Object extraction techniques and visual image search with Semantic web techniques

Aninda Maulik

Supervisors: Prof. Pierre Maret
Dennis Diefenbach

Cyber Physical and Social Systems
University of Jean Monnet

May 2020

Content

- Implementation of an Algorithm for object extraction.
- Design of a semantic web modelling for extracted data.
- Implementation of a visual image search engine through Qanswer.

Abstract

- My work is divided into three parts and the first part of it is implementation of an already existing algorithm for object extraction. The algorithm chosen here, is YOLO(You Only Look Once). This algorithm predicts bounding boxes of objects and class probabilities of the detected objects directly from full images in one evaluation. Later YOLO would be combined with another alogorithm named (RSIDE)"Revisiting Single Image Depth Estimation" to detect the depth of the detected objects in order to better facilitate the detection of object position.

Abstract contd...

- The second part of my work is applying semantic web modelling to the generated data, to later query the same from Qanswer. The semantic web modelling of data depends on type of data that we want to use to convert to rdf. Here, we choose CSV format. CSV format is simple and practical, but it is difficult to express the relationships between data fields, and user access approaches/rights, etc. In order to make the CSV data semantically structured, interoperable, accessible and reusable for various Web applications, they need to be converted into the Resource Description Framework (RDF) format that provides superior data assimilation and query functionality.
- Finally we query the resultant RDF using Qanswer.

Implementation of an Algorithm for object extraction.

- Initiate a computer vision api and implement an algorithm in order to send a set of images only to receive annotated version of the same set
- Identify the computer vision algorithm
- The identified computer vision algorithm should be within the PyTorch framework because the majority of depth estimation algorithms are implemented in PyTorch. In this way we could use object detection and depth estimation in the same platform.
- The initial identified computer vision algorithm for object detection is YOLO(You Only Look Once) under PyTorch Framework
- The output of YOLO :

```
124 87 183 150 surfboard 0.988910973072052
108 67 167 142 person 0.9777714610099792
134 88 184 149 surfboard 0.39430269598960876
```

- The corresponding identified computer vision algorithm for depth estimation of the detected objects is (RSIDE)"Revisiting Single Image Depth Estimation"

Implementation of an Algorithm for object extraction.

YOLO+RSIDE and more:

- The expected output for YOLO+RSIDE:

Nb of images	100	
Nb of objects	15	Max: 26
Image size X axis	1	500
Image size Y axis	1	500
Depth scale	1	10

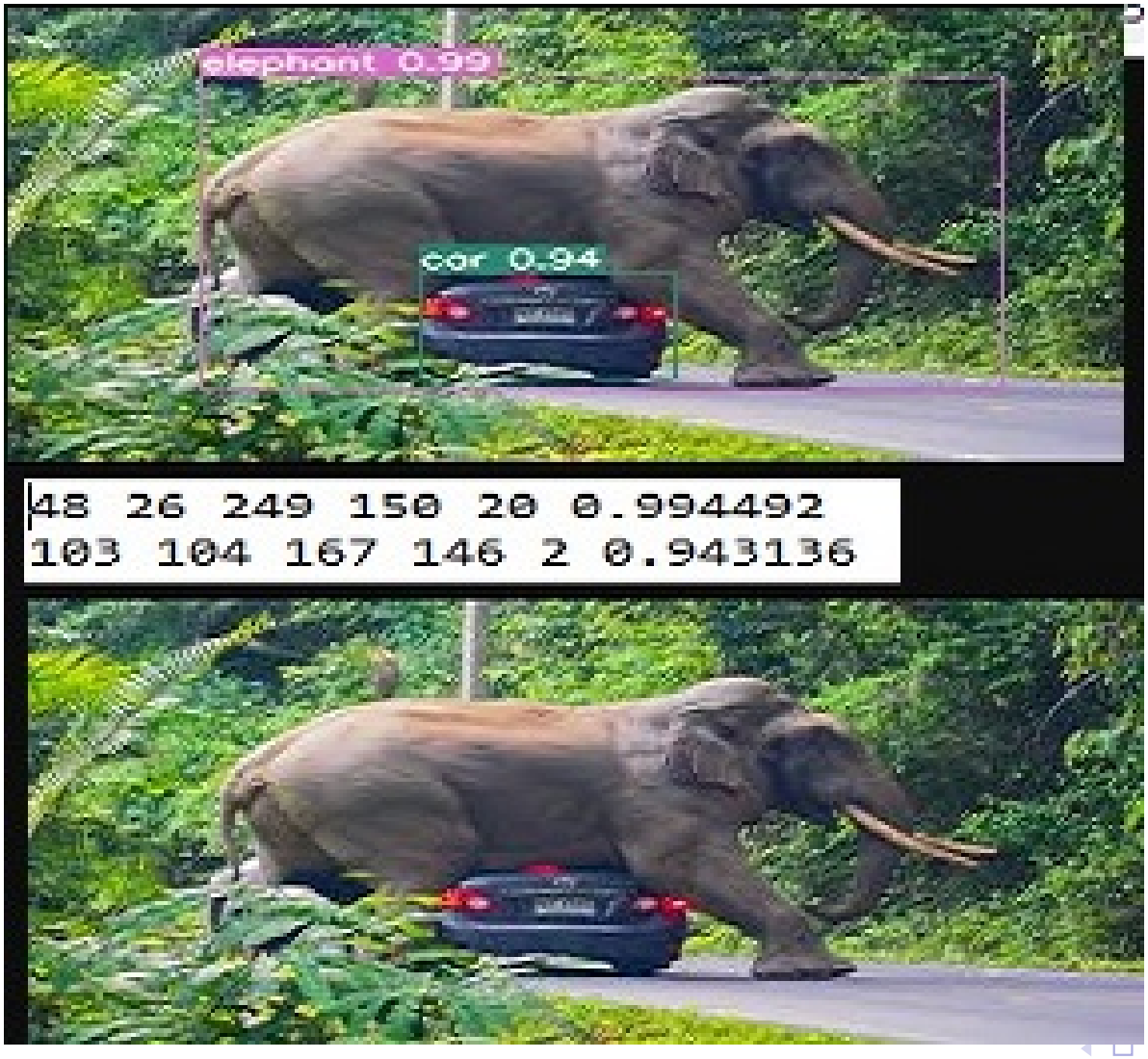
Random generation of image numbers, object names (A, B, C...), rectangle coordinates (X1,Y1)(X2,Y2), and depth.

Image number	Object name	X1	Y1	X2	Y2	Depth
48	H	300	12	24	192	2
5	C	198	44	401	41	5
19	O	418	324	441	48	7
14	C	107	388	454	220	3
85	O	316	117	237	97	7

Implementation of an Algorithm for object extraction.

YOLO+RSIDE and more:

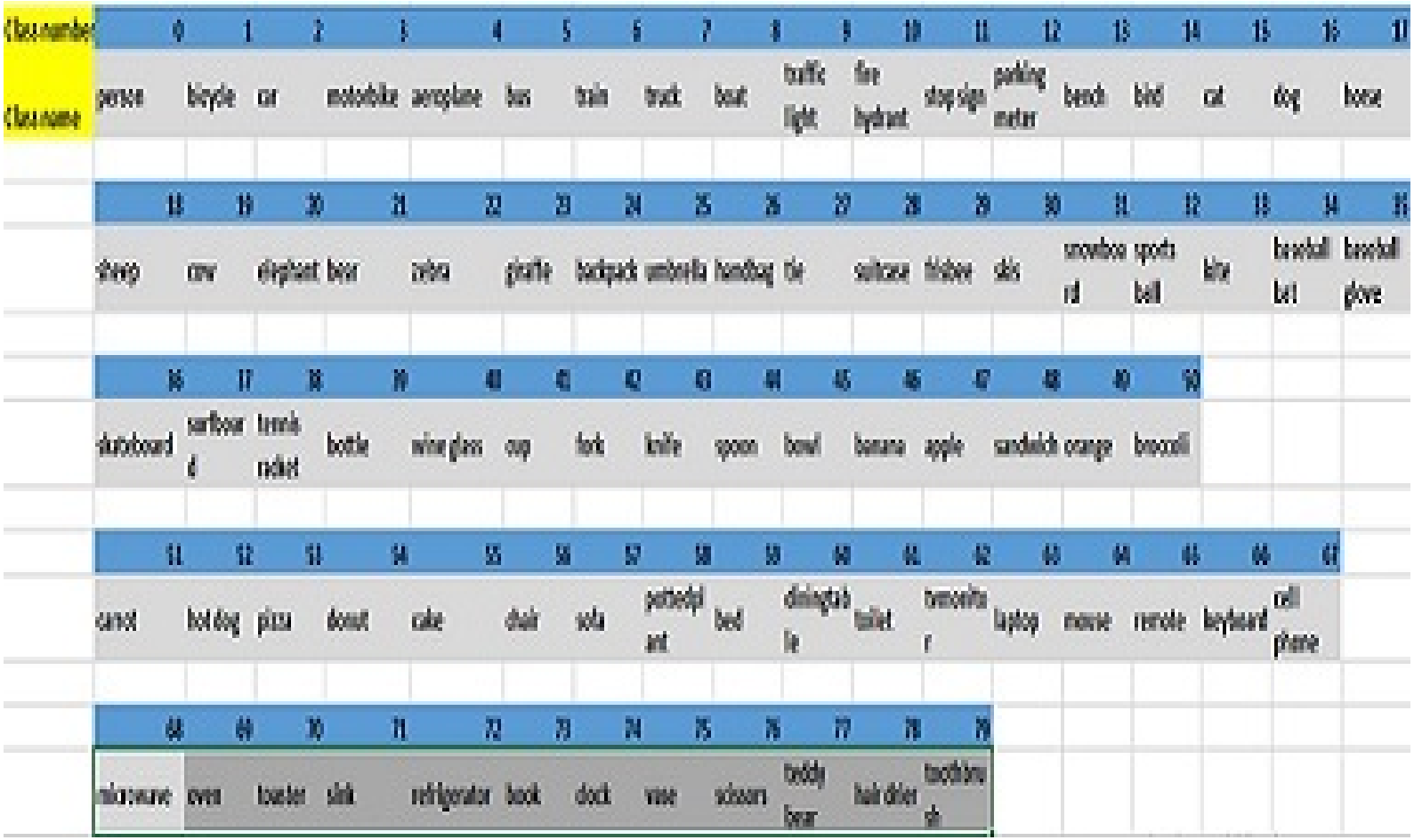
- The current output for YOLO:



Implementation of an Algorithm for object extraction.

YOLO objects and their corresponding class number:

- The current classes for YOLO:



class number	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
class name	person	bicycle	car	motorbike	aeroplane	bus	train	truck	boat	traffic light	fire hydrant	stop sign	parking meter	bench	bird	cat	dog	horse

	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
	sheep	cow	elephant	bear	zebra	giraffe	backpack	umbrella	handbag	tie	suitcase	frisbee	skis	snowboard	sports ball	kite	baseball bat	baseball glove

	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
	skateboard	surfboard	tennis racket	bottle	wine glass	cup	fork	knife	spoon	bowl	banana	apple	sandwich	orange	broccoli

	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67
	carrot	hot dog	pizza	donut	cake	chair	sofa	pottedplant	bed	diningtable	toilet	tvmonitor	laptop	mouse	remote	keyboard	cell phone

	68	69	70	71	72	73	74	75	76	77	78	79
	microwave	oven	toaster	sink	refrigerator	book	clock	vase	scissors	teddy bear	hair drier	toothbrush

Implementation of an Algorithm for object extraction.

Study of the results of YOLO+RSIDE :

- Following this, we try to prepare data for semantic web modelling. In this regard, we have prepared a blueprint to study the use of results of object detection.
- We try to identify the object position in the image by using the results of YOLO. However, the blue marked relation is restricted to RSIDE

image	relation	property value
	has on the left	object
	has on the right	object
	has on the top	object
	has on the buttom	object
	has in the center	object
	has in the back	object
	has in the front	object
	has isolated object	object
	nb of objects	value
	contains	object
	has background	object

Study of the results of YOLO contd...:

- This figure shows various object relations and these objects are not consolidated based on the image names

Image name	object name	has on the left	has on the right	has on the top	has on the bottom	has in the centre
download (1).jpg.txt	car	na	car	na	car	car
download (10).jpg.txt	cat	na	cat	na	cat	cat
download (10).jpg.txt	cup	na	na	cup	na	cup
download (11).jpg.txt	cat	na	cat	na	cat	cat
download (12).jpg.txt	cat	na	cat	na	cat	cat
download (12).jpg.txt	diningtable	na	diningtable	na	diningtable	diningtable
download (13).jpg.txt	cat	na	cat	na	na	cat
download (14).jpg.txt	elephant	na	elephant	na	elephant	elephant
download (14).jpg.txt	elephant	elephant	na	na	elephant	elephant
download (15).jpg.txt	elephant	na	elephant	na	elephant	elephant
download (15).jpg.txt	cow	na	cow	na	cow	cow
download (15).jpg.txt	elephant	na	elephant	na	elephant	elephant
download (15).jpg.txt	cow	cow	na	na	cow	cow
download (15).jpg.txt	cow	na	cow	na	cow	cow
download (15).jpg.txt	cow	na	cow	na	cow	cow

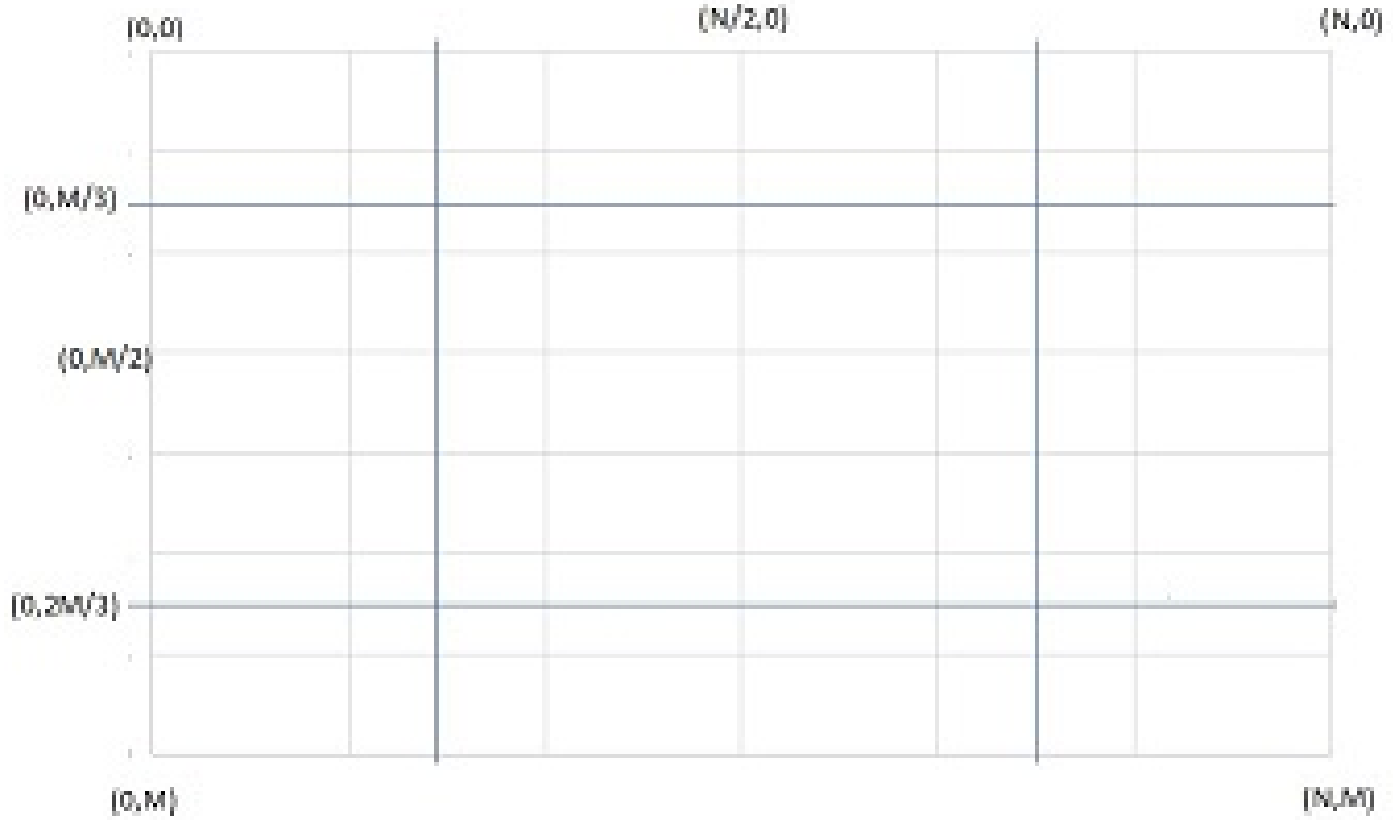
Study of the results of YOLO contd...:

- This figure shows a mathematical validation for choosing a particular object to the right, left,centre,top or bottom

Image name	Image Dimention	X1	Y1	X2	Y2	object name	X-centre	Y-centre
download (1).jpg.txt	(168, 299, 3)	26	22	277	144	car	151.5	83
download (10).jpg.txt	(168, 300, 3)	47	6	276	164	cat	161.5	85
download (10).jpg.txt	(168, 300, 3)	1	0	83	114	cup	42	57
download (11).jpg.txt	(168, 300, 3)	18	3	265	164	cat	141.5	83.5
download (12).jpg.txt	(224, 225, 3)	1	10	181	217	cat	91	113.5
download (12).jpg.txt	(224, 225, 3)	8	9	221	218	diningtable	114.5	113.5
download (13).jpg.txt	(159, 318, 3)	35	0	257	159	cat	146	79.5
download (14).jpg.txt	(194, 259, 3)	84	34	209	170	elephant	146.5	102
download (14).jpg.txt	(194, 259, 3)	4	36	88	162	elephant	46	99
download (15).jpg.txt	(168, 300, 3)	196	55	255	114	elephant	225.5	84.5
download (15).jpg.txt	(168, 300, 3)	77	58	155	115	cow	116	86.5
download (15).jpg.txt	(168, 300, 3)	254	50	299	116	elephant	276.5	83
download (15).jpg.txt	(168, 300, 3)	7	74	71	122	cow	39	98
download (15).jpg.txt	(168, 300, 3)	171	83	194	118	cow	182.5	100.5
download (15).jpg.txt	(168, 300, 3)	241	52	300	115	cow	270.5	83.5
download (15).jpg.txt	(168, 300, 3)	69	74	88	118	elephant	78.5	96
download (15).jpg.txt	(168, 300, 3)	79	53	162	114	elephant	120.5	83.5

Study of the results of YOLO contd...:

- This figure shows a mathematical detail for choosing a particular object to the right, left,centre,top or bottom



Study of the results of YOLO contd...:

- We define the following relations:
- If X-centre $<$N/3(N=X co-ordinate of the image), then object is on the left
- If X-centre $>$2N/3(N=X co-ordinate of the image), then object is on the right
- If Y-centre $<$M/3(M=Y co-ordinate of the image), then object is on the top
- If Y-centre $>$M/3(M=Y co-ordinate of the image), then object is on the bottom
- If N/3 $>$X-centre $>$2N/3 and M/3$>$Y-centre $>$2N/3, then object is at the centre

Study of the results of YOLO contd...:

- This figure shows various object relations and these objects are consolidated based on the image names. Here we would just see a glimpse of the isolated objects, number of objects corresponding to each image, and finally what are all the objects that the image contains.

image name	contains	isolated objects	no of
download (1).jpg.txt	car	car	1
download (10).jpg.txt	cat, cup	na	2
download (11).jpg.txt	cat	cat	1
download (12).jpg.txt	cat, diningtable	na	2
download (13).jpg.txt	cat	cat	1
download (14).jpg.txt	elephant, elephant	na	2
download (15).jpg.txt	elephant, cow, elephant, cow, cow, cow, elephant, elephant, elephant	na	9
download (16).jpg.txt	elephant, elephant	na	2
download (17).jpg.txt	elephant, car, elephant	na	3
download (18).jpg.txt	car, dog	na	2
download (19).jpg.txt	elephant, car	na	2
download (2).jpg.txt	car, car	na	2
download (3).jpg.txt	person, car	na	2
download (4).jpg.txt	car	car	1
download (5).jpg.txt	car	car	1
download (6).jpg.txt	bus	bus	1
download (7).jpg.txt	bus, traffic light	na	2
download (8).jpg.txt	bus	bus	1

Study of the results of YOLO+RSIDE contd...:

- We try to identify an object position with respect to another object in a particular image by using the results of YOLO. However, the blue marked relation is restricted to RSIDE. Moreover, we would also try to establish relationship between more than 2 objects(3 objects)

Object	relation	property value
	left of	object
	right of	object
	top of	object
	buttom of	object
	further than	object
	closer than	object
	close to	object
	far from	object
	overlaping with	object
	behind	object
	in front of	object
	greater than	object
	smaler than	object
	related-to	object
	has color	color
	%of image	value
relation between 3 objects		
	inbetween in plan	
	inbetween in depth	