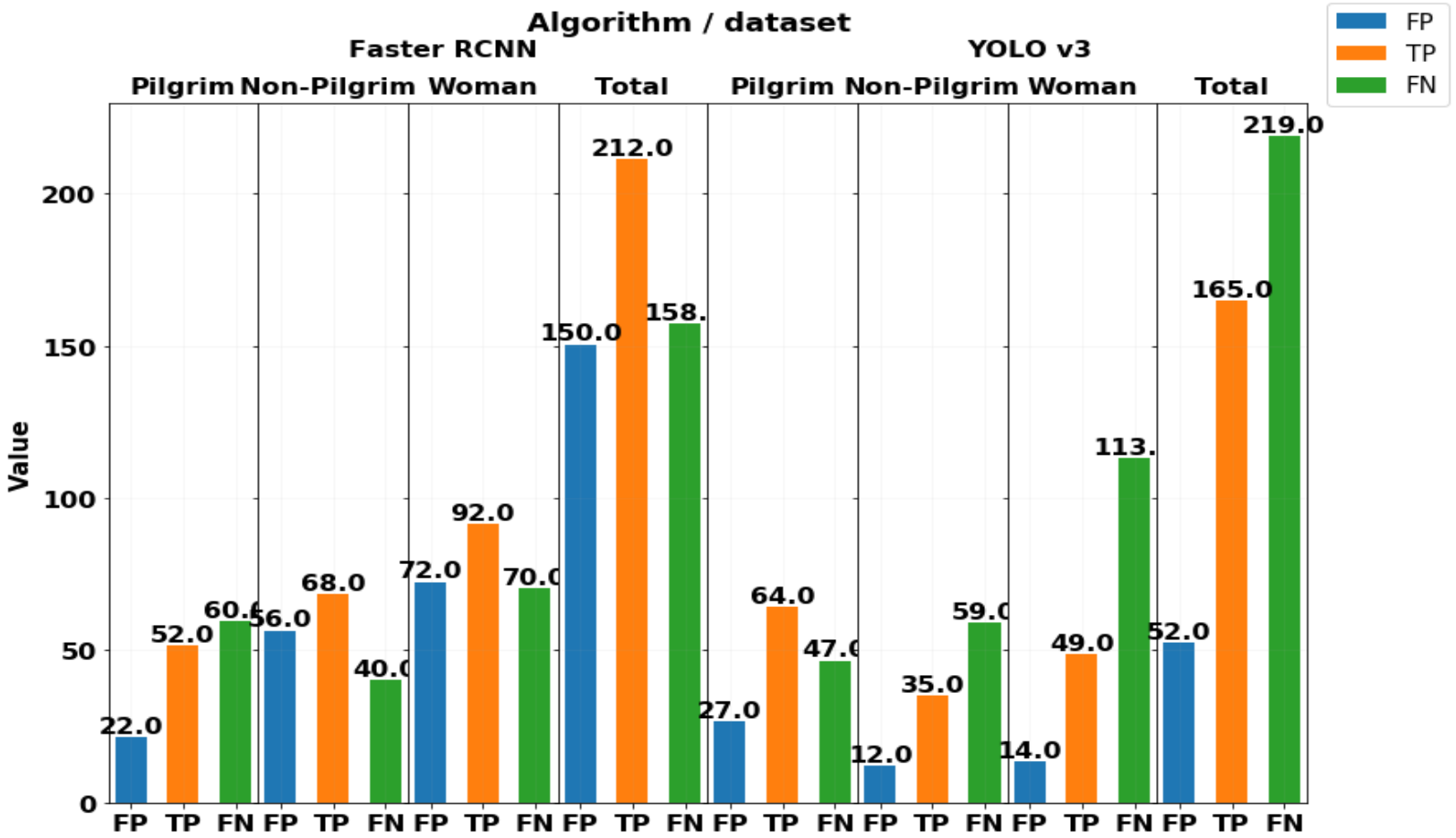
For the evaluation used six metrics :

- **True Positive (TP)**
- **False Positive (FP)**
- **False Negative (FN)**
- **Precision** = TP/(TP + FP)
- **Recall** = TP/(TP + FN)
- **F1score** = (2 ∗ Precision ∗ Recall)/(Precision + Recall)
- **Quality** = TP/(TP + FP + FN)
- **mIoU**: mean of the Intersection over Union.
- **mAP**: mean Average Precision
- **FPS**: frame per second.

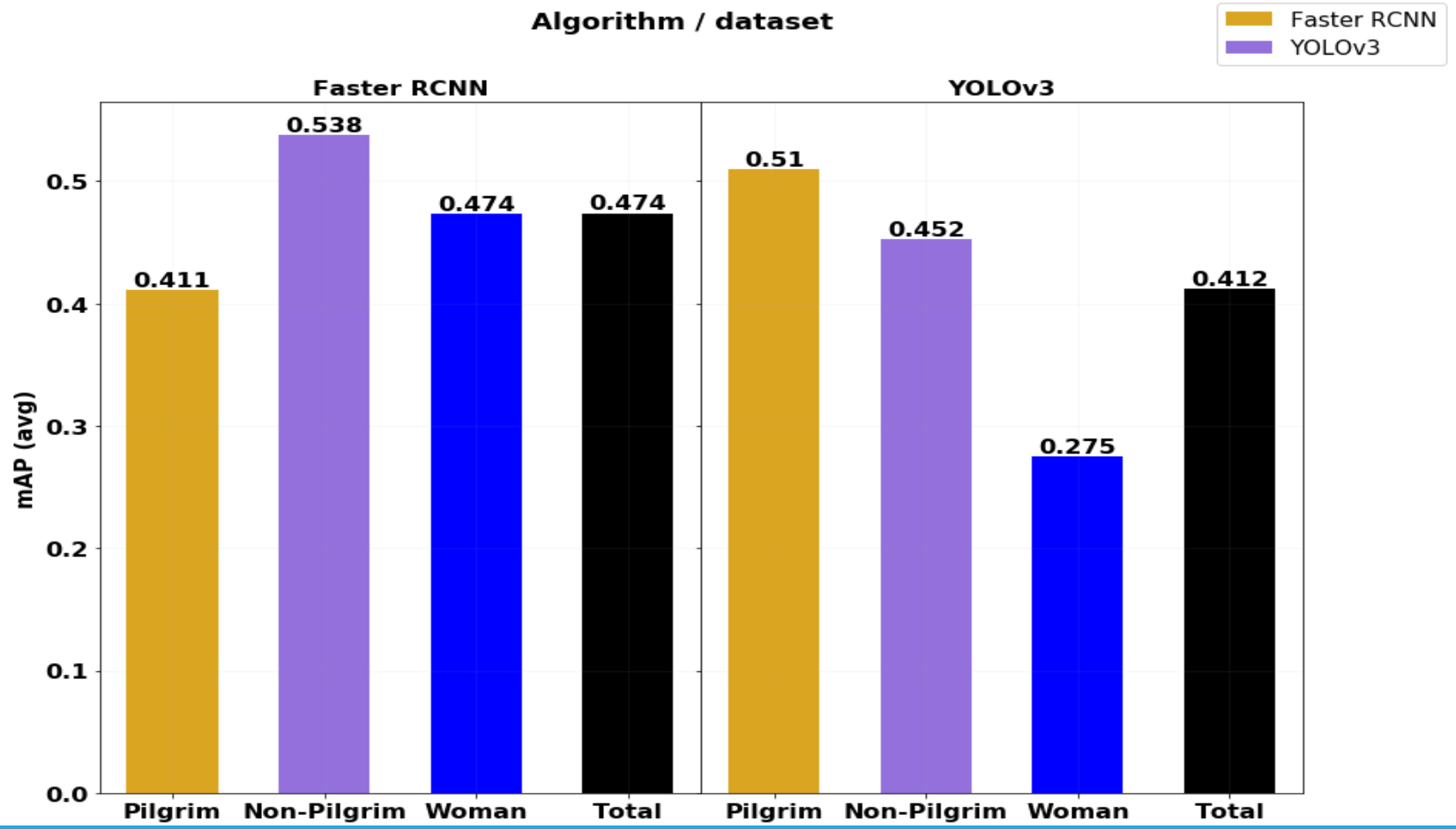
Comparison between Faster R-CNN and YOLO v3

Algorithm	YOLOv3 (320x320)px	YOLOv3 (416x416)px	YOLOv3 (608x608)px	Faster R-CNN (Inception v2)	Faster R-CNN (ResNet 50)
FP	40	66	51	164	137
TP	159	171	165	228	195
FN	225	213	219	156	189
Precision	0.8058	0.7288	0.7735	0.6091	0.6022
Recall	0.4349	0.4577	0.4557	0.5929	0.5042
Quality	0.3905	0.3849	0.3923	0.4172	0.3744
F1score	0.5541	0.5546	0.5546	0.5887	0.5446
mAP	0.3999	0.4152	0.4214	0.5162	0.4317
mIoU	0.6352	0.5988	0.6192	0.5710	0.5850
FPS	91.28	65.31	43.84	3.35	3.8

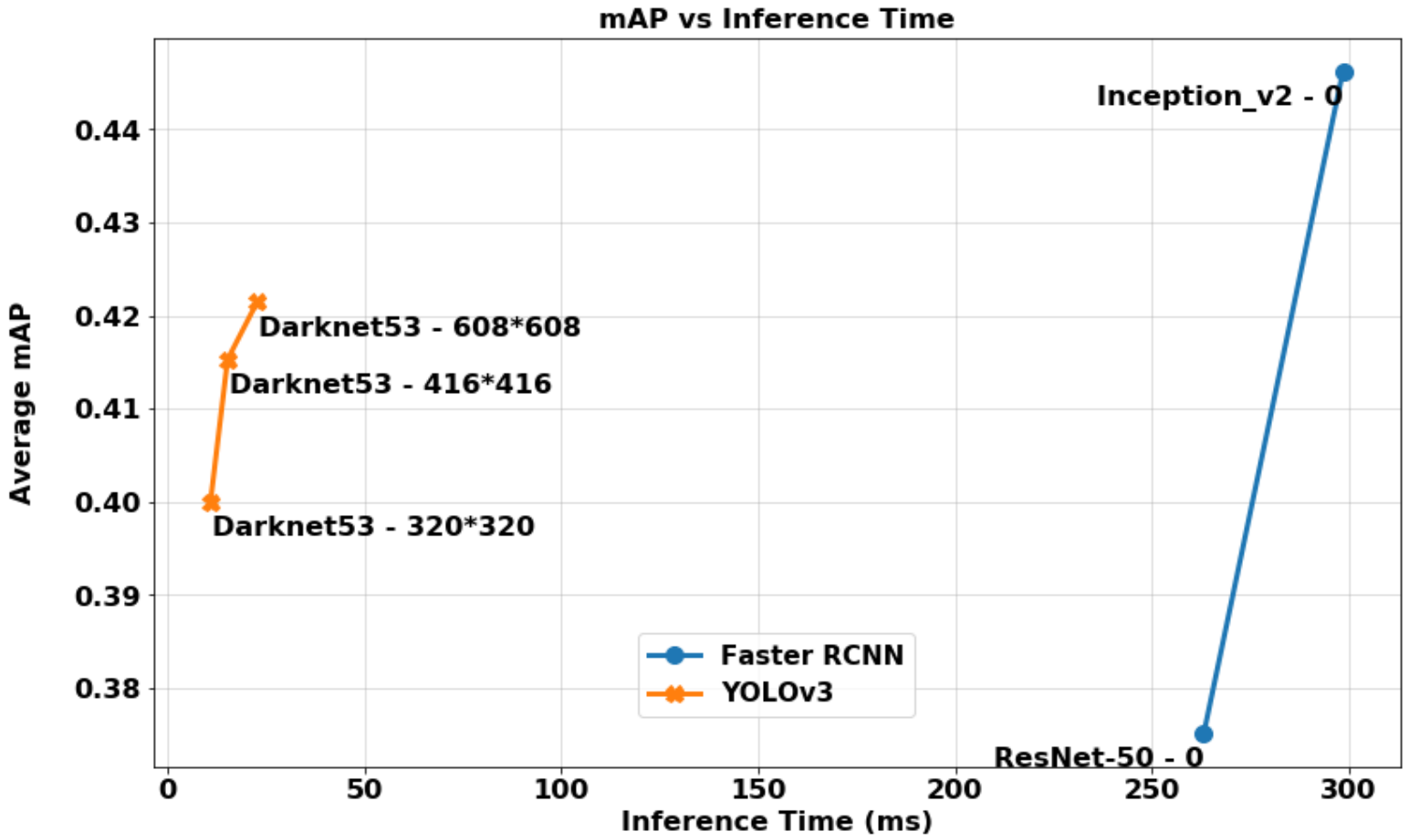
TP, FP and FN


Algorithm / dataset
Faster RCNN
YOLO v3
Pilgrim
Non-Pilgrim
Woman
Total
FP
TP
FN
Value
0
50
100
150
200
22.0
52.0
56.0
68.0
72.0
92.0
150.0
212.0
27.0
64.0
12.0
35.0
14.0
49.0
52.0
165.0
219.0

mAP: mean Average Precision



FPS: frame per second.



CONCLUSION

- We developed convolutional neural network models for pilgrim detection for Al-Hajj based on YOLOv3 and Faster RCNN.
- We have built a dataset containing three classes of pilgrims, non-pilgrims and women.
- Experimental results show that Faster RCNN with Inception v2 feature extractor provides the best mean average precision over all classes with 51%, comparable to state-of-the-art object detection algorithms.

Perspective

- We will extend the dataset to have several tens of thousands of instances to improve the overall accuracy and precision,
- We will consider more classes.
- We aim at developing a search application for lost people during Hajj and Umrah.

Demo

Link of demo: https://www.youtube.com/watch?v=L-nmYBY2pvE

Thanks for your attention