OBJECT DETECTION USING YOLO

DR.S. Venkatesh M.E, Ph. D[#1] Mohammed Yousuf.N[#2], Nitish Kumar.R[*3], Renganathan.J[#4]

1st Assistant Professor, Department of Computer Science And Engineering Jeppiaar Engineering College venkyjep2019@gmail.com	2nd Student of Department of Computer Science and Engineering, Jeppiaar Engineering College yousuf4christ@gmail.com	3rd Student of Department of Computer Science and Engineering Jeppiaar Engineering College nithishkumarrnk0@gmail.com

4th Student of Department
of Computer Science
and Engineering,
Jeppiaar Engineering College
rengarenga180@gmail.com

ABSTRACT --- OBJECT DETECTION IS A VAST, VIBRANT AND COMPLEX AREA OF COMPUTER VISION. IF THERE IS A SINGLE OBJECT TO BE DETECTED IN AN IMAGE, IT IS KNOWN AS IMAGE LOCALIZATION AND IF THERE ARE MULTIPLE OBJECTS IN AN IMAGE, THEN IT IS OBJECT DETECTION. THIS DETECTS THE SEMANTIC OBJECTS OF A CLASS IN DIGITAL IMAGES AND VIDEOS. THE APPLICATIONS OF OBJECT DETECTION INCLUDE TRACKING OBJECTS, PEDESTRIAN DETECTION, PEOPLE COUNTING, SELF-DRIVING CARS, FACE DETECTION, BALL TRACKING IN SPORTS AND MANY MORE. CONVOLUTION NEURAL NETWORKS IS A REPRESENTATIVE TOOL OF MACHINE LEARNING TO DETECT OBJECTS USING OPEN CV (OPEN SOURCE COMPUTER VISION), WHICH IS A LIBRARY OF PROGRAMMING FUNCTIONS MAINLY AIMED AT REAL TIME COMPUTER VISION.

INTRODUCTION

TIGHT SECURITY IN PUBLIC PLACES SUCH AS IN AIRPORT TERMINALS, TRAIN STATIONS, AND COMMERCIAL ESTABLISHMENTS IS IMPORTANT NOWADAYS DUE TO THE INCREASING TERRORIST ACTIVITIES LIKE BOMBING. THESE WILL PREVENT SEVERAL CASUALTIES THAT CAN POSSIBLY HAPPEN ANYTIME. WHEN A CHILD BEGINS TO WALK THE FIRST THING THAT IT TRIES TO DO IS UNDERSTAND THE SURROUNDING AROUND ITSELF, SIMILAR IS THE CASE WITH OBJECT DETECTION. THERE ARE MULTIPLE SENSORS THAT AID FOR THE PURPOSE OF UNDERSTANDING THE SURROUNDINGS BUT THE MOST EFFICIENT ONES ARE THE CAMERA SENSORS. ON A COMPARATIVE OF ALL THE SENSORS THAT EXIST, CAMERA SENSORS TEND TO BE THE BEST PERFORMERS AT AN EFFICIENT COMPREHENSION OF THE SURROUNDING. MULTIPLE COMPANIES HAVE BEEN SUCCESSFUL IN DEVELOPING OBJECT DETECTION ALGORITHMS SUCH AS NVIDIA[2], VINOTION, ETC. WHICH ARE BEING USED BY THE AUTONOMOUS DRIVING GIANTS LIKE TESLA BUT INFORMATION REGARDING THESE AREN'T WIDELY AVAILABLE.

LITERATURE SURVEY

"OBJECT DETECTION ALGORITHMS FOR VIDEOSURVEILLANCE APPLICATIONS" BY APOORVA RAGHUNANDAN, MOHANA, PAKALA RAGHAV AND H. V. RAVISH ARADHYA OBJECT DETECTION MAINLY DEALS WITH IDENTIFICATION OF REAL-WORLD OBJECTS SUCH AS PEOPLE, ANIMALS, AND OBJECTS OF SUSPENSE OR THREATENING OBJECTS. OBJECT DETECTION ALGORITHMS USE A WIDE RANGE OF IMAGE PROCESSING APPLICATIONS FOR EXTRACTING THE OBJECT'S DESIRED PORTION. IT IS COMMONLY USED IN APPLICATIONS SUCH AS IMAGE RETRIEVAL, SECURITY, MEDICALFIELD, AND DEFENSE.

"AN OBJECT DETECTION ALGORITHM WITH DISPARITY VALUES" BY CHUNG-CHENG CHIU ANDWEN-CHIEN LO PROPOSES AN OBJECT DETECTION ALGORITHM WITH DISPARITY VALUES TO SOLVE THE PROBLEM. THE PROPOSED ALGORITHM USES THE CONNECTED COMPONENT ALGORITHM AND DISPARITY VALUES TO CONNECT PIXELS WITH SIMILAR DISPARITIES, AND THE RECURSIVEOBJECT MERGE USES TO MERGE THE SIMILAR NEIGHBORING CONTOURS FOR LOCATING AND BOUNDING THE CONTOURS OF AN OBJECT. EXPERIMENTAL RESULTS SHOW THAT THE PROPOSED ALGORITHM CAN RESOLVE THE PROBLEM OF OBJECT DETECTION IN OVERLAPPING OBJECTS, AND THE DETECTION RESULTS CAN BE USED FOR SUBSEQUENT OBJECT RECOGNITION APPLICATIONS.

"OBJECT DETECTION FOR AUTONOMOUS DRIVING USING YOLO [YOU ONLY LOOK ONCE] ALGORITHM"BY ABHISHEK SARDA, DR. SHUBHRA DIXIT AND DR. ANUPAMA BHAN, ONE OF THE PRIMARY PREREQUISITES AND A HUGE PART OF AUTONOMOUS DRIVING IS DEPENDENT ON OBJECT DETECTION THROUGH COMPUTER VISION, THIS PAPER AIMS AT AIDING TOWARDS THE FIELD OF AUTONOMOUS DRIVING BY HELPING DETECT OBJECTS WITH THE USE OF DEEP LEARNING ALGORITHMS. RESEARCH WORK USED STATE-OF-THE-ART ALGORITHM YOLO (YOU ONLY LOOK ONCE) TO DETECT DIFFERENT OBJECTS THAT APPEAR ON THE ROAD AND CLASSIFIED INTO THE CATEGORY THAT THEY BELONG TO WITH THE HELP OF BOUNDING BOXES. THE WEIGHTS OF THE YOLO V4 IS UTILIZED TO CUSTOM TRAIN OUR MODEL TO DETECT THE OBJECTS AND THE DATA WILL BE COLLECTED FROM THE OPEN IMAGES DATASET USING ITS YOLO TOOLKIT.

"PREDICTION OF TRAFFIC DENSITY USING YOLO OBJECT DETECTION AND IMPLEMENTED IN RASPBERRY PI 3B +AND INTEL NCS 2" by ROSA ANDRIE ASMARA, BIMO SYAHPUTRO, ANIK NUR HANDAYANI AND DODIT SUPRIYANTO. AIMS TO MEASURE HOW THE ERROR RATE OBTAINED FROM THE RESULTS OF FORECASTING. VEHICLE IDENTIFICATION AS THE OBJECT CAUSING THE TRAFFIC JAM USES THE CONVOLUTIONAL NEURAL NETWORK (CNN) OBJECT DETECTION METHOD, ESPECIALLY THE YOU ONLY LOOK ONCE (YOLO) METHOD. THE SYSTEM DESIGN INVOLVES A RASPBERRY PI 3 DEVICE, AN INTEL NCS 2 DEVICE, AND A WEBSITE APPLICATION. BASED ON THE TEST RESULTS OBTAINED, ACCURATE FORECASTING ACCURACY.

IN" YOLO-BASED THREAT OBJECT DETECTION IN X-RAY IMAGES" BY REAGAN L. GALVEZ, ELMER P. DADIOS, ARGEL A. BANDALA, AND RYAN RHAY P. VICERRA THEY HAVE PROPOSED THE SYSTEM THAT MANUAL DETECTION OF THREAT OBJECTS IN AN X-RAY MACHINE IS A TEDIOUS TASK FOR THE BAGGAGE INSPECTORS IN AIRPORTS, TRAIN STATIONS, AND ESTABLISHMENTS. OBJECTS INSIDE THE BAGGAGE SEEN BY THE X-RAY MACHINE ARE COMMONLY OCCLUDED AND

DIFFICULT TO RECOGNIZE WHEN ROTATED. BECAUSE OF THIS, THERE IS A HIGH CHANCE OF MISSED DETECTION, PARTICULARLY DURING RUSH HOUR. AS A SOLUTION, THIS PAPER PRESENTS A YOU ONLY LOOK ONCE (YOLO)-BASED OBJECT DETECTOR FOR THE AUTOMATED DETECTION OF THREAT OBJECTS IN AN X-RAY IMAGE.

PROPOSED SYSTEM

USING THRESHOLD, WE TRY TO SEE THE MINIMUM THRESHOLD VALUE FOR WHICH WE CAN ELIMINATE THIS PROBLEM. (I.E.) SAME OBJECT HAVING TWO LABELS. AFTER ESTABLISHING THE DIFFERENT THRESHOLD MODELS, WE COMPARED THE EFFECTS OF DIFFERENT THRESHOLD MODELS ON OBJECT DETECTION. WE USED THE YOLO NETWORK TO TRAIN A ROBUST MODEL TO IMPROVE THE AVERAGE PRECISION (AP) OF DETECTION IN REAL SCENES.

PROPOSED ALOGORITHM

THE GENERAL FLOW OF THE PROGRAM WILL BE AS FOLLOW

1. **OBJECT DETECTION:**
   a. DOWNLOAD THE YOLO MODEL WITH ITS WEIGHTS, LABELS, AND CONFIGURATION FILES.
   b. SET THE PATH FOR WEIGHTS, LABELS AND CONFIGURATION FILES.
   c. LOAD THE YOLO MODEL
   d. LOAD THE IMAGE
   e. PRE-PROCESS AND RESIZE THE IMAGE AND CONVERT TO NUMPY BLOB
   f. PASS THE IMAGE TO YOLO
   g. GET THE LAYERS CO-ORDINATES AND THEIR CORRESPONDING LABELS AND CONFIDENCE LEVELS.
   h. FILTER THE LAYER WITH REQUIRED CONFIDENCE LEVELS.
   i. DRAW THE BOX AND LABEL IT

2. **GUI FRONT END:**
   a. CREATE A BOX TO CONTAIN THREE BUTTONS
   b. CREATE TWO BUTTONS TO CALL THE TWO FUNCTIONS(DEFINED ABOVE)
   c. CREATE A BUTTON THE EXIT

METHODOLOGY

THE WHOLE PROGRAM IS DIVIDED INTO THREE PARTS.
1. FUNCTION TO PROCESS A SINGLE IMAGE
2. FUNCTION TO PROCESS MULTIPLE IMAGES

## FUNCTION TO PROCESS SINGLE IMAGE:

THE DEFINITION OF THE FUNCTION STARTS WITH DEFINING THE VALUES FOR OUR CONFIDENCE LEVELS AND THRESHOLD LEVELS. WE HAVE SET THAT AT 50% (0.5). THEN THE PATHS FOR CONFIGURATION FILE, YOLO WEIGHT FILE AND YOLO LABEL FILE ARE PROPERLY CONFIGURED. TO DRAW THE BOXES IN IMAGES IN DIFFERENT COLOURS FOR DIFFERENT OBJECTS, RANDOM COLOUR GENERATION IS SET. FINALLY YOLO MODEL IS LOADED. USING THE TKINTER, WINDOWS BASED FILE OPENING SYSTEM IS USED TO ASK THE USER TO OPEN A SINGLE IMAGE. THE IMAGE IS READ, STORED AND RESIZED TO 416*416 PIXEL WITH RED AND BLUE LAYER SWAPPED. THIS IS CONVERTED INTO A BLOB BY NUMPY. THIS BLOB IS PASSED TO YOLO MODEL AND THE LAYER DETAILS (CO-ORDINATES, CONFIDENCE LEVEL, AND CLASS-IDS) ARE OBTAINED FROM IT.

FUNCTION TO PROCESS MULTIPLE IMAGE:

THE DEFINITION OF THE FUNCTION STARTS WITH DEFINING THE VALUES FOR OUR CONFIDENCE LEVELS AND THRESHOLD LEVELS. WE HAVE SET THAT AT 50% (0.5). THEN THE PATHS FOR CONFIGURATION FILE, YOLO WEIGHT FILE AND YOLO LABEL FILE ARE PROPERLY CONFIGURED. TO DRAW THE

BOXES IN IMAGES IN DIFFERENT COLOURS FOR DIFFERENT OBJECTS, RANDOM COLOUR GENERATION IS SET. FINALLY YOLO MODEL IS LOADED. USING THE TKINTER, WINDOWS BASED FILE OPENING SYSTEM IS USED TO ASK THE USER TO OPEN A SINGLE IMAGE. THE IMAGE IS READ, STORED AND RESIZED TO 416*416 PIXEL WITH RED AND BLUE LAYER SWAPPED. THIS IS CONVERTED INTO A BLOB BY NUMPY. THIS BLOB IS PASSED TO YOLO MODEL AND THE LAYER DETAILS (CO-ORDINATES, CONFIDENCE LEVEL, AND CLASS-IDS) ARE OBTAINED FROM IT. USING OUR PREDEFINED CUT-OFF FOR CONFIDENCE LEVEL, THE CO-ORDINATES AND CLASS-IDS WITH LESS CONFIDENCE LEVELS ARE REJECTED. FOR THE REMAINING DETAILS, A LOOP IS STARTED. INSIDE THE LOOP EACH CO-ORDINATES VALUES ARE READ, CONVERTED TO BOX CO-ORDINATE SYSTEM AND THEN DRAWN ON THE GIVEN IMAGE WITH A RANDOM COLOUR (WHICH WE HAVE PRE-DEFINED AT THE START OF THE FUNCTION) AND ALSO LABELLED WITH THE CLASS-IDS. THIS PROCESS IS REPEATED FOR ALL THE SELECTED CLASS-IDS.ONCE THE PROCESS IS COMPLETED THE IMAGE IS STORED AT THE SAME PATH AS ORIGINAL, WITH "_YOLO4" ADDED TO ITS NAME.
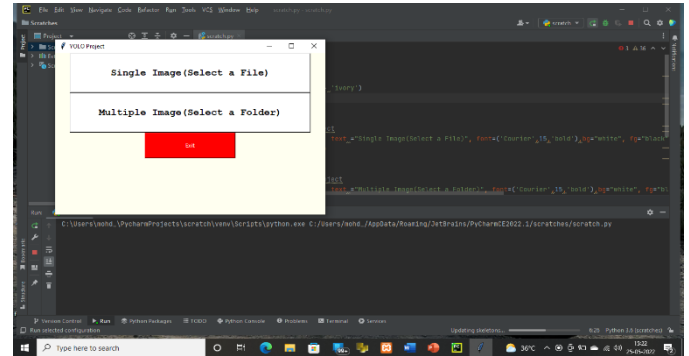
Block diagram for training process:
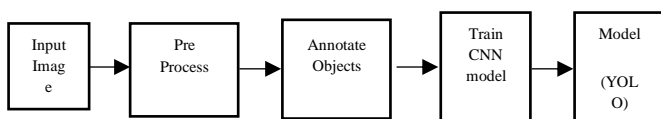
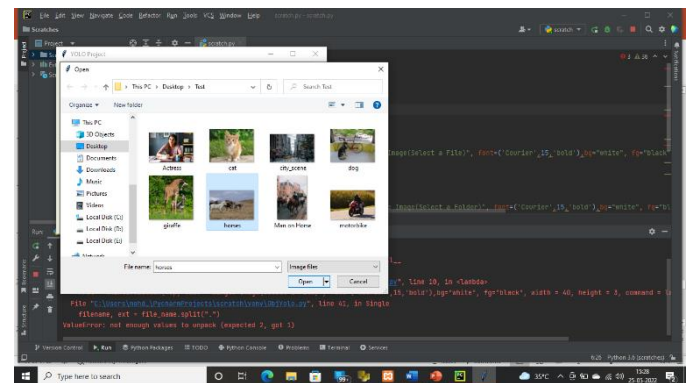**Block diagram for testing process:**
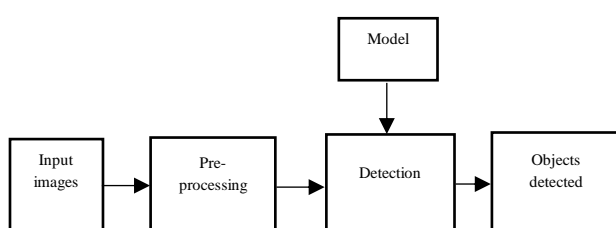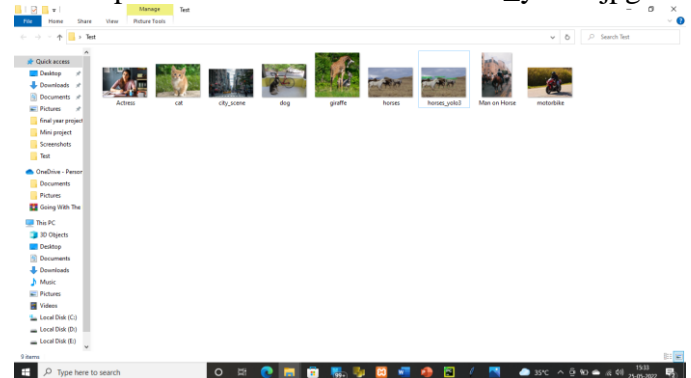

OUTPUTED SCREENSHOTS
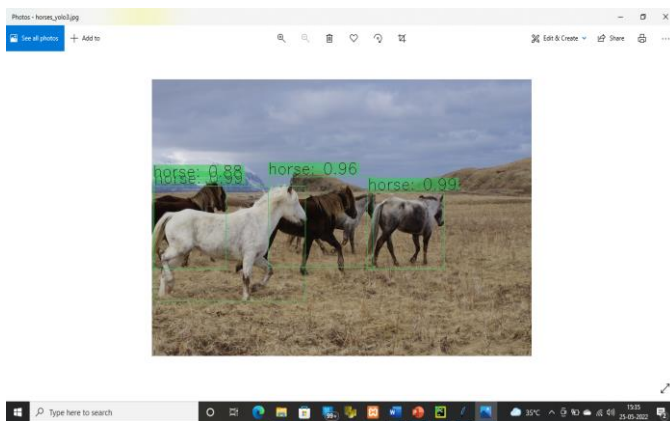
FRONT END GUI Interface:
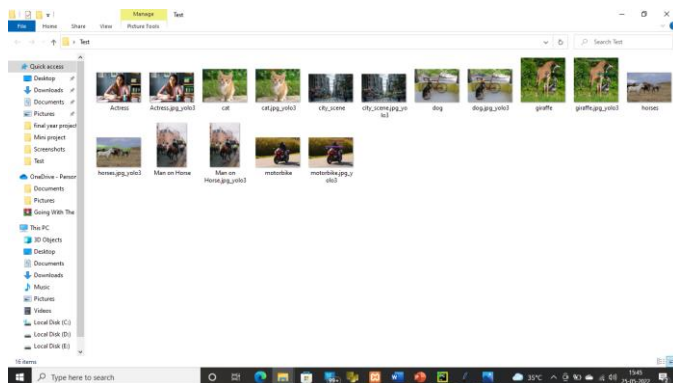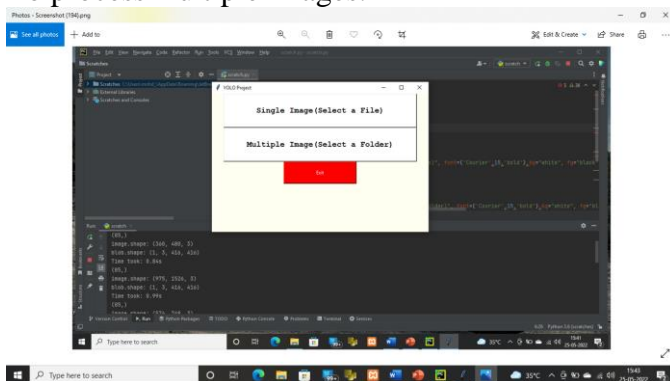

TO PROCESS A SINGLE IMAGE:

The output file will be shown as name_yolo3.jpg

To process multiple images:

REFERENCES

AGARWAL, S., AWAN, A., AND ROTH, D. (2004). LEARNING TO DETECT OBJECTS IN IMAGES VIAASPARSE,PART-BASED REPRESENTATION. IEEE TRANS. PATTERN ANAL. MACH. INTELL. 26,1475–1490.DOI:10.1109/TPAMI.2004.108

ALEXE, B., DESELAERS, T., AND FERRARI, V. (2010). "WHAT IS AN OBJECT?," IN COMPUTERVISIONANDPATTERNRECOGNITION (CVPR), 2010 IEEE CONFERENCE ON (SAN FRANCISCO,CA: IEEE), 73–80.DOI:10.1109/CVPR.2010.5540226

ALOIMONOS, J., WEISS, I., AND BANDYOPADHYAY, A. (1988). ACTIVE VISION. INT. J.COMPUT. VIS.1,333–356. DOI:10.1007/BF00133571

ANDREOPOULOS, A., AND TSOTSOS, J. K. (2013). 50 YEARS OF OBJECT RECOGNITION: DIREC-TIONSFORWARD.COMPUT. VIS. IMAGE UNDERST. 117, 827–891. DOI:10.1016/J.CVIU.2013.04.005

CONCLUSIONS

OBJECT DETECTION FINDS SCOPE IN VARIOUS AREAS SUCH AS DEFENSE AND BORDER SECURITY, MEDICAL IMAGE PROCESSING, VIDEO SURVEILLANCE, ASTRONOMY AND OTHER SECURITY RELATED APPLICATIONS.THE PROPOSED METHOD CAN RESOLVE THE PROBLEM OF OBJECT DETECTION IN OVERLAPPING OBJECTS, AND IT CAN PERFORM BETTER RESULTS FOR SUBSEQUENT OBJECT RECOGNITION APPLICATIONS. THEY FURTHER CAN BE IMPROVED IN ORDER TO DETECT OBJECT WITH MUCH GREATER EFFICIENCY IN ACCURACY