**AI Project**

|  |  |
| --- | --- |
| **ID** | **الاسم** |
| **201900439** | **عبد الرحمن محمود محمود** |
| **201900419** | **عبد الرحمن سيد فتحي** |
| **201900384** | **ضحي علي مصطفي عبد الرحمن** |
| **201900600** | **ماريان نادر موريس جرجس** |
| **201900149** | **أسماء محمد عابدين توفيق** |

***Object Detection using Artificial Neural Networks***

**(1) Project idea in details**

**Object detection is a computer vision technique that allows us to identify and locate objects in an image or video. With this kind of identification and localization, object detection can be used to count objects in a scene and determine and track their precise locations, all while accurately labeling them.**

**Object detection allows us to at once classify the types of things found while also locating instances of them within the image**.

صورة تحتوي على نص, مبنى, طريق, خارجي

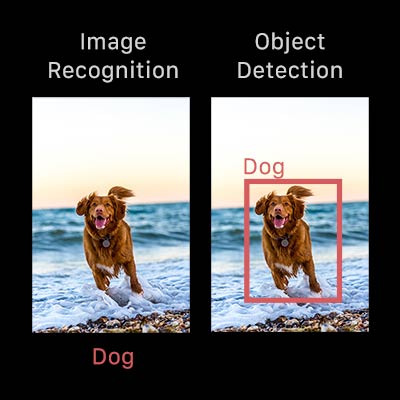
تم إنشاء الوصف تلقائياً

**Object detection is commonly confused with image recognition, so before we proceed, it’s important that we clarify the distinctions between them.**

**Image recognition assigns a label to an image. A picture of a dog receives the label “dog”. A picture of two dogs, still receives the label “dog”.**

**Object detection, on the other hand, draws a box around each dog and labels the** **box “dog”. The model predicts where each object is and what label should be applied. In that way, object detection provides more information about an image than recognition.**

**Here’s an example of how this distinction looks in practice:**



**Object detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos. Object detection is a computer vision technique for locating instances of objects in images or videos.**

**Object detection algorithms typically leverage machine learning or deep learning to produce meaningful results. When humans look at images or video, we can recognize and locate objects of interest within a matter of moments. The goal of object detection is to replicate this intelligence using a computer.**

**Object recognition is an extremely complex problem that will, no matter what the approach, at least for now be bound by its high computational costs. One benefit of trying to tackle object recognition from the AI point of view is that such an approach, be it at times less exact, can often lead to a significant reduction in computational costs. Furthermore, since object recognition is a task we humans are particularly good at, it can certainly be beneficial to look at object recognition from a cognitive point of view.** **In our opinion the most interesting question in AI has always been how to generalize.**

**Humans have the great ability to generalize from one situation to another, even if nothing is exactly the same. How do we know when two situations require the same approach, and when they don’t? Where lies our ability to see whether things are similar or not?**

**(2) Main functionalities**

**Broadly speaking, object detection can be broken down into Machine Learning-based approaches and Deep Learning-based approaches.**

**In more traditional ML-based approaches, computer vision techniques are used to look at various features of an image, such as the color histogram or edges, to identify groups of pixels that may belong to an object. These features are then fed into a regression model that predicts the location of the object along with its label.**

**On the other hand, Deep Learning-based approaches employ convolutional neural networks (CNNs) to perform end-to-end, unsupervised object detection, in which features don’t need to be defined and extracted separately.**

**Neural networks**

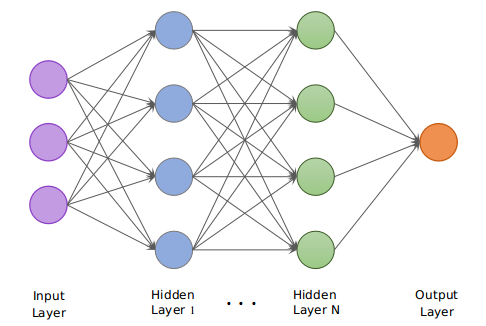
**NN is algorithms are inspired by the human brain to performs a particular task or functions. NN perform computations through a process by learning. The neural network is a set of connected input/output units in which each connection has a weight associated with it. In the learning phase, the network learns by adjusting the weights to predict the correct class label of the given inputs.**

صورة تحتوي على ساعة حائط

تم إنشاء الوصف تلقائياً

**Here, x1,x2....xn are input variables. w1,w2....wn are weights of respective inputs. b is the bias, which is summed with the weighted inputs to form the net inputs. Bias and weights are both adjustable parameters of the neuron. Parameters are adjusted using some learning rules. The output of a neuron can range from -inf to +inf.**

**Block Diagram of the Algorithm**



**A neural network consists of an input layer, output layer, and hidden layer, which has many units that transform the input into something meaningful that can be used by the output layer; hidden layers are required to separate the data non-linearly.**

**Artificial neural networks are similar to the human brain.**

**An artificial neural network (ANN) is a popular tool used in machine learning. ANNs are used in facial recognition, image classification, object detection, weather forecasting, financial forecasting, etc. To perform such tasks, neural networks are trained by giving them thousands of input data so that a model can differentiate between similar things. To process such wide-ranging data, GPU must be used.**

**By implementing object recognition using neural networks we hope to create a system that is able to generalize well among objects, that is adaptable so that it can easily learn to recognize new objects and that is limited in its computational cost so that it can easily be incorporated into existing systems. We have chosen neural networks for this task because of their favorable properties that make them an excellent choice for object recognition.**

**The most important of these properties are:**

**• Generalization**

**Small distortions can be handled easily, a necessity for object recognition. Accounting for small changes under different conditions** **(lighting, rotation, etc.) can to a certain extent be left up to training the network under as many conditions as possible.**

**• Expandability**

**Another great benefit of a neural network is that it can easily be expanded. In order for it to learn new objects there is no need to start all over and redefine distance measures and distributions.** **Learning a different set of objects will require hardly any change to the structure of the program.**

• **Representing multiple samples**

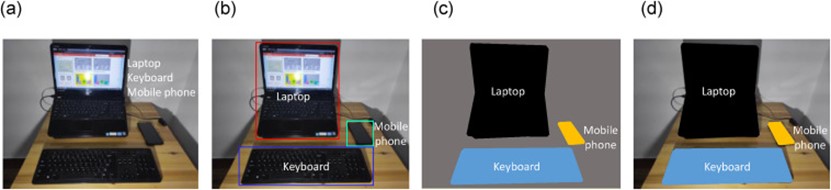
**A class of objects can easily be represented by multiple samples under multiple conditions. Because a neural network incorporates in its structure what it learns, the recognition of an object becomes a single step. In this way there is no need for multiple comparisons as in many conventional systems. The network determines in one single step to what class the object belongs.**

**• Memory Saving**

**An advantage stemming from the previously mentioned characteristic is that there is no need to store all the standard images to be used for comparison. Once a network is trained properly it contains the necessary information and the image data becomes expendable and can be removed from memory.**

**(3) Similar applications in the market**

**1-Image recognition**

****

**2- Text classification, information extraction, and speech and character recognition.**

**صورة تحتوي على نص, سبورة بيضاء

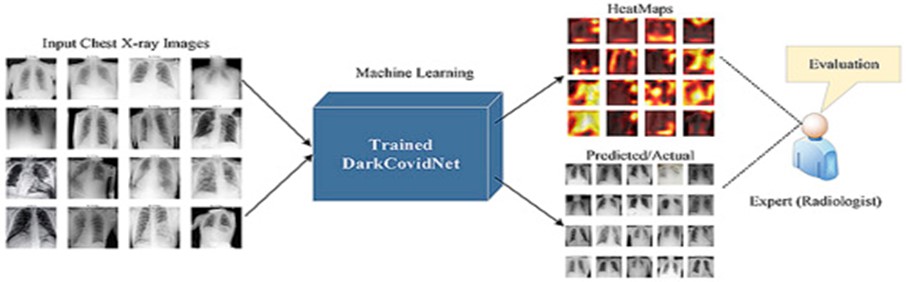
تم إنشاء الوصف تلقائياً**

**3-** **Face detection**

صورة تحتوي على نص, شخص, ذكر

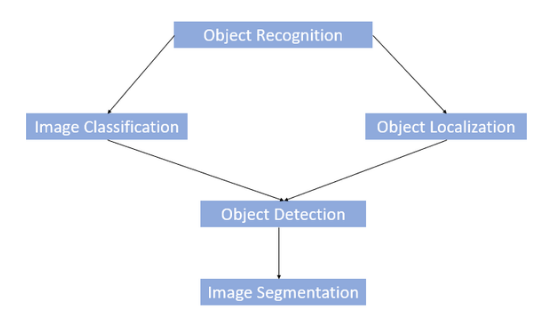
تم إنشاء الوصف تلقائياً

**4-** **Automated detection of COVID-19 cases using deep neural networks with X-ray images**



**(4) An initial literature review of Academic publications (papers) relevant to the idea (Object detection)**

**Computer vision is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs, and take actions or make recommendations based on that information. And there are 4 important computer vision tasks that are classification, localization, object detection and instance segmentation**

****

**Object detection is one of the modern technologies that all the world focus on, and it is the future of the world, as it is recently applications is focused on artificial intelligence and computer vision and deep learning, as the object detection is a computer vision technique that allows us to locate and identify objects in an image or video. and can be approached as either a classification problem or a regression problem. As a classification problem, the image is divided into small patches, each of which will be run through a classifier to determine whether there are objects in the patch. Then the bounding boxes will be assigned to locate around patches that are classified with a high probability of present of an object. In the regression approach, the whole image will be run through a convolutional neural network to directly generate one or more bounding boxes for objects in the images.**

**صورة تحتوي على نص, الثدييات, كلب

تم إنشاء الوصف تلقائياً**

**Body and summarize:**

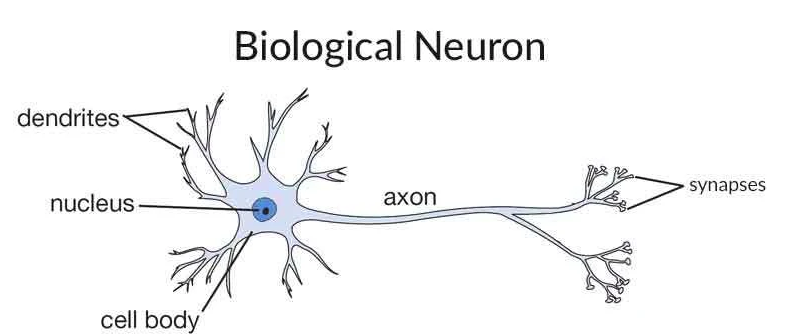
**Using artificial intelligence technique, object detection can be used to count objects in a scene and determine and track their precise locations, all while accurately labeling them,** **Specifically, object detection draws bounding boxes around these detected objects, also refers to the detection or classification of objects in image or video data.**

**And we will use Artificial Neural Networks for Object Detection, as ANN is consisting of several different layers such as the input layer, at least one hidden layer, and an output layer. They are best used in object detection for recognizing patterns such as edges (vertical/horizontal), shapes, colors, and textures, as it is the most current deep learning approaches are based on a network architecture known. These networks use the backpropagation method for the actual learning process, which is also referred to as training.**

**Recently, object recognition methods using deep learning or artificial neural networks (ANNs). Artificial Neural networks (ANN) or neural networks are computational algorithms. It intended to simulate the behavior of biological systems composed of “neurons”. ANNs are computational models inspired by an animal’s central nervous systems. It is capable of machine learning as well as pattern recognition. These presented as systems of interconnected “neurons” which can compute values from inputs.**

**A neural network is an oriented graph. It consists of nodes which in the biological analogy represent neurons, connected by arcs. It corresponds to dendrites and synapses. Each arc associated with a weight while at each node. Apply the values received as input by the node and define Activation function along the incoming arcs, adjusted by the weights of the arcs. And it is a machine learning algorithm based on the model of a human neuron.**

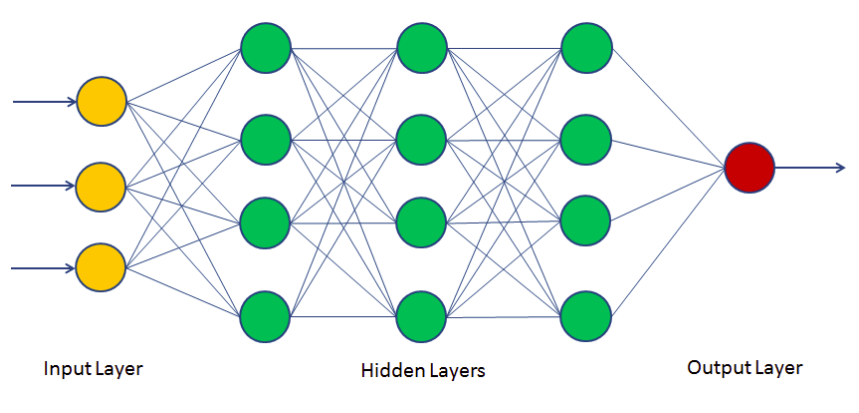
**The human brain consists of millions of neurons. It sends and process signals in the form of electrical and chemical signals. These neurons are connected with a special structure known as synapses. Synapses allow neurons to pass signals. From large numbers of simulated neurons neural networks forms.**

****

**An Artificial Neural Network is an information processing technique. It works like the way human brain processes information. ANN includes a large number of connected processing units that work together to process information. They also generate meaningful results from it, we can apply Neural network not only for classification. It can also apply for regression of continuous target attributes. As ANN is a great application in data mining used in sectors. For example pattern recognition. It can be also used for data classification in a large amount of data after careful training.**

**ANNs have a strong representational power that enables them to learn complex hierarchical representation of images, increasingly representing abstract concepts. They are powerful at generalizing to never seen data. This is especially the case when it comes to recognizing a multitude of different objects whose appearance also varies greatly. ANNs are also more robust against variations, e.g., masking, fading of colors, damage or soiling. As early as the end of the 1990s, ANNs were used to solve simple tasks such as handwriting recognition for object recognition primarily in that the objects or object classes to be recognized are not described by a set of manually predefined features, but the feature representation of an object is learned during training from the neural network.**

**Artificial Neural network is typically organized in 3 layers, Layers are being made up of many interconnected ‘nodes’ which contain an ‘activation function’. A neural network may contain the following 3 layers, input and hidden and output layers, as we are going to discuss them:**

****

**The Input layer: The purpose of the input layer is to receive as input the values of the explanatory attributes for each observation. Usually, the number of input nodes in an input layer is equal to the number of explanatory variables. ‘Input layer’ presents the patterns to the network, which communicates to one or more ‘hidden layers’ The nodes of the input layer are passive, meaning they do not change the data. They receive a single value on their input and duplicate the value to their many outputs. From the input layer, it duplicates each value and sent to all the hidden nodes.**

**The Hidden layer: The Hidden layers apply given transformations to the input values inside the network. In this, incoming arcs that go from other hidden nodes or from input nodes connected to each node. It connects with outgoing arcs to output nodes or to other hidden nodes. In hidden layer, the actual processing is done via a system of weighted ‘connections’. There may be one or more hidden layers. The values entering a hidden node multiplied by weights, a set of predetermined numbers stored in the program. The weighted inputs are then added to produce a single number.**

**The Output layer: The hidden layers then link to an ‘output layer‘. Output layer receives connections from hidden layers or from input layer. It returns an output value that corresponds to the prediction of the response variable. In classification problems, there is usually only one output node. The active nodes of the output layer combine and change the data to produce the output values.**

**The ability of the neural network to provide useful data manipulation lies in the proper selection of the weights. This is different from conventional information processing. The structure of a neural network also referred to as its ‘architecture’ or ‘topology’. It consists of the number of layers, Elementary units. The choice of the structure determines the results which are going to obtain. It is the most critical part of the implementation of a neural network.**

**The simplest structure is the one in which units distributes in two layers: An input layer and an output layer. Each unit in the input layer has a single input and a single output which is equal to the input. The output unit has all the units of the input layer connected to its input, with a combination function and a transfer function. There may be more than 1 output unit. In this case, resulting model is a linear or logistic regression. This is depending on whether transfer function is linear or logistic. The weights of the network are regression coefficients. By adding 1 or more hidden layers between the input and output layers and units in this layer the predictive power of neural network increases. But a number of hidden layers should be as small as possible. This ensures that neural network does not store all information from learning set but can generalize it to avoid overfitting. Overfitting can occur. It occurs when weights make the system learn details of learning set instead of discovering structures. This happens when size of learning set is too small in relation to the complexity of the model. A hidden layer is present or not, the output layer of the network can sometimes have many units, when there are many classes to predict.**

**And there a lot of libraries to design and create object detection program code like OpenCV, TensorFlow, OCR,** **YOLO and so on, and we are going to use OpenCV as it is s the huge open-source library for computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in our today’s systems. As it is a great tool for image processing and performing computer vision tasks, one can process images and videos to identify objects, faces, or even the handwriting of a human.**

**And it is programmed with python programming language. As also OpenCV provides many machine learning algorithms bundled into it. Some of the algorithms include Bayes Classifier, K-Nearest Neighbors, Support Vector Machines, Decision Trees, Neural Networks.**

صورة تحتوي على نص, كلب, مختلف, حزمة

تم إنشاء الوصف تلقائياً

**(5) the Dataset employed (preferably a publicly available dataset)**

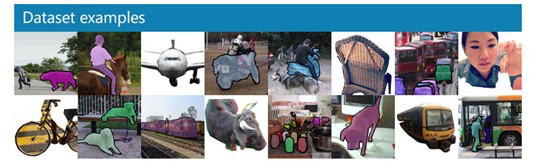
**The Dataset that we use in our Ai project (Object Detection) is COCO dataset**

**COCO is a large-scale object detection, segmentation, and captioning dataset. COCO has several features:**

* **Object segmentation**
* **Recognition in context**

**And a lot ….**

**As it is the most popular dataset for pose estimation is the COCO dataset. It has around 80 categories of images and around 250,000 instances of people.**

****

**(6) Details of the algorithm(s)/approach(es) used and the results of the experiments**

**The program has been written completely in python Using ANN algorithm && CNN algorithm**

**Structure of Algorithm 1**

**The program is divided into one packages and 4 files**

* **webcamdetection**
* **ImageDetectionwithGUI**
* **VideoDetectionwithGUI**
* **Dataset(cocoNames)**

**Webcamdetection:**

**The Webcamdetection file holds the other two file:**

* + **webcamdetection function**
  + **main function ( which contains the GUI of which contains ‘interface’)**

**ImageDetectionwithGUI:**

**The ImageDetectionwithGUI file holds:**

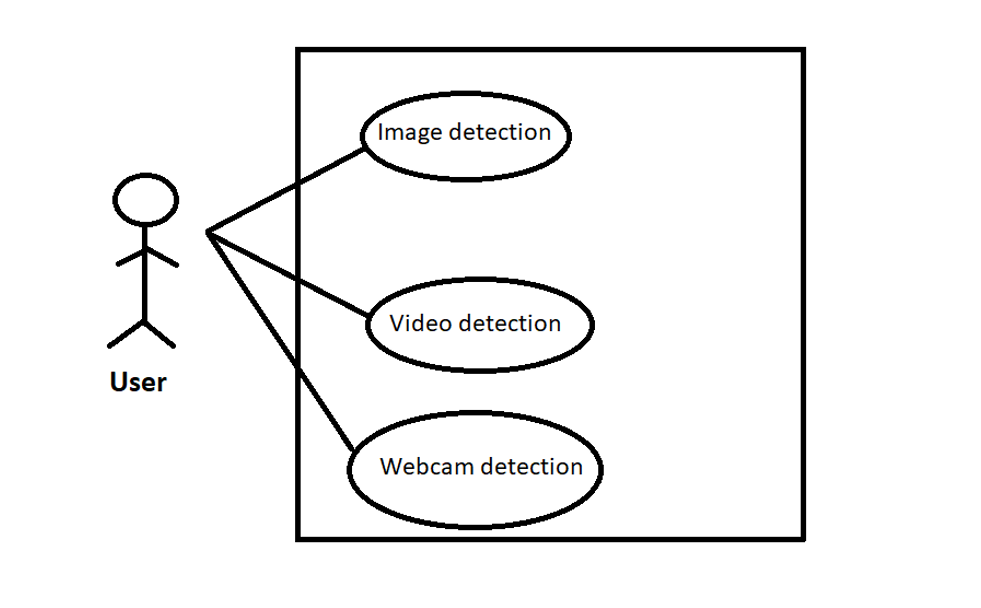
* + **detectImage function(which contain the ANN algorithm that detect image )**
  + **main function which contains the GUI of ImageDetectionwithGUI file**

**VideoDetectionwithGUI:**

**The VideoDetectionwithGUI file holds:**

* + **videoDetect function (which contain the ANN algorithm that detect video object)**
  + **main function which contains the GUI of VideoDetectionwithGUI file**

**Use-Case Diagram**



**Structure of Algorithm 2**

**The program is divided into one packages and 3 files**

* **Object Detection**
* **CNN**
* **Test File**
* **Dataset(Person Images, Car Images, Cat Images, Dog Images)**

**Object Detection**

**Take the dataset and encode it and convert it into X\_pickle and y\_pickle**

**CNN**

**Enter the X\_pickle and y\_pickle into the CNN algorithm , then train them , then put them into CNN.model**

**Test**

**Enter the CNN file to test the image to detect it**

**Packages and libraries in project**

1. **OpenCV**

**OpenCV (Open-Source Computer Vision Library) is an open-source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products. Being a BSD-licensed product, OpenCV makes it easy for businesses to utilize and modify the code.**

1. **PySimpleGUI**

**PySimpleGUI is a Python package that enables Python programmers of all levels to create GUIs. You specify your GUI window using a "layout" which contains widgets (they're called "Elements" in PySimpleGUI). Your layout is used to create a window using one of the 4 supported frameworks to display and interact with your window. Supported frameworks include tkinter, Qt, WxPython, or Remi. The term "wrapper" is sometimes used for these kinds of packages.**

1. **OS**

**This module provides a portable way of using operating system dependent functionality. If you just want to read or write a file see**[**open()**](https://docs.python.org/3/library/functions.html#open)**, if you want to manipulate paths, see the**[**os.path**](https://docs.python.org/3/library/os.path.html#module-os.path)**module, and if you want to read all the lines in all the files on the command line see the**[**fileinput**](https://docs.python.org/3/library/fileinput.html#module-fileinput)**module. For creating temporary files and directories see the**[**tempfile**](https://docs.python.org/3/library/tempfile.html#module-tempfile)**module, and for high-level file and directory handling see the [shutil](https://docs.python.org/3/library/shutil.html" \l "module-shutil" \o "shutil: High-level file operations, including copying.) module.**

1. **TesnorFlow**

**TensorFlow offers multiple levels of abstraction so you can choose the right one for your needs. Build and train models by using the high-level Keras API, which makes getting started with TensorFlow and machine learning easy.**

**If you need more flexibility, eager execution allows for immediate iteration and intuitive debugging. For large ML training tasks, use the Distribution Strategy API for distributed training on different hardware configurations without changing the model definition.**

1. **Pickle**

**Python pickle module is used for serializing and de-serializing python object structures. The process to converts any kind of python objects (list, dict, etc.) into byte streams (0s and 1s) is called pickling or serialization or flattening or marshalling. We can converts the byte stream (generated through pickling) back into python objects by a process called as unpickling.**

1. **NumPy**

**NumPy is a Python library used for working with arrays.**

**It also has functions for working in domain of linear algebra, fourier transform, and matrices.**

**NumPy was created in 2005 by Travis Oliphant. It is an open source project and you can use it freely.**

**NumPy stands for Numerical Python.**

**(7) Development platform**

* **Platform:**

**anaconda (spyder), visual studio, pycharm**

* **Libraries and packages:**

**opencv, PySimpleGUI, os**

* **Operating System and Computer Hardware:**

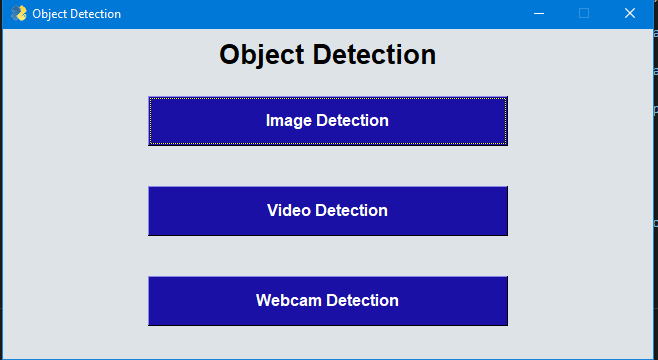
**Laptop run on windows 10**

* **Python Version:**

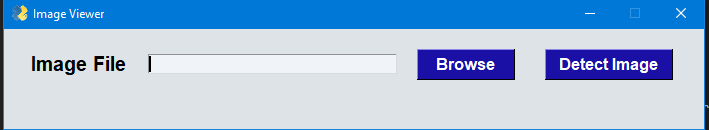
**3.9.8**

**(8)** **Experiments & Results**

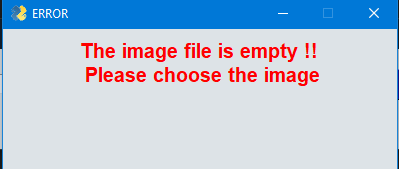
**Firstly, We run the Main File which contain the three options Image Detection , Video Detection, Webcam Detection**

****

**Then, If we press on the Image Detection button, then Enter the path of the image**

****

**Then if the Image File Text is empty and you press on the Detect Image button, the ERROR message will appear**

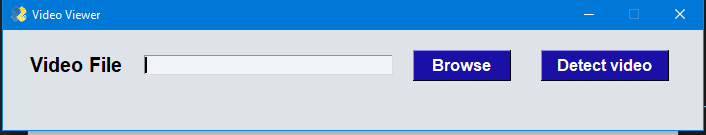
****

**Else, It will Detect the Image**

**صورة تحتوي على نص, دراجة بخارية, عشب, خارجي

تم إنشاء الوصف تلقائياً**

**Then, If we go back and press on the Video Detection button to enter the path of the video**

****

**And again, if the Video File Text is empty and you press on Detect Video button, the ERROR message will appear**

**صورة تحتوي على نص

تم إنشاء الوصف تلقائياً**

**Else It will Detect the Video**

**صورة تحتوي على نص, مبنى, خارجي

تم إنشاء الوصف تلقائياً**

**Then, if we go back and press on the Webcam Detection it will open the camera of the device and will detect the objects.**

**(9) Analysis, Discussion, and Future Work**

* **Analysis of the results, what are the insights?**

**The code can detect the object from the image, video or the webcam**

* **What are the advantages / disadvantages of Object Detection Using ANN ?**
* **The Advantages:**

1. **Improve Accuracy**
2. **Deliver Faster Results**
3. **Reduce Costs**
4. **Provide Unbiased Results**
5. **Offer a Unique Customer Experience**

* **The Disadvantages:**

**One of the most controversial aspects of object detection projects is the potential for invasion of privacy. Facial recognition software is especially a contentious issue, particularly for individuals concerned about privacy invasion through surveillance online or in the actual world.**

* **Why did the algorithm behave in such a way? What might be the future modifications you’d like to try when solving this problem?**
* **As we using ANN ahgorithm**
* **We'd like to try Open-World Learning and Active Vision**

**An important problem is to incrementally learn, to detect new classes, or to incrementally learn to distinguish among subclasses after the “main” class has been learned. If this can be done in an unsupervised way, we will be able to build new classifiers based on existing ones, without much additional effort, greatly reducing the effort required to learn new object classes. Note that humans are continuously inventing new objects, fashion changes, etc., and therefore detection systems will need to be continuously updated, adding new classes, or updating existing ones. Some recent works have addressed these issues, mostly based on deep learning and transfer learning methods [e.g., [Bengio (2012)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B9), [Mesnil et al. (2012)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B45), and [Kotzias et al. (2014)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B39)]. This open-world learning is of particular importance in robot applications, case where active vision mechanisms can aid in the detection and learning [e.g., [Paletta and Pinz (2000)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B53) and**[**Correa et al. (2012)**](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B13)**].**

**Or we would try Pixel-Level Detection (Segmentation) and Background Objects**

**In many applications, we may be interested in detecting objects that are usually considered as background. The detection of such “background objects,” such as rivers, walls, mountains, has not been addressed by most of the here mentioned approaches. In general, this kind of problem has been addressed by first segmenting the image and later labeling each segment of the image [e.g.,**[**Peng et al. (2013)**](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B57)**]. Of course, for successfully detecting all objects in a scene, and to completely understand the scene, we will need to have a pixel level detection of the objects, and further more, a 3D model of such scene. Therefore, at some point object detection and image segmentation methods may need to be integrated. We are still far from attaining such automatic understanding of the world, and to achieve this, active vision mechanisms might be required [e.g., [Aloimonos et al. (1988)](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full" \l "B3) and**[**Cadena et al. (2015)**](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full#B12)**].**

***Recourses***

[**https://www.datacamp.com/community/tutorials/neural-network-models-r**](https://www.datacamp.com/community/tutorials/neural-network-models-r)

[**https://www.fritz.ai/object-detection/#part-basics**](https://www.fritz.ai/object-detection/#part-basics)

[**www.datasciencecentral.com**](http://www.datasciencecentral.com)

[**https://opencv.org/about/**](https://opencv.org/about/)

[**https://pysimplegui.readthedocs.io/en/latest/**](https://pysimplegui.readthedocs.io/en/latest/)

[**https://pypi.org/project/PySimpleGUI/**](https://pypi.org/project/PySimpleGUI/)

[**https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full**](https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full)

[**https://www.upgrad.com/blog/trending-object-detection-project-ideas/**](https://www.upgrad.com/blog/trending-object-detection-project-ideas/)

**mathematica.stackexchange.com**

**scholar.google.com**

**www.geeksforgeeks.org**

**ieeexplore.ieee.org**

**arxiv.org**

**Youtube.com**

**The Link of the project on the drive:**

[**https://github.com/AbdelrahmanSayed672001/ObjectDetection**](https://github.com/AbdelrahmanSayed672001/ObjectDetection)