ImageAI is a python library built to empower developers, researchers and students to build applications and systems with self-contained Deep Learning and Computer Vision capabilities using simple and few lines of code. This documentation is provided to provide detailed insight into all the classes and functions available in ImageAI, coupled with a number of code examples. ImageAI is a project developed by Moses Olafenwa.

The Official GitHub Repository of ImageAI is https://github.com/OlafenwaMoses/ImageAI
CHAPTER 1

ImageAI now uses PyTorch backend.

For full details on this and if you plan on using existing Tensorflow pretrained models, custom models and Pascal VOC dataset, visit the BACKEND_MIGRATION.md documentation.
Chapter 1. ImageAI now uses PyTorch backend.
Installing ImageAI

ImageAI requires that you have Python 3.7.6 installed as well as some other Python libraries and frameworks. Before you install ImageAI, you must install the following dependencies.

- Download and Install Python 3.7, Python 3.8, Python 3.9 or Python 3.10
- Install Dependencies (CPU)

```
pip install cython pillow>=7.0.0 numpy>=1.18.1 opencv-python>=4.1.2 torch>=1.9.0 --extra-index-url https://download.pytorch.org/wb/cpu torchvision>=0.10.0 --extra-index-url https://download.pytorch.org/wb/cpu pytest==7.1.3 tqdm==4.64.1 --scipy>=1.7.3 matplotlib>=3.4.3 mock==4.0.3
```

- Install Dependencies (GPU/CUDA)

```
pip install cython pillow>=7.0.0 numpy>=1.18.1 opencv-python>=4.1.2 torch>=1.9.0 --extra-index-url https://download.pytorch.org/wb/cpu torchvision>=0.10.0 --extra-index-url https://download.pytorch.org/wb/cpu pytest==7.1.3 tqdm==4.64.1 --1 scipy>=1.7.3 matplotlib>=3.4.3 mock==4.0.3
```

- If you plan to train custom AI models, run the command below to install an extra dependency

```
pip install pycocotools@git+https://github.com/gautamchitnis/cocoapi.git@cocodataset-master#subdirectory=PythonAPI
```

- ImageAI

```
pip install imageai --upgrade
```

Once ImageAI is installed, you can start running very few lines of code to perform very powerful computer visions tasks as seen below.

**Image Recognition**

*Recognize 1000 different objects in images*

- convertible : 52.459555864334106
• sports_car : 37.61284649372101
• pickup : 3.1751200556755066
• car_wheel : 1.817505806684494
• minivan : 1.7487050965428352

Visit Documentation

**Image Object Detection**

*Detect 80 most common everyday objects in images.*

Visit Documentation

**Video Object Detection**

*Detect 80 most common everyday objects in videos.*

Visit Documentation

**Video Detection Analysis**

*Generate time based analysis of objects detected in videos.*

Visit Documentation

**Custom Image Recognition Training and Inference**

*Train new image new deep learning models on recognize custom objects*

Visit Documentation

**Custom Objects Detection Training and Inference**
When a Software Engineer get married to another Software Engineer then this happens.
Train new YOLOv3 models to detect custom objects

Visit Documentation

Follow the links in the Content section below to see all the code samples and full documentation of available classes and functions.

2.1 Prediction Classes

ImageAI provides very powerful yet easy to use classes to perform Image Recognition tasks. You can perform all of these state-of-the-art computer vision tasks with python code that ranges between just 5 lines to 12 lines. Once you have Python, other dependencies and ImageAI installed on your computer system, there is no limit to the incredible applications you can create. Find below the classes and their respective functions available for you to use. These classes can be integrated into any traditional python program you are developing, be it a website, Windows/Linux/MacOS application or a system that supports or part of a Local-Area-Network.

======= imageai.Classification.ImageClassification ======

The ImageClassification class provides you the functions to use state-of-the-art image recognition models like MobileNetV2, ResNet50, InceptionV3 and DenseNet121 that were pre-trained on the ImageNet-1000 dataset. This means you can use this class to predict/recognize 1000 different objects in any image or number of images. To initiate the class in your code, you will create a new instance of the class in your code as seen below

```
from imageai.Classification import ImageClassification
prediction = ImageClassification()
```

We have provided pre-trained MobileNetV2, ResNet50, InceptionV3 and DenseNet121 image recognition models which you use with your ImageClassification class to recognize images. Find below the link to download the pre-trained models. You can download the model you want to use.

Download MobileNetV2 Model
Download ResNet50 Model
Download InceptionV3 Model
Download DenseNet121 Model

After creating a new instance of the `ImageClassification` class, you can use the functions below to set your instance property and start recognizing objects in images.

- `.setModelTypeAsMobileNetV2()`, This function sets the model type of the image recognition instance you created to the `MobileNetV2` model, which means you will be performing your image prediction tasks using the pre-trained “MobileNetV2” model you downloaded from the links above. Find example code below

```python
prediction.setModelTypeAsMobileNetV2()
```

- `.setModelTypeAsResNet50()`, This function sets the model type of the image recognition instance you created to the `ResNet50` model, which means you will be performing your image prediction tasks using the pre-trained “ResNet50” model you downloaded from the links above. Find example code below

```python
prediction.setModelTypeAsResNet50()
```

- `.setModelTypeAsInceptionV3()`, This function sets the model type of the image recognition instance you created to the `InceptionV3` model, which means you will be performing your image prediction tasks using the pre-trained “InceptionV3” model you downloaded from the links above. Find example code below

```python
prediction.setModelTypeAsInceptionV3()
```

- `.setModelTypeAsDenseNet121()`, This function sets the model type of the image recognition instance you created to the `DenseNet121` model, which means you will be performing your image prediction tasks using the
pre-trained “DenseNet121” model you downloaded from the links above. Find example code below

```python
prediction.setModelTypeAsDenseNet121()
```

- **.setModelPath()**, This function accepts a string which must be the path to the model file you downloaded and must corresponds to the model type you set for your image prediction instance. Find example code, and parameters of the function below

```python
prediction.setModelPath("resnet50-19c8e357.pth")
```

— parameter **model_path** (required) : This is the path to your downloaded model file.

- **.loadModel()**, This function loads the model from the path you specified in the function call above into your image prediction instance. Find example code below

```python
prediction.loadModel()
```

- **.classifyImage()**, This is the function that performs actual classification of an image. It can be called many times on many images once the model as been loaded into your prediction instance. Find example code, parameters of the function and returned values below

```python
predictions, probabilities = prediction.classifyImage("image1.jpg", result_count=10)
```

— parameter **image_input** (required) : This refers to the path to your image file, Numpy array of your image or image file stream of your image, depending on the input type you specified.

— parameter **result_count** (optional) : This refers to the number of possible predictions that should be returned. The parameter is set to 5 by default.

— returns **prediction_results** (a python list) : The first value returned by the `classifyImage` function is a list that contains all the possible prediction results. The results are arranged in descending order of the percentage probability.

— returns **prediction_probabilities** (a python list) : The second value returned by the `classifyImage` function is a list that contains the corresponding percentage probability of all the possible predictions in the `prediction_results`.

- **.useCPU()**, This function loads the model in CPU and forces processes to be done on the CPU. This is because by default, ImageAI will use GPU/CUDA if available else default to CPU. Find example code:

```python
prediction.useCPU()
```

**Sample Codes**

Find below sample code for predicting one image

```python
from imageai.Classification import ImageClassification
import os

execution_path = os.getcwd()
prediction = ImageClassification()
prediction.setModelTypeAsResNet50()
prediction.setModelPath(os.path.join(execution_path, "resnet50-19c8e357.pth"))
prediction.loadModel()
predictions, probabilities = prediction.classifyImage(os.path.join(execution_path, "image1.jpg"), result_count=10)
```

(continues on next page)
for eachPrediction, eachProbability in zip(predictions, probabilities):
    print(eachPrediction, " : ", eachProbability)

2.2 Detection Classes

ImageAI provides very powerful yet easy to use classes and functions to perform Image Object Detection and Extraction.

ImageAI allows you to perform all of these with state-of-the-art deep learning algorithms like RetinaNet, YOLOv3 and TinyYOLOv3. With ImageAI you can run detection tasks and analyse images.

Find below the classes and their respective functions available for you to use. These classes can be integrated into any traditional python program you are developing, be it a website, Windows/Linux/MacOS application or a system that supports or part of a Local-Area-Network.

======= imageai.Detection.ObjectDetection =======

This ObjectDetection class provides you function to perform object detection on any image or set of images, using pre-trained models that was trained on the COCO dataset. The models supported are RetinaNet, YOLOv3 and TinyYOLOv3. This means you can detect and recognize 80 different kind of common everyday objects. To get started, download any of the pre-trained model that you want to use via the links below.

Download RetinaNet Model - retinanet_resnet50_fpn_coco-eeacb38b.pth
Download YOLOv3 Model - yolov3.pt
Download TinyYOLOv3 Model - tiny-yolov3.pt

Once you have downloaded the model of your choice, you should create a new instance of the ObjectDetection class as seen in the sample below:
from imageai.Detection import ObjectDetection

detector = ObjectDetection()

Once you have created an instance of the class, you can use the functions below to set your instance property and start detecting objects in images.

- `.setModelTypeAsRetinaNet()` , This function sets the model type of the object detection instance you created to the RetinaNet model, which means you will be performing your object detection tasks using the pre-trained “RetinaNet” model you downloaded from the links above. Find example code below:

```
detector.setModelTypeAsRetinaNet()
```

- `.setModelTypeAsYOLOv3()` , This function sets the model type of the object detection instance you created to the YOLOv3 model, which means you will be performing your object detection tasks using the pre-trained “YOLOv3” model you downloaded from the links above. Find example code below:

```
detector.setModelTypeAsYOLOv3()
```

- `.setModelTypeAsTinyYOLOv3()` , This function sets the model type of the object detection instance you created to the TinyYOLOv3 model, which means you will be performing your object detection tasks using the pre-trained “TinyYOLOv3” model you downloaded from the links above. Find example code below:

```
detector.setModelTypeAsTinyYOLOv3()
```

- `.setModelPath()` , This function accepts a string which must be the path to the model file you downloaded and must corresponds to the model type you set for your object detection instance. Find example code and parameters of the function below:

```
detector.setModelPath("yolov3.pt")
```

- `parameter model_path` (required) : This is the path to your downloaded model file.

- `.loadModel()` . This function loads the model from the path you specified in the function call above into your object detection instance. Find example code below:

```
detector.loadModel()
```

- `.detectObjectsFromImage()` . This is the function that performs object detection task after the model as loaded. It can be called many times to detect objects in any number of images. Find example code below:

```
detections = detector.detectObjectsFromImage(input_image="image.jpg", output_image_path="imagenew.jpg", minimum_percentage_probability=30)
```

- `parameter input_image` (required) : This refers to the path to image file which you want to detect. You can set this parameter to the Numpy array of File stream of any image if you set the parameter `input_type` to “array” or “stream”

  - `parameter output_image_path` (required only if `input_type` = “file”) : This refers to the file path to which the detected image will be saved. It is required only if `input_type` = “file”

  - `parameter minimum_percentage_probability` (optional ) : This parameter is used to determine the integrity of the detection results. Lowering the value shows more objects while increasing the value ensures objects with the highest accuracy are detected. The default value is 50.

  - `parameter output_type` (optional ) : This parameter is used to set the format in which the detected image will be produced. The available values are “file” and “array”. The default value is “file”. If this
parameter is set to “array”, the function will return a Numpy array of the detected image. See sample below:

```python
returned_image, detections = detector.detectObjectsFromImage(input_image="image.jpg", output_type="array", minimum_percentage_probability=30)
```

— *parameter* **display_percentage_probability** (optional) : This parameter can be used to hide the percentage probability of each object detected in the detected image if set to False. The default values is True.

— *parameter* **display_object_name** (optional) : This parameter can be used to hide the name of each object detected in the detected image if set to False. The default values is True.

— *parameter* **extract_detected_objects** (optional) : This parameter can be used to extract and save/return each object detected in an image as a separate image. The default values is False.

— *returns* : The returned values will depend on the parameters parsed into the detectObjectsFromImage() function. See the comments and code below

```python

""
If all required parameters are set and 'output_image_path' is set to a file path you want the detected image to be saved, the function will return:
1. an array of dictionaries, with each dictionary corresponding to the objects detected in the image. Each dictionary contains the following property:
   * name (string)
   * percentage_probability (float)
   * box_points (tuple of x1,y1,x2 and y2 coordinates)
""

detections = detector.detectObjectsFromImage(input_image="image.jpg", output_image_path="imagenew.jpg", minimum_percentage_probability=30)

""
If all required parameters are set and output_type = 'array', the function will return
1. a numpy array of the detected image
2. an array of dictionaries, with each dictionary corresponding to the objects detected in the image. Each dictionary contains the following property:
   * name (string)
   * percentage_probability (float)
   * box_points (list of x1,y1,x2 and y2 coordinates)
""

returned_image, detections = detector.detectObjectsFromImage(input_image="image.jpg", output_type="array", minimum_percentage_probability=30)

""
If extract_detected_objects = True and 'output_image_path' is set to a file path you want the detected image to be saved, the function will return:
1. an array of dictionaries, with each dictionary corresponding to the objects
""
```

(continues on next page)
detected in the image. Each dictionary contains the following property:
  * name (string)
  * percentage_probability (float)
  * box_points (list of x1,y1,x2 and y2 coordinates)
2. an array of string paths to the image of each object extracted from the image

""

detections, extracted_objects = detector.
detectObjectsFromImage(input_image="image.jpg", output_image_path="imagenew.jpg", extract_detected_objects=True, minimum_percentage_probability=30)

""

If extract_detected_objects = True and output_type = 'array', the function will return:
1. a numpy array of the detected image
2. an array of dictionaries, with each dictionary corresponding to the objects detected in the image. Each dictionary contains the following property:
  * name (string)
  * percentage_probability (float)
  * box_points (list of x1,y1,x2 and y2 coordinates)
3. an array of numpy arrays of each object detected in the image

""

returned_image, detections, extracted_objects = detector.
detectObjectsFromImage(input_image="image.jpg", output_type="array", extract_detected_objects=True, minimum_percentage_probability=30)

• .useCPU(), This function loads the model in CPU and forces processes to be done on the CPU. This is because by default, ImageAI will use GPU/CUDA if available else default to CPU. Find example code:

    detector.useCPU()

• .CustomObjects(), This function is used when you want to detect only a selected number of objects. It returns a dictionary of objects and their True or False values. To detect selected objects in an image, you will have to use the dictionary returned by the this function with the detectCustomObjectsFromImage() function. Find the details in the comment and code sample below:

""
There are 80 possible objects that you can detect with the ObjectDetection class, and they are as seen below.

    person, bicycle, car, motorcycle, airplane, bus, train, truck, boat, traffic light, fire hydrant, stop_sign, parking meter, bench, bird, cat, dog, horse, sheep, cow, elephant, bear, zebra, giraffe, backpack, umbrella, handbag, tie, suitcase, frisbee, skis, snowboard, sports ball, kite, baseball bat, baseball glove, skateboard, surfboard, tennis racket, bottle, wine glass, cup, fork, knife, spoon, bowl, banana, apple, sandwich, orange, (continues on next page)

2.2. Detection Classes 15
broccoli, carrot, hot dog, pizza, donut, cake, chair, couch,
potted plant, bed, dining table, toilet, tv, laptop, mouse, remote, keyboard, cell,
phone, microwave, oven, toaster, sink, refrigerator, book, clock, vase, scissors,
teddy bear, hair dryer, toothbrush.

To detect only some of the objects above, you will need to call the CustomObjects
function and set the name of the object(s) you want to detect to through. The rest are False by default. In below example, we detected only chose detect only person and dog.

```python
custom = detector.CustomObjects(person=True, dog=True)
```

- `detectCustomObjectsFromImage()`. This function have all the parameters and returns all the values the `detectObjectsFromImage()` functions does but a slight difference. This function let detect only selected objects in an image. Unlike the normal `detectObjectsFromImage()` function, this needs an extra parameter which is “custom_object” which accepts the dictionary returned by the `CustomObjects()` function. In the sample below, we set the detection funtion to report only detections on persons and dogs:

```python
custom = detector.CustomObjects(person=True, dog=True)
detections = detector.detectCustomObjectsFromImage( custom_objects=custom, input_image=os.path.join(execution_path , "image3.jpg"), output_image_path=os.path.join(execution_path , "image3new-custom.jpg"), minimum_percentage_probability=30)
```

Sample Image Object Detection code

Find below a code sample for detecting objects in an image:

```python
from imageai.Detection import ObjectDetection
import os

execution_path = os.getcwd()
detector = ObjectDetection()
detector.setModelTypeAsYOLOv3()
detector.setModelPath( os.path.join(execution_path , "yolov3.pt"))
detector.loadModel()
detections = detector.detectObjectsFromImage(input_image=os.path.join(execution_path , "image.jpg"), output_image_path=os.path.join(execution_path , "imagenew.jpg"), minimum_percentage_probability=30)

for eachObject in detections:
    print(eachObject["name"] , " : ", eachObject["percentage_probability"] , " : ", eachObject["box_points"])
    print("-----------------------------")
```

2.3 Video and Live-Feed Detection and Analysis

ImageAI provided very powerful yet easy to use classes and functions to perform Video Object Detection and Track-
ing and Video analysis. ImageAI allows you to perform all of these with state-of-the-art deep learning algorithms like RetinaNet, YOLOv3 and TinyYOLOv3. With ImageAI you can run detection tasks and analyse videos and live-video feeds from device cameras and IP cameras. Find below the classes and their respective functions available for you to use. These classes can be integrated into any traditional python program you are developing, be it a website, Windows/Linux/MacOS application or a system that supports or part of a Local-Area-Network.

====== imageai.Detection.VideoObjectDetection ======

This VideoObjectDetection class provides you function to detect objects in videos and live-feed from device cameras and IP cameras, using pre-trained models that was trained on the COCO dataset. The models supported are RetinaNet, YOLOv3 and TinyYOLOv3. This means you can detect and recognize 80 different kind of common everyday objects in any video. To get started, download any of the pre-trained model that you want to use via the links below.

Download RetinaNet Model - retinanet_resnet50_fpn_coco-eeacb38b.pth
Download YOLOv3 Model - yolov3.pt
Download TinyYOLOv3 Model - tiny-yolov3.pt

Once you have downloaded the model you chose to use, create an instance of the VideoObjectDetection as seen below:

```python
from imageai.Detection import VideoObjectDetection
detector = VideoObjectDetection()
```

Once you have created an instance of the class, you can call the functions below to set its properties and detect objects in a video.

- **.setModelTypeAsRetinaNet()**, This function sets the model type of the object detection instance you created to the RetinaNet model, which means you will be performing your object detection tasks using the pre-trained “RetinaNet” model you downloaded from the links above. Find example code below:

  ```python
detector.setModelTypeAsRetinaNet()
```

- **.setModelTypeAsYOLOv3()**, This function sets the model type of the object detection instance you created to the YOLOv3 model, which means you will be performing your object detection tasks using the pre-trained “YOLOv3” model you downloaded from the links above. Find example code below:

  ```python
detector.setModelTypeAsYOLOv3()
```

- **.setModelTypeAsTinyYOLOv3()**, This function sets the model type of the object detection instance you created to the TinyYOLOv3 model, which means you will be performing your object detection tasks using the pre-trained “TinyYOLOv3” model you downloaded from the links above. Find example code below:

  ```python
detector.setModelTypeAsTinyYOLOv3()
```

- **.setModelPath()**, This function accepts a string which must be the path to the model file you downloaded and must corresponds to the model type you set for your object detection instance. Find example code, and parameters of the function below:

  ```python
detector.setModelPath("yolov3.pt")
```

  – parameter **model_path** (required): This is the path to your downloaded model file.

- **.loadModel()**, This function loads the model from the path you specified in the function call above into your object detection instance. Find example code below:

  ```python
detector.loadModel()
```
parameter detection_speed (optional) : This parameter allows you to reduce the time it takes to detect objects in a video by up to 80% which leads to slight reduction in accuracy. This parameter accepts string values. The available values are “normal”, “fast”, “faster”, “fastest” and “flash”. The default values is “normal”.

- useCPU() , This function loads the model in CPU and forces processes to be done on the CPU. This is because by default, ImageAI will use GPU/CUDA if available else default to CPU. Find example code:

```python
detector.useCPU()
```

- detectObjectsFromVideo() , This is the function that performs object detection on a video file or video live-feed after the model has been loaded into the instance you created. Find a full sample code below:

```python
from imageai.Detection import VideoObjectDetection
import os
execution_path = os.getcwd()
detector = VideoObjectDetection()
detector.setModelTypeAsYOLOv3()
detector.setModelPath(os.path.join(execution_path, "yolov3.pt"))
detector.loadModel()

video_path = detector.detectObjectsFromVideo(input_file_path=os.path.join(execution_path, "traffic.mp4"), output_file_path=os.path.join(execution_path, "traffic_detected"), frames_per_second=20, log_progress=True)
print(video_path)
```

- parameter input_file_path (required if you did not set camera_input) : This refers to the path to the video file you want to detect.

- parameter output_file_path (required if you did not set save_detected_video = False) : This refers to the path to which the detected video will be saved. By default, this function saves video .avi format.

- parameter frames_per_second (optional , but recommended) : This parameters allows you to set your desired frames per second for the detected video that will be saved. The default value is 20 but we recommend you set the value that suits your video or camera live-feed.

- parameter log_progress (optional) : Setting this parameter to True shows the progress of the video or live-feed as it is detected in the CLI. It will report every frame detected as it progresses. The default value is False.

- parameter return_detected_frame (optional) : This parameter allows you to return the detected frame as a Numpy array at every frame, second and minute of the video detected. The returned Numpy array will be parsed into the respective per_frame_function, per_second_function and per_minute_function (See details below)

- parameter camera_input (optional) : This parameter can be set in replacement of the input_file_path if you want to detect objects in the live-feed of a camera. All you need is to load the camera with OpenCV’s VideoCapture() function and parse the object into this parameter.

See a full code sample below:
execution_path = os.getcwd()
camera = cv2.VideoCapture(0)
detector = VideoObjectDetection()
detector.setModelTypeAsYOLOv3()
detector.setModelPath(os.path.join(execution_path, "yolov3.pt"))
detector.loadModel()

video_path = detector.detectObjectsFromVideo(camera_input=camera,
    output_file_path=os.path.join(execution_path, "camera_detected_video"),
    frames_per_second=20, log_progress=True, minimum_percentage_probability=30)
print(video_path)

—parameter minimum_percentage_probability (optional): This parameter is used to determine the integrity of the detection results. Lowering the value shows more objects while increasing the value ensures objects with the highest accuracy are detected. The default value is 50.

—parameter display_percentage_probability (optional): This parameter can be used to hide the percentage probability of each object detected in the detected video if set to False. The default values is True.

—parameter display_object_name (optional): This parameter can be used to hide the name of each object detected in the detected video if set to False. The default values is True.

—parameter save_detected_video (optional): This parameter can be used to or not to save the detected video or not to save it. It is set to True by default.

—parameter per_frame_function (optional): This parameter allows you to parse in the name of a function you define. Then, for every frame of the video that is detected, the function will be parsed into the parameter will be executed and any analytical data of the video will be parsed into the function. The data returned can be visualized or saved in a NoSQL database for future processing and visualization.

See the sample code below:

```python
""
This parameter allows you to parse in a function you will want to execute after each frame of the video is detected. If this parameter is set to a function, after every video frame is detected, the function will be executed with the following values parsed into it:
  -- position number of the frame
  -- an array of dictionaries, with each dictionary corresponding to each object detected.
    Each dictionary contains 'name', 'percentage_probability' and 'box_points'
    -- a dictionary with with keys being the name of each unique object and value
    are the number of instances of each of the objects present
  -- If return_detected_frame is set to True, the numpy array of the detected frame will be parsed as the fourth value into the function
""
```
from imageai.Detection import VideoObjectDetection
import os

def forFrame(frame_number, output_array, output_count):
    print("FOR FRAME " , frame_number)
    print("Output for each object : ", output_array)
    print("Output count for unique objects : ", output_count)
    print("----------END OF A FRAME ----------")

video_detector = VideoObjectDetection()
video_detector.setModelTypeAsYOLOv3()
video_detector.setModelPath(os.path.join(execution_path, "yolov3.pt"))
video_detector.loadModel()

video_detector.detectObjectsFromVideo(input_file_path=os.path.join(execution_path, "traffic.mp4"),
                                      output_file_path=os.path.join(execution_path, "video_frame_analysis"),
                                      frames_per_second=20,
                                      per_frame_function=forFrame,
                                      minimum_percentage_probability=30)

In the above example, once every frame in the video is processed and detected, the function
will receive and prints out the analytical data for objects detected in the video frame as you can see below:

```
Output for each object : [{'box_points': (362, 295, 443, 355), 'name': 'boat', 'percentage_probability': 26.666194200515747}, {'box_points': (319, 245, 386, 296), 'name': 'boat', 'percentage_probability': 30.052968859672546}, {'box_points': (219, 308, 341, 358), 'name': 'boat', 'percentage_probability': 47.46982455253601}, {'box_points': (589, 198, 621, 241), 'name': 'bus', 'percentage_probability': 24.62330162525177}, {'box_points': (519, 181, 583, 263), 'name': 'bus', 'percentage_probability': 27.446213364601135}, {'box_points': (493, 197, 561, 272), 'name': 'bus', 'percentage_probability': 59.81815455253601}, {'box_points': (432, 187, 491, 240), 'name': 'bus', 'percentage_probability': 64.42965269088745}, {'box_points': (157, 225, 220, 255), 'name': 'car', 'percentage_probability': 21.150341629981995}, {'box_points': (152, 275, 260, 327), 'name': 'car', 'percentage_probability': 24.089913070201874}, {'box_points': (184, 226, 233, 260), 'name': 'car', 'percentage_probability': 37.205660343170166}, {'box_points': (184, 226, 233, 260), 'name': 'car', 'percentage_probability': 30.34144319456482}, {'box_points': (3, 296, 134, 359), 'name': 'car', 'percentage_probability': 38.5252535341702}, {'box_points': (3, 296, 134, 359), 'name': 'car', 'percentage_probability': 47.80363142490387}, {'box_points': (357, 302, 439, 359), 'name': 'car', 'percentage_probability': 47.94844686985016}, {'box_points': (481, 266, 546, 314), 'name': 'car', 'percentage_probability': 65.858578619458}, {'box_points': (597, 269, 624, 318), 'name': 'person', 'percentage_probability': 27.125394344329834}]
Output count for unique objects : {'bus': 4, 'boat': 3, 'person': 1, 'car': 8}
```
Below is a full code that has a function that tasks the analytical data and visualizes it and the detected frame in real time as the video is processed and detected:

```python
from imageai.Detection import VideoObjectDetection
import os
from matplotlib import pyplot as plt

execution_path = os.getcwd()


resized = False

def forFrame(frame_number, output_array, output_count, returned_frame):
    plt.clf()
    this_colors = []
    labels = []
    sizes = []
    counter = 0
```

(continues on next page)
for eachItem in output_count:
    counter += 1
    labels.append(eachItem + " = " + str(output_count[eachItem]))
    sizes.append(output_count[eachItem])
    this_colors.append(color_index[eachItem])

    global resized
    if (resized == False):
        manager = plt.get_current_fig_manager()
        manager.resize(width=1000, height=500)
        resized = True
        plt.subplot(1, 2, 1)
        plt.title("Frame : " + str(frame_number))
        plt.axis("off")
        plt.imshow(returned_frame, interpolation="none")
        plt.subplot(1, 2, 2)
        plt.title("Analysis: " + str(frame_number))
        plt.pie(sizes, labels=labels, colors=this_colors, shadow=True,
                startangle=140, autopct="%.1f%%")
        plt.pause(0.01)

video_detector = VideoObjectDetection()
video_detector.setModelTypeAsYOLOv3()
video_detector.setModelPath(os.path.join(execution_path, "yolov3.pt"))
video_detector.loadModel()
plt.show()

video_detector.detectObjectsFromVideo(input_file_path=os.path.join(execution_path, "traffic.mp4"), output_file_path=os.path.join(execution_path, "video_frame_analysis"), frames_per_second=20, per_frame_function=forFrame, minimum_percentage_probability=30, return_detected_frame=True)

—parameter per_second_function (optional) : This parameter allows you to parse in the name of a function you define. Then, for every second of the video that is detected, the function will be parsed into the parameter will be executed and analytical data of the video will be parsed into the function. The data returned can be visualized or saved in a NoSQL database for future processing and visualization.

See the sample code below:

""
This parameter allows you to parse in a function you will want to
execute after
each second of the video is detected. If this parameter is set to a
function, after every second of a video
is detected, the function will be executed with the following values,
-- parsed into it:
-- position number of the second
"""
-- an array of dictionaries whose keys are position number of each frame present in the last second, and the value for each key is the array for each frame that contains the dictionaries for each object detected in the frame.

-- an array of dictionaries, with each dictionary corresponding to each frame in the past second, and the keys of each dictionary are the name of the number of unique objects detected in each frame, and the key values are the number of instances of the objects found in the frame.

-- a dictionary with its keys being the name of each unique object detected throughout the past second, and the key values are the average number of instances of the object found in all the frames contained in the past second.

-- If return_detected_frame is set to True, the numpy array of the detected frame will be parsed as the fifth value into the function.

```python
import os
from imageai.Detection import VideoObjectDetection

def forSeconds(second_number, output_arrays, count_arrays, average_output_count):
    print("SECOND : ", second_number)
    print("Array for the outputs of each frame ", output_arrays)
    print("Array for output count for unique objects in each frame : ", count_arrays)
    print("Output average count for unique objects in the last second: ", average_output_count)
    print("-------------END OF A SECOND --------------")

video_detector = VideoObjectDetection()
video_detector.setModelTypeAsYOLOv3()
video_detector.setModelPath(os.path.join(execution_path, "yolov3.pt"))
video_detector.loadModel()

video_detector.detectObjectsFromVideo(input_file_path=os.path.join(execution_path, "traffic.mp4"), output_file_path=os.path.join(execution_path, "video_second_analysis"), frames_per_second=20, per_second_function=forSecond, minimum_percentage_probability=30)
```

In the above example, once every second in the video is processed and detected, the function will receive and print out the analytical data for objects detected in the video as you can see below:

```
[{'box_points': (362, 295, 443, 355), 'name': 'boat', 'percentage_probability': 26.666194200515747},
{'box_points': (319, 245, 386, 296), 'name': 'boat', 'percentage_probability': 30.052968859672546},
{'box_points': (219, 308, 341, 358), 'name': 'boat', 'percentage_probability': 47.46982455253601},
{'box_points': (589, 198, 621, 241), 'name': 'bus', 'percentage_probability': 24.62330162525177},
{'box_points': (519, 181, 583, 263), 'name': 'bus', 'percentage_probability': 27.446213364601135},
{'box_points': (493, 197, 561, 272), 'name': 'bus', 'percentage_probability': 59.81854734005458},
{'box_points': (432, 187, 491, 240), 'name': 'bus', 'percentage_probability': 64.42965269088745},
{'box_points': (157, 225, 225, 205), 'name': 'car', 'percentage_probability': 21.150341629981995},
{'box_points': (324, 249, 377, 293), 'name': 'car', 'percentage_probability': 24.089913070201874},
{'box_points': (152, 275, 260, 327), 'name': 'car', 'percentage_probability': 30.341443419456482},
{'box_points': (433, 198, 485, 244), 'name': 'car', 'percentage_probability': 37.205660343170166},
{'box_points': (184, 226, 233, 260), 'name': 'car', 'percentage_probability': 38.52525353431702},
{'box_points': (3, 296, 134, 359), 'name': 'car', 'percentage_probability': 47.80363142490387},
{'box_points': (357, 302, 439, 359), 'name': 'car', 'percentage_probability': 47.94844686985016},
{'box_points': (481, 266, 546, 314), 'name': 'car', 'percentage_probability': 65.8585786819458},
{'box_points': (597, 269, 624, 318), 'name': 'person', 'percentage_probability': 27.125394344329834}]
```
Array for output count for unique objects in each frame: 

```json
[{
    'bus': 4,
    'boat': 3,
    'person': 1,
    'car': 8
},
{
    'truck': 1,
    'bus': 4,
    'boat': 3,
    'person': 1,
    'car': 7
},
{
    'bus': 5,
    'boat': 2,
    'person': 1,
    'car': 5
},
{
    'bus': 5,
    'boat': 1,
    'person': 1,
    'car': 9
},
{
    'truck': 1,
    'bus': 2,
    'car': 6,
    'person': 1
},
{
    'truck': 2,
    'bus': 4,
    'boat': 2,
    'person': 1,
    'car': 7
},
{
    'truck': 1,
    'bus': 3,
    'car': 7,
    'person': 1,
    'umbrella': 1
},
{
    'bus': 4,
    'car': 7,
    'person': 1,
    'umbrella': 2
},
{
    'bus': 3,
    'car': 6,
    'boat': 1,
    'person': 1,
    'umbrella': 3
},
{
    'bus': 3,
    'car': 4,
    'boat': 1,
    'person': 1,
    'umbrella': 2
}]
```

Output average count for unique objects in the last second: 

```json
{'truck': 0.5, 'bus': 3.7, 'umbrella': 0.8, 'boat': 1.5, 'person': 1.0, 'car': 6.6}
```
Below is a full code that has a function that tasks the analytical data and visualizes it and the detected frame at the end of the second in real time as the video is processed and detected:

```python
from imageai.Detection import VideoObjectDetection
import os
from matplotlib import pyplot as plt

execution_path = os.getcwd()

color_index = {'bus': 'red', 'handbag': 'steelblue', 'giraffe':
  'elephant': 'pink', 'truck': 'indigo', 'motorcycle': 'azure',
  'refrigerator': 'gold', 'keyboard': 'violet', 'cow': 'magenta',
  'mouse': 'crimson', 'sports ball': 'raspberry', 'horse': 'maroon',
  'cat': 'orchid', 'boat': 'slateblue', 'hot dog': 'navy', 'apple':
  'cobalt', 'parking meter': 'aliceblue', 'sandwich': 'skyblue',
  'skis': 'deepskyblue', 'microwave': 'peacock', 'knife': 'cadetblue
  'baseball bat': 'cyan', 'oven': 'lightcyan', 'carrot': 'coldgrey
  'scissors': 'seagreen', 'sheep': 'deepgreen', 'toothbrush':
  'cobaltgreen', 'fire hydrant': 'limegreen', 'remote': 'forestgreen
  'bicycle': 'olivedrab', 'toilet': 'ivory', 'tv': 'khaki',
  'skateboard': 'palegoldenrod', 'train': 'cornsilk', 'zebra': 'wheat
  'tie': 'burlywood', 'orange': 'melon', 'bird': 'bisque',
  'dining table': 'chocolate', 'hair drier': 'sandybrown', 'cell
  phone': 'sienna', 'sink': 'coral', 'bench': 'salmon', 'bottle':
  'brown', 'car': 'silver', 'bowl': 'maroon', 'tennis racket':
  'palevioletred', 'airplane': 'lavenderblush', 'pizza': 'hotpink',
  'umbrella': 'deepink', 'bear': 'plum', 'fork': 'purple', 'laptop
  ': 'indigo', 'vase': 'mediumpurple', 'baseball glove': 'slateblue',
  'traffic light': 'mediumpurple', 'bed': 'navy', 'broccoli':
  'royalblue', 'backpack': 'slategray', 'snowboard': 'skyblue', 'kite
  ': 'cadetblue', 'teddy bear': 'peacock', 'clock': 'lightcyan',
  'wine glass': 'teal', 'frisbee': 'aquamarine', 'donut': 'mincream',
  'suitcase': 'seagreen', 'dog': 'springgreen', 'banana':
  'emeraldgreen', 'person': 'honeydew', 'surfboard': 'palegreen',
  'cake': 'sagegreen', 'book': 'lawngreen', 'potted plant':
  'greenyellow', 'toaster': 'ivory', 'stop sign': 'beige', 'couch':
  'khaki'}

resized = False

def forSecond(frame2_number, output_arrays, count_arrays, average_
  count, returned_frame):

  plt.clf()

  this_colors = []
  labels = []
  sizes = []
  counter = 0

  (continues on next page)
```
for eachItem in average_count:
    counter += 1
    labels.append(eachItem + " = " + str(average_count[eachItem]))
    sizes.append(average_count[eachItem])
    this_colors.append(color_index[eachItem])

if resized == False:
    manager = plt.get_current_fig_manager()
    manager.resize(width=1000, height=500)
    resized = True

plt.subplot(1, 2, 1)
plt.title("Second : " + str(frame_number))
plt.axis("off")
plt.imshow(returned_frame, interpolation="none")
plt.subplot(1, 2, 2)
plt.title("Analysis: " + str(frame_number))
plt.pie(sizes, labels=labels, colors=this_colors, shadow=True, startangle=140, autopct="%.1f%%")
plt.pause(0.01)

video_detector = VideoObjectDetection()
video_detector.setModelTypeAsYOLOv3()
video_detector.setModelPath(os.path.join(execution_path, "yolov3.pt"))
video_detector.loadModel()

plt.show()

video_detector.detectObjectsFromVideo(input_file_path=os.path.join(execution_path, "traffic.mp4"), output_file_path=os.path.join(execution_path, "video_second_analysis"), frames_per_second=20, per_second_function=forSecond, minimum_percentage_probability=30, return_detected_frame=True, log_progress=True)

—parameter per_minute_function (optional): This parameter allows you to parse in the name of a function you define. Then, for every frame of the video that is detected, the function which was parsed into the parameter will be executed and analytical data of the video will be parsed into the function. The data returned has the same nature as the per_second_function; the difference is that it covers all the frames in the past 1 minute of the video.

See a sample function for this parameter below:

def forMinute(minute_number, output_arrays, count_arrays, average_output_count):
    print("MINUTE : ", minute_number)
    print("Array for the outputs of each frame ", output_arrays)
    print("Array for output count for unique objects in each frame : ", count_arrays)
print("Output average count for unique objects in the last minute: ", average_output_count)
print("-------------END OF A MINUTE -------------")

---parameter video_complete_function (optional)---: This parameter allows you to parse in the name of a function you define. Once all the frames in the video is fully detected, the function will be parsed into the parameter will be executed and analytical data of the video will be parsed into the function. The data returned has the same nature as the per_second_function and per_minute_function; the differences are that no index will be returned and it covers all the frames in the entire video.

See a sample function for this parameter below:

```python
def forFull(output_arrays, count_arrays, average_output_count):
    print("Array for the outputs of each frame ", output_arrays)
    print("Array for output count for unique objects in each frame : ", count_arrays)
    print("Output average count for unique objects in the entire video: ", average_output_count)
    print("-------------END OF THE VIDEO -------------")
```

---parameter detection_timeout (optional)---: This function allows you to state the number of seconds of a video that should be detected after which the detection function stop processing the video.

See a sample code for this parameter below:

```python
from imageai.Detection import VideoObjectDetection
import os
import cv2

execution_path = os.getcwd()

camera = cv2.VideoCapture(0)
detector = VideoObjectDetection()
detector.setModelTypeAsRetinaNet()
detector.setModelPath(os.path.join(execution_path, "retinanet_resnet50_fpn_coco-eeacb38b.pth"))
detector.loadModel()

video_path = detector.detectObjectsFromVideo(camera_input=camera,
                                           output_file_path=os.path.join(execution_path, "camera_detected_video"),
                                           frames_per_second=20, log_progress=True,
                                           minimum_percentage_probability=40, detection_timeout=120)
```

2.4 Custom Training: Prediction

ImageAI provides very powerful yet easy to use classes to train state-of-the-art deep learning algorithms like SqueezeNet, ResNet, InceptionV3 and DenseNet on your own image datasets using as few as 5 lines of code to generate your own custom models. Once you have trained your own custom model, you can use the CustomImagePrediction class provided by ImageAI to use your own models to recognize/predict any image or set of images.

====== imageai.Classification.Custom.ClassificationModelTrainer ======
The **ClassificationModelTrainer** class allows you to train any of the 4 supported deep learning algorithms (**MobileNetV2**, **ResNet50**, **InceptionV3** and **DenseNet121**) on your own image dataset to generate your own custom models. Your image dataset must contain at least 2 different classes/types of images (e.g. cat and dog) and you must collect at least 500 images for each of the classes to achieve maximum accuracy.

The training process generates a JSON file that maps the objects types in your image dataset and creates lots of models. You will then peak the model with the highest accuracy and perform custom image prediction using the model and the JSON file generated.

Because model training is a compute intensive tasks, we strongly advise you perform this experiment using a computer with a NVIDIA GPU and the GPU version of Tensorflow installed. Performing model training on CPU will my take hours or days. With NVIDIA GPU powered computer system, this will take a few hours. You can use Google Colab for this experiment as it has an NVIDIA K80 GPU available. To train a custom prediction model, you need to prepare the images you want to use to train the model. You will prepare the images as follows:

- Create a dataset folder with the name you will like your dataset to be called (e.g pets)
  — In the dataset folder, create a folder by the name train

- In the dataset folder, create a folder by the name test

- In the train folder, create a folder for each object you want to the model to predict and give the folder a name that corresponds to the respective object name (e.g dog, cat, squirrel, snake)

- In the test folder, create a folder for each object you want to the model to predict and give the folder a name that corresponds to the respective object name (e.g dog, cat, squirrel, snake)

- In each folder present in the train folder, put the images of each object in its respective folder. This images are the ones to be used to train the model

- To produce a model that can perform well in practical applications, I recommend you about 500 or more images per object. 1000 images per object is just great

- In each folder present in the test folder, put about 100 to 200 images of each object in its respective folder. These images are the ones to be used to test the model as it trains

- Once you have done this, the structure of your image dataset folder should look like below

```markdown
pets//train//dog//dog-train-images
pets//train//cat//cat-train-images
pets//train//squirrel//squirrel-train-images
pets//train//snake//snake-train-images

pets//test//dog//dog-test-images
pets//test//cat//cat-test-images
pets//test//squirrel//squirrel-test-images
pets//test//snake//snake-test-images
```

Once your dataset is ready, you can proceed to creating an instance of the **ModelTraining** class. Find the example below

```python
from imageai.Classification.Custom import ClassificationModelTrainer

model_trainer = ClassificationModelTrainer()
```

Once you have created an instance above, you can use the functions below to set your instance property and start the training process.

- `.setModelTypeAsMobileNetV2()` , This function sets the model type of the training instance you created to the **MobileNetV2** model, which means the **MobileNetV2** algorithm will be trained on your dataset. Find example code below
model_trainer.setModelTypeAsMobileNetV2()

- **.setModelTypeAsResNet50()**, This function sets the model type of the training instance you created to the ResNet50 model, which means the ResNet50 algorithm will be trained on your dataset. Find example code below

model_trainer.setModelTypeAsResNet()

- **.setModelTypeAsInceptionV3()**, This function sets the model type of the training instance you created to the InceptionV3 model, which means the InceptionV3 algorithm will be trained on your dataset. Find example code below

model_trainer.setModelTypeAsInceptionV3()

- **.setModelTypeAsDenseNet121()**, This function sets the model type of the training instance you created to the DenseNet121 model, which means the DenseNet121 algorithm will be trained on your dataset. Find example code below

model_trainer.setModelTypeAsDenseNet121()

- **.setDataDirectory()**, This function accepts a string which must be the path to the folder that contains the test and train subfolder of your image dataset. Find example code and parameters of the function below

prediction.setDataDirectory(r"C:/Users/Moses/Documents/Moses/AI/Custom Datasets/pets")

- **parameter data_directory** (required): This is the path to the folder that contains your image dataset.

- **.trainModel()**, This is the function that starts the training process. Once it starts, it will create a JSON file in the dataset/json folder which contains the mapping of the classes of the dataset. The JSON file will be used during custom prediction to produce results. Find example code below

model_trainer.trainModel(num_experiments=100, batch_size=32)

- **parameter num_experiments** (required): This is the number of times the algorithm will be trained on your image dataset. The accuracy of your training does increases as the number of times it trains increases. However, it does peak after a certain number of trainings; and that point depends on the size and nature of the dataset.

- **parameter batch_size** (optional): During training, the algorithm is trained on a set of images in parallel. Because of this, the default value is set to 32. You can increase or reduce this value if you understand well enough to know the capacity of the system you are using to train. Should you intend to change this value, you should set it to values that are in multiples of 8 to optimize the training process.

**Sample Code for Custom Model Training**

Find below a sample code for training custom models for your image dataset

```python
from imageai.Classification.Custom import ClassificationModelTrainer

model_trainer = ClassificationModelTrainer()
model_trainer.setModelTypeAsResNet50()
model_trainer.setDataDirectory(r"C:/Users/Moses/Documents/Moses/AI/Custom Datasets/pets")
model_trainer.trainModel(num_objects=10, num_experiments=100, enhance_data=True,
batch_size=32, show_network_summary=True)
```

Below is a sample of the result when the training starts
Let us explain the details shown above:

1. The line Epoch 1/100 means the network is training the first experiment of the targeted 100
2. The line 1/25 [>.........................] - ETA: 52s - loss: 2.3026 - acc: 0.2500 represents the number of
   batches that has been trained in the present experiment
3. The best model is automatically saved to <dataset-directory>/models>

Once you are done training your custom model, you can use the CustomImagePrediction class described below to
perform image prediction with your model.

= imageai.Classification.Custom.CustomImageClassification =

This class can be considered a replica of the imageai.Classification.CustomImageClassification as it has all the same
functions, parameters and results. The only differences are that this class works with your own trained model, you will
need to specify the path to the JSON file generated during the training and will need to specify the number of classes
in your image dataset when loading the model. Below is an example of creating an instance of the class

```python
from imageai.Classification.Custom import CustomImageClassification

prediction = CustomImageClassification()
```

Once you have created the new instance, you can use the functions below to set your instance property and start
recognizing objects in images.

- `setModelTypeAsMobileNetV2()`, This function sets the model type of the image recognition instance you
  created to the MobileNetV2 model, which means you will be performing your image prediction tasks using the
  “MobileNetV2” model generated during your custom training. Find example code below

```python
prediction.setModelTypeAsMobileNetV2()
```

- `setModelTypeAsResNet50()`, This function sets the model type of the image recognition instance you created
to the ResNet50 model, which means you will be performing your image prediction tasks using the “ResNet”
model model generated during your custom training. Find example code below

```python
prediction.setModelTypeAsResNet50()
```

- `setModelTypeAsInceptionV3()`, This function sets the model type of the image recognition instance you
  created to the InceptionV3 model, which means you will be performing your image prediction tasks using the
  “InceptionV3” model generated during your custom training. Find example code below

```python
prediction.setModelTypeAsInceptionV3()
```
prediction.setModelTypeAsInceptionV3()

- .setModelTypeAsDenseNet121() , This function sets the model type of the image recognition instance you created to the DenseNet121 model, which means you will be performing your image prediction tasks using the “DenseNet” model generated during your custom training. Find example code below

prediction.setModelTypeAsDenseNet121()

- .setModelPath() , This function accepts a string which must be the path to the model file generated during your custom training and must corresponds to the model type you set for your image prediction instance. Find example code, and parameters of the function below

prediction.setModelPath("resnet50-idenprof-test_acc_0.78200_epoch-91.pt")

  – parameter model_path (required) : This is the path to your downloaded model file.

- .setJsonPath() , This function accepts a string which must be the path to the JSON file generated during your custom training. Find example code and parameters of the function below

prediction.setJsonPath("idenprof_model_classes.jsonn")

  – parameter model_path (required) : This is the path to your downloaded model file.

- .loadModel() , This function loads the model from the path you specified in the function call above into your image prediction instance. You will have to set the parameter num_objects to the number of classes in your image dataset. Find example code and parameter details below

prediction.loadModel()

- .classifyImage() , This is the function that performs actual prediction of an image. It can be called many times on many images once the model as been loaded into your prediction instance. Find example code, parameters of the function and returned values below

predictions, probabilities = prediction.classifyImage("image1.jpg", result_count=2)

  – parameter image_input (required) : This refers to the path to your image file, Numpy array of your image or image file stream of your image, depending on the input type you specified.

  --- parameter result_count (optional) : This refers to the number of possible predictions that should be returned. The parameter is set to 5 by default.

  -- returns prediction_results (a python list) : The first value returned by the predictImage function is a list that contains all the possible prediction results. The results are arranged in descending order of the percentage probability.

  --- returns prediction_probabilities (a python list) : The second value returned by the predictImage function is a list that contains the corresponding percentage probability of all the possible predictions in the prediction_results.

- .useCPU() , This function loads the model in CPU and forces processes to be done on the CPU. This is because by default, ImageAI will use GPU/CUDA if available else default to CPU. Find example code:

prediction.useCPU()

Sample Codes

Find below sample code for custom prediction
from imageai.Classification.Custom import CustomImageClassification
import os

execution_path = os.getcwd()

prediction = CustomImageClassification()
prediction.setModelTypeAsResNet50()
prediction.setModelPath(os.path.join(execution_path, "resnet50-idenprof-test_acc_0._78200_epoch-91.pt"))
prediction.setJsonPath(os.path.join(execution_path, "idenprof_model_classes.json"))
prediction.loadModel()

predictions, probabilities = prediction.classifyImage(os.path.join(execution_path, "4.jpg"), result_count=5)

for eachPrediction, eachProbability in zip(predictions, probabilities):
    print(eachPrediction, " : ", eachProbability)

---

2.5 Custom Object Detection: Training and Inference

ImageAI provides the simple and powerful approach to training custom object detection models using the YOLOv3 architecture. This allows you to train your own model on any set of images that corresponds to any type of object of interest.

You can use your trained detection models to detect objects in images, videos and perform video analysis.

======== imageai.Detection.Custom.DetectionModelTrainer ========

This is the Detection Model training class, which allows you to train object detection models on image datasets that are in YOLO annotation format, using the YOLOv3 and TinyYOLOv3 model. The training process generates a JSON file that maps the objects names in your image dataset and the detection anchors, as well as creates lots of models.
To get started, you need to prepare your dataset in the **YOLO annotation format** and organize it as detailed below:

- Decide the type of object(s) you want to detect and collect about 200 (minimum recommendation) or more pictures of each of the object(s).

- Once you have collected the images, you need to annotate the object(s) in the images. You can use a tool like LabelIMG to generate the annotations for your images.

- Once you have the annotations for all your images, create a folder for your dataset (**E.g headsets**) and in this parent folder, create child folders **train** and **validation**.

- In the **train** folder, create **images** and **annotations** sub-folders. Put about 70-80% of your dataset of each object's images in the **images** folder and put the corresponding annotations for these images in the **annotations** folder.

- In the **validation** folder, create **images** and **annotations** sub-folders. Put the rest of your dataset images in the **images** folder and put the corresponding annotations for these images in the **annotations** folder.

- Once you have done this, the structure of your image dataset folder should look like below:

```plaintext
>> train >> images >> img_1.jpg (shows Object_1)
>> images >> img_2.jpg (shows Object_2)
>> images >> img_3.jpg (shows Object_1, Object_3 and Object_n)
>> annotations >> img_1.txt (describes Object_1)
>> annotations >> img_2.txt (describes Object_2)
>> annotations >> img_3.txt (describes Object_1, Object_3 and Object_n)

>> validation >> images >> img_151.jpg (shows Object_1, Object_3 and Object_n)
>> images >> img_152.jpg (shows Object_2)
>> images >> img_153.jpg (shows Object_1)
>> annotations >> img_151.txt (describes Object_1, Object_3 and Object_n)
>> annotations >> img_152.txt (describes Object_2)
>> annotations >> img_153.txt (describes Object_1)
```

- You can train your custom detection model completely from scratch or use transfer learning (**recommended for better accuracy**) from a pre-trained YOLOv3 model or TinyYOLOv3. Also, we have provided a sample annotated Hololens and Headsets (Hololens and Oculus) dataset for you to train with. Download the pre-trained YOLOv3 or TinyYOLOv3 model and the sample dataset.

Below is the code to train new detection models on your dataset:

```python
from imageai.Detection.Custom import DetectionModelTrainer

trainer = DetectionModelTrainer()
trainer.setModelTypeAsYOLOv3()
trainer.setDataDirectory(data_directory="hololens")
trainer.setTrainConfig(object_names_array=["hololens"], batch_size=4, num_experiments=200, train_from_pretrained_model="yolov3.pt")
trainer.trainModel()
```

In the first 2 lines, we imported the **DetectionModelTrainer** class and created an instance of it.

```python
from imageai.Detection.Custom import DetectionModelTrainer

trainer = DetectionModelTrainer()
```

Then we called the following functions:

- `.setModelTypeAsYOLOv3()` - This function sets the model type of the object detection training instance to the YOLOv3 model:
trainer.setModelTypeAsYOLOv3()

- **trainer.setDataDirectory()**. This function sets the path to your dataset’s folder:

  \[
  \text{trainer.setDataDirectory()}
  \]

  - **parameter** `data_directory` (required) : This is the path to your dataset folder.

- **trainer.setTrainConfig()**. This function sets the properties for the training instances:

  \[
  \text{trainer.setTrainConfig()}
  \]

  - **parameter** `object_names_array` (required) : This is a list of the names of all the different objects in your dataset.
  - **parameter** `batch_size` (optional) : This is the batch size for the training instance.
  - **parameter** `num_experiments` (required) : Also known as epochs, it is the number of times the network will train on all the training.
  - **parameter** `train_from_pretrained_model` (optional) : This is used to perform transfer learning by specifying the path to a pre-trained YOLOv3 model

When you run the training code, **ImageAI** will perform the following actions:

- generate a **.json** in the **dataset_folder/json** folder. Please note that the **JSON** file generated in a training session can only be used with the **detection models** saved in the training session.
- saves new models n the **dataset_folder/models** folder as the training loss reduces.

As the training progresses, the information displayed in the terminal will look similar to the sample below:

```
Generating anchor boxes for training images...
thr=0.25: 1.0000 best possible recall, 6.93 anchors past thr
n=9, img_size=416, metric_all=0.463/0.856-mean/best, past_thr=0.549-mean:
===============
Pretrained YOLOv3 model loaded to initialize weights
===============
Epoch 1/100
----------
Train: 30it [00:14,  2.09it/s]
  box loss-> 0.09820, object loss-> 0.27985, **class loss**-> 0.00000
Validation: 15it [01:45, 7.05s/it]
  recall: 0.085714 precision: 0.000364 mAP@0.5: 0.000186, mAP@0.5-0.95: 0.000030
Epoch 2/100
----------
Train: 30it [00:07, 4.25it/s]
  box loss-> 0.08691, object loss-> 0.07011, **class loss**-> 0.00000
Validation: 15it [01:37, 6.53s/it]
  recall: 0.214286 precision: 0.000854 mAP@0.5: 0.000516, mAP@0.5-0.95: 0.000030
```

For each increase in the mAP0.5 after an experiment, a model is saved in the hololens-yolo/models folder. The higher the mAP0.5, the better the model.

```
======= imageai.Detection.Custom.CustomObjectDetection =======
```

2.5. **Custom Object Detection: Training and Inference**
CustomObjectDetection class provides very convenient and powerful methods to perform object detection on images and extract each object from the image using your own custom YOLOv3 model and the corresponding .json generated during the training.

To test the custom object detection, you can download a sample custom model we have trained to detect the Hololens headset and its detection_config.json file via the links below:

Hololens Detection Model
detection_config.json

- Sample Image

Once you download the custom object detection model file, you should copy the model file to the your project folder where your .py files will be. Then create a python file and give it a name; an example is FirstCustomDetection.py. Then write the code below into the python file:

```python
from imageai.Detection.Custom import CustomObjectDetection

detector = CustomObjectDetection()
detector.setModelTypeAsYOLOv3()
detector.setModelPath("yolov3_hololens-yolo_mAP-0.82726_epoch-73.pt")
detector.setJsonPath("hololens-yolo_yolov3_detection_config.json")
detector.loadModel()
detections = detector.detectObjectsFromImage(input_image="holo1.jpg", output_image_path="holo1-detected.jpg")

for detection in detections:
    print(detection["name"], " : ", detection["percentage_probability"], " : ", detection["box_points"])
```

When you run the code, it will produce a result similar to the one below:

```
hololens : 39.69653248786926 : [611, 74, 751, 154]
hololens : 87.6643180847168 : [23, 46, 90, 79]
```
See more details below:

- `.setModelTypeAsYOLOv3()` , This specifies that you are using a trained YOLOv3 model

```python
detector.setModelTypeAsYOLOv3()
```

- `.setModelPath()` , This is used to set the file path to your trained model

```python
detector.setModelPath()
```

  - parameter `model_path` (required) : This is path to your model file

- `.setJsonPath()` , This is used to set the file path to your configuration json file

```python
detector.setJsonPath()
```

  - parameter `configuration_json` (required) : This is path to `.json` file

- `.loadModel()` , This is load the detection model:

```python
detector.loadModel()
```

- `.detectObjectsFromImage()` , This is the function that performs object detection task after the model as loaded. It can be called many times to detect objects in any number of images. Find example code below:

```python
detections = detector.detectObjectsFromImage(input_image="image.jpg", output_image_path="imagenew.jpg", minimum_percentage_probability=30)
```
– **parameter** `input_image` *(required)*: This refers to the path to image file which you want to detect. You can set this parameter to the Numpy array of File stream of any image if you set the parameter `input_type` to “array” or “stream”

– **parameter** `output_image_path` *(required only if input_type = “file”)*: This refers to the file path to which the detected image will be saved. It is required only if `input_type` = “file”

– **parameter** `minimum_percentage_probability` *(optional)*: This parameter is used to determine the integrity of the detection results. Lowering the value shows more objects while increasing the value ensures objects with the highest accuracy are detected. The default value is 50.

– **parameter** `output_type` *(optional)*: This parameter is used to set the format in which the detected image will be produced. The available values are “file” and “array”. The default value is “file”. If this parameter is set to “array”, the function will return a Numpy array of the detected image. See sample below

– **parameter** `display_percentage_probability` *(optional)*: This parameter can be used to hide the percentage probability of each object detected in the detected image if set to False. The default values is True.

– **parameter** `display_object_name` *(optional)*: This parameter can be used to hide the name of each object detected in the detected image if set to False. The default values is True.

– **parameter** `extract_detected_objects` *(optional)*: This parameter can be used to extract and save/return each object detected in an image as a separate image. The default values is False.

– **returns**: The returned values will depend on the parameters parsed into the `detectObjectsFromImage()` function. See the comments and code below

- `useCPU()`, This function loads the model in CPU and forces processes to be done on the CPU. This is because by default, ImageAI will use GPU/CUDA if available else default to CPU. Find example code:

```python
detector.useCPU()
```

====== imageai.Detection.Custom.CustomVideoObjectDetection ======

CustomVideoObjectDetection class provides very convenient and powerful methods to perform object detection on videos and obtain analytical from the video, using your own custom YOLOv3 model and the corresponding detection_config.json generated during the training.

To test the custom object detection, you can download a sample custom model we have trained to detect the Hololens headset and its detection_config.json file via the links below:

Hololens Detection Model
detection_config.json

Download a sample video of the Hololens in the link below.

Sample Hololens Video

Then run the code below in the video:

```python
from imageai.Detection.Custom import CustomVideoObjectDetection
import os

execution_path = os.getcwd()

video_detector = CustomVideoObjectDetection()
video_detector.setModelTypeAsYOLOv3()
video_detector.setModelPath("yolov3_hololens-yolo_mAP-0.82726_epoch-73.pt")
```
video_detector.setJsonPath("hololens-yolo_yolov3_detection_config.json")
video_detector.setModelTypeAsYOLOv3()
video_detector.setModelPath()
video_detector.detectObjectsFromVideo(input_file_path="holol.mp4",
output_file_path=os.path.join(execution_path,"holol-detected"),
frames_per_second=30,
minimum_percentage_probability=40,
log_progress=True)

See details on the available functions below

• .setModelTypeAsYOLOv3() , This specifies that you are using a trained YOLOv3 model

• .setModelPath() , This is used to set the file path to your trained model

• .setJsonPath() , This is used to set the file path to your configuration json file

• .loadModel() , This is load the detection model:

• .useCPU() , This function loads the model in CPU and forces processes to be done on the CPU. This is because by default, ImageAI will use GPU/CUDA if available else default to CPU. Find example code:

• .detectObjectsFromVideo() , This is the function that performs object detection on a video file or video live-feed after the model has been loaded into the instance you created. Find a full sample code below:

  – parameter input_file_path (required if you did not set camera_input) : This refers to the path to the video file you want to detect.

  —parameter output_file_path (required if you did not set save_detected_video = False) : This refers to the path to which the detected video will be saved. By default, this function saves video .avi format.

  – parameter frames_per_second (optional , but recommended) : This parameters allows you to set your desired frames per second for the detected video that will be saved. The default value is 20 but we recommend you set the value that suits your video or camera live-feed.

  —parameter log_progress (optional) : Setting this parameter to True shows the progress of the video or live-feed as it is detected in the CLI. It will report every frame detected as it progresses. The default value is False.

  – parameter return_detected_frame (optional) : This parameter allows you to return the detected frame as a Numpy array at every frame, second and minute of the video detected. The returned Numpy array will be parsed into the respective per_frame_function, per_second_function and per_minute_function (See details below)
—parameter camera_input (optional): This parameter can be set in replacement of the input_file_path if you want to detect objects in the live-feed of a camera. All you need is to load the camera with OpenCV’s VideoCapture() function and parse the object into this parameter.

— parameter minimum_percentage_probability (optional): This parameter is used to determine the integrity of the detection results. Lowering the value shows more objects while increasing the value ensures objects with the highest accuracy are detected. The default value is 50.

— parameter display_percentage_probability (optional): This parameter can be used to hide the percentage probability of each object detected in the detected video if set to False. The default values is True.

— parameter display_object_name (optional): This parameter can be used to hide the name of each object detected in the detected video if set to False. The default values is True.

— parameter save_detected_video (optional): This parameter can be used to or not to save the detected video or not to save it. It is set to True by default.

— parameter per_frame_function (optional): This parameter allows you to parse in the name of a function you define. Then, for every frame of the video that is detected, the function will be parsed into the parameter will be executed and analytical data of the video will be parsed into the function. The data returned can be visualized or saved in a NoSQL database for future processing and visualization.

See a sample function for this parameter below .. code-block:

```python
This parameter allows you to parse in a function you will want to execute after each frame of the video is detected. If this parameter is set to a function, after every video frame is detected, the function will be executed with the following values parsed into it:
- position number of the frame
- an array of dictionaries, with each dictionary corresponding to each object detected. Each dictionary contains 'name', 'percentage_probability' and 'box_points'
- a dictionary with keys being the name of each unique objects and value are the number of instances of each of the objects present
- If return_detected_frame is set to True, the numpy array of the detected frame will be parsed as the fourth value into the function

def forFrame(frame_number, output_array, output_count):
    print("FOR FRAME ", frame_number)
    print("Output for each object : ", output_array)
    print("Output count for unique objects : ", output_count)
    print("--------END OF A FRAME ---------")
```

— parameter per_second_function (optional): This parameter allows you to parse in the name of a function you define. Then, for every second of the video that is detected, the function will be parsed into the parameter will be executed and analytical data of the video will be parsed into the function. The data returned can be visualized or saved in a NoSQL database for future processing and visualization.

See a sample function for this parameter below .. code-block:

```python
This parameter allows you to parse in a function you will want to execute after
```
each second of the video is detected. If this parameter is set to a function, after every second of a video is detected, the function will be executed with the following values parsed into it:
-- position number of the second
-- an array of dictionaries whose keys are position number of each frame present in the last second, and the value for each key is the array for each frame that contains the dictionaries for each object detected in the frame
-- an array of dictionaries, with each dictionary corresponding to each frame in the past second, and the keys of each dictionary are the name of the number of unique objects detected in each frame, and the key values are the number of instances of the objects found in the frame
-- a dictionary with its keys being the name of each unique object detected throughout the past second, and the key values are the average number of instances of the object found in all the frames contained in the past second
-- If return_detected_frame is set to True, the numpy array of the detected frame will be parsed as the fifth value into the function

def forSeconds(second_number, output_arrays, count_arrays, average_output_count):
    print("SECOND : ", second_number)
    print("Array for the outputs of each frame ", output_arrays)
    print("Array for output count for unique objects in each frame :", count_arrays)
    print("Output average count for unique objects in the last second: ", average_output_count)
    print("------------END OF A SECOND --------------")

—parameter per_minute_function (optional): This parameter allows you to parse in the name of a function you define. Then, for every frame of the video that is detected, the function which was parsed into the parameter will be executed and analytical data of the video will be parsed into the function. The data returned has the same nature as the per_second_function; the difference is that it covers all the frames in the past 1 minute of the video.

See a sample function for this parameter below:

def forMinute(minute_number, output_arrays, count_arrays, average_output_count):
    print("MINUTE : ", minute_number)
    print("Array for the outputs of each frame ", output_arrays)
    print("Array for output count for unique objects in each frame :", count_arrays)
    print("Output average count for unique objects in the last minute: ", average_output_count)
    print("------------END OF A MINUTE --------------")

—parameter video_complete_function (optional): This parameter allows you to parse in the name of a function you define. Once all the frames in the video is fully detected, the function which was parsed into the parameter will be executed and analytical data of the video will be parsed into the function. The data returned has the same nature as the per_second_function and per_minute_function; the differences are that no index will be returned and it covers all the frames in the entire video.

2.5. Custom Object Detection: Training and Inference 41
See a sample function for this parameter below: .. code-block:

```python
def forFull(output_arrays, count_arrays, average_output_count):
    print("Array for the outputs of each frame ", output_arrays)
    print("Array for output count for unique objects in each frame :
        ", count_arrays)
    print("Output average count for unique objects in the entire video: ", average_output_count)
    print("--------------END OF THE VIDEO --------------")
```

---

**parameter detection_timeout** (optional): This function allows you to state the number of seconds of a video that should be detected after which the detection function stop processing the video.
CHAPTER 3

Indices and tables

- genindex
- modindex
- search