

输入输出基准

1、介绍

设计一组基于JSON格式的语义描述基准，该基准用来统一描述模型inference得到的结果，以便Ti-Brain统一调用。下面针对图像分类，图像检测，人脸检测与识别，OCR五种模型类别进行输入输出标准化。后续根据随着类型增加继续动态扩充。

2、图像分类模型

分类模型的输出结果可以包含以下key：每个分类结果的类别信息（category），及其对应的置信度（confidence）

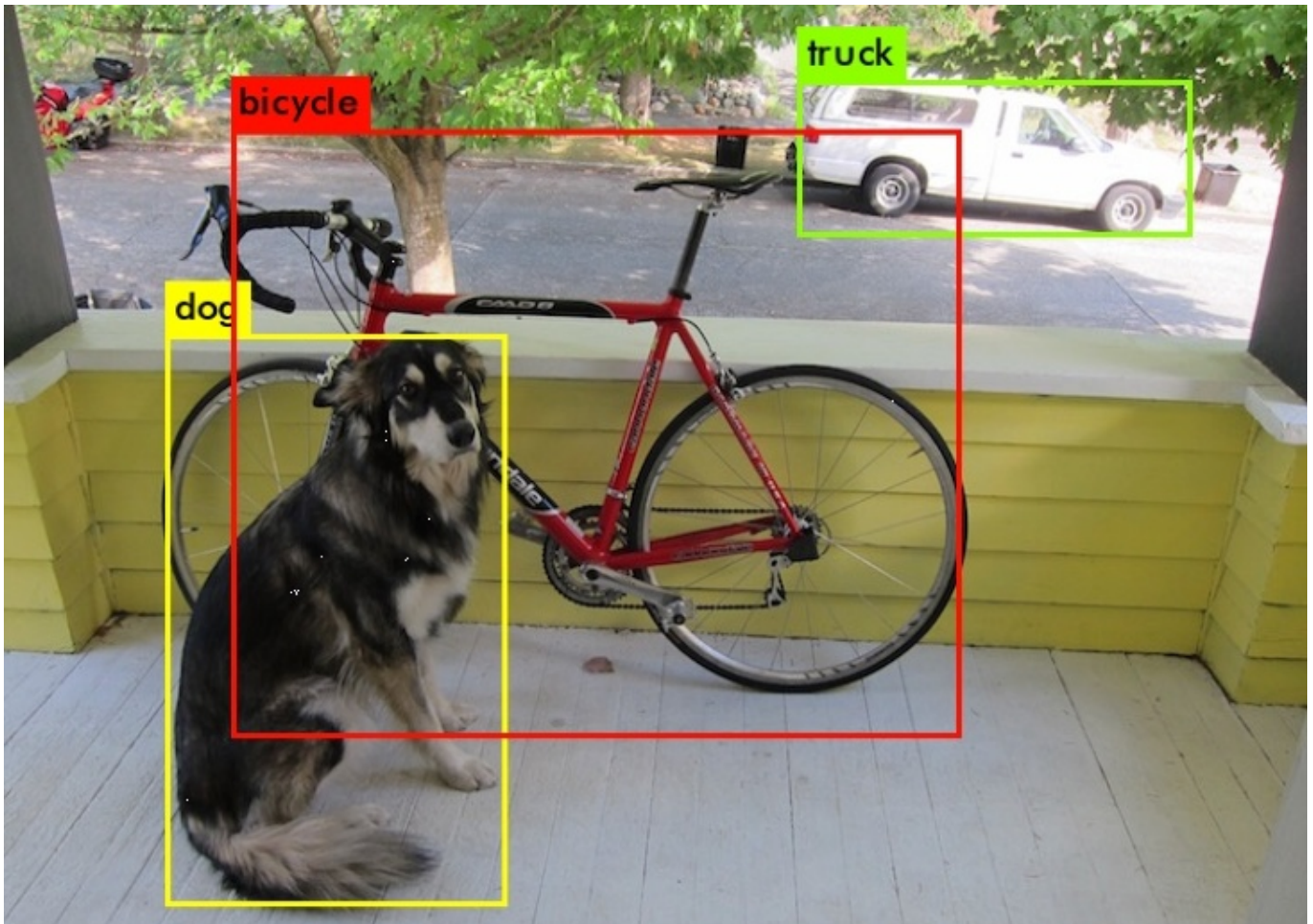


例如当输入图片为下图所示时，图像分类模型的输出为：

```
{ "cls_res": [
  { "category": "cat", "confidence": "0.9" },
  { "category": "tiger", "confidence": "0.08" },
  { "category": "lion", "confidence": "0.02" }
]
```

3、图像检测模型

检测模型的输出结果可以包含以下key：输入图像中所有物体的类别（category），置信度（confidence）及位置信息（bbox），其中bbox以具体左上角，右上角，左下角，右下角的四个精确坐标顺时针表示，例如（0,0），（100,0），（100,100），（0,100）。



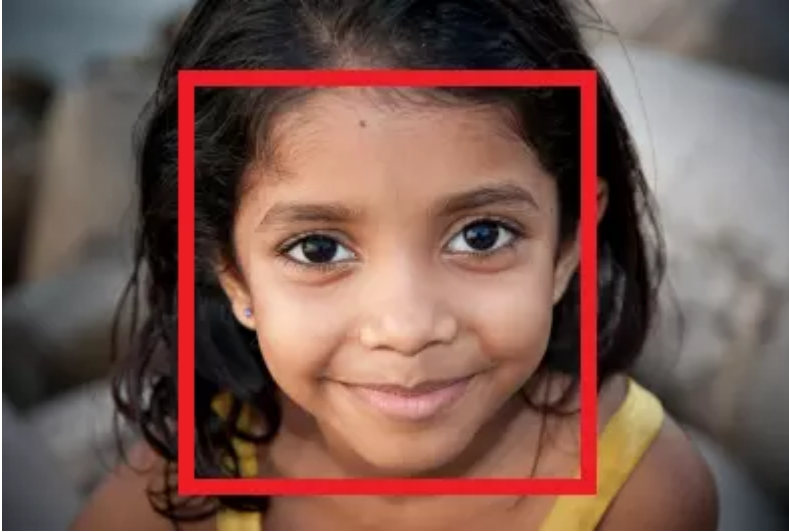
例如当输入图片为下图所示时，检测分类模型的输出为：

```
{ "det_res": [
  { "category": "dog", "confidence": "0.98", "bbox": "[ (20,40), (60,40), (60,120), (20,120)]" },
  { "category": "bicycle", "confidence": "0.88", "bbox": "[ (22,10), (120,10), (120,100), (22,100)]" },
  { "category": "truck", "confidence": "0.92", "bbox": "[ (110,5), (140,5), (140,15), (110,15)]" }
]
```

4、人脸检测与识别模型

4.1 人脸检测

人脸检测模型的输出结果可以包含以下key：输入图像中所有人脸的置信度（confidence）及位置信息（bbox），其中bbox以具体左上角，右上角，左下角，右下角的四个精确坐标顺时针表示，例如（0,0），（100,0），（100,100），（0,100）。



例如当输入图片为下图所示时，检测分类模型的输出为：

```
{ "face_det_res": [
  { "confidence": "0.98", "bbox": "[ (20,40), (60,40), (60,120), (20,120) ]" }
]
```

4.2 人脸识别

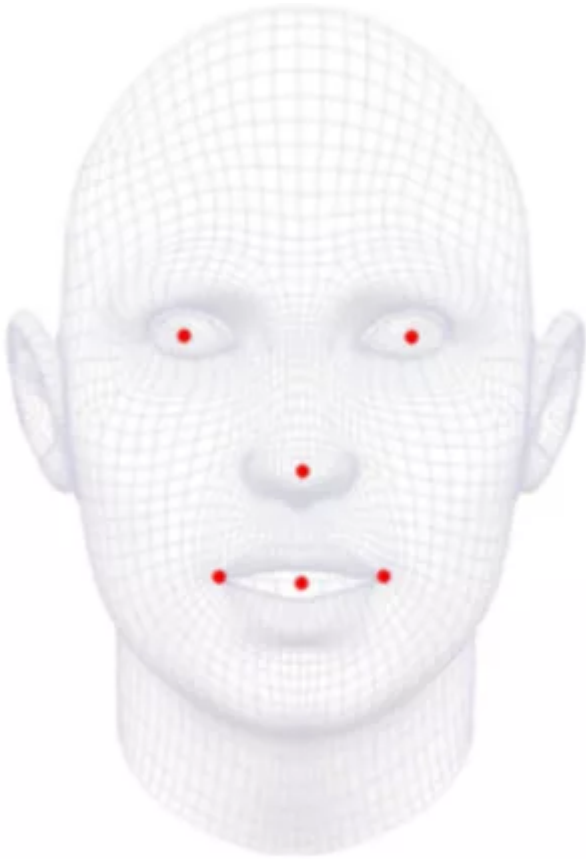
人脸识别模型接受一张人脸源图像，一张待对比的人脸图像作为输入，输出结果为两张人脸的特征距离

示例：

```
{ "face_rec_res": [
  { "feature_distance": "0.38" }
]
```

4.3 人脸关键点检测

人脸关键点检测模型的输出结果可以包含以下key：输入图像中所有人脸关键点的名称或序号（name），置信度（confidence）及位置信息（locat），其中locat以x, y坐标表示。



(g) AFW

例如当输入图片为下图所示时，检测分类模型的输出为：

```
{ "face_rec_res": [
  { "name": "left_eye", "confidence": "0.96", "locat": "[100,200]" },
  { "name": "right_eye", "confidence": "0.91", "locat": "[200,200]" },
  { "name": "nose", "confidence": "0.96", "locat": "[200,150]" },
  { "name": "left_mouth", "confidence": "0.92", "locat": "[110,300]" },
  { "name": "mid_mouth", "confidence": "0.91", "locat": "[150,310]" },
  { "name": "right_mouth", "confidence": "0.96", "locat": "[200,300]" },
]
}
```

5、OCR模型

5.1、文本检测

文字检测模型的输出结果可以包含以下key：输入图像中所有文字的置信度（confidence）及位置信息（bbox），其中 bbox以具体左上角，右上角，左下角，右下角的四个精确坐标顺时针表示，例如（0,0），（100,0），（100,100），（0,100）。

示例：

```
{ "character_det_res": [
  { "confidence": "0.98", "bbox": "[ (20,40), (60,40), (60,120), (20,120) ]" }
]
```

5.2、文本识别

文字识别模型接受一张文字图片输入，输出结果为文本识别出的字符信息。



示例：

```
{ "character_rec_res": [
  { "character_info": "google" }
]
```