

Report Name: Data Annotation and Dataset Preparation

Objective: The goal of this task was to accurately annotate images and prepare a dataset suitable for training a computer vision model.

Annotation Process

Tools Used:

- Websites for image collection: Flickr, Google, Unsplash, Pexels, Pascal Visual Object Classes.
- Annotation tool: LabelImg (set to Pascal VOC mode).

Steps Taken:

1. The LabelImg tool was set to Pascal VOC mode for annotation.
2. A rectangle box was drawn around each object, and the appropriate label was assigned.
3. Default file names were kept unchanged.
4. This process was repeated for all images.

Precautions Taken:

- High-definition images with a minimum resolution of 500×500 pixels were used.
- Blurry images were avoided.
- Images with cluttered backgrounds were not selected.
- Bounding boxes were carefully drawn to cover the object properly.

Extracting Annotation Data from XML Files

Steps Taken:

1. A list of XML files was generated in Python.
2. XML files were parsed using an XML parsing library.
3. Data extracted from each XML file:
 - Filename
 - Image dimensions (width, height)
 - Object details (name, xmin, xmax, ymin, ymax)
4. The extracted data was stored in a structured format (list or dictionary).

Conversion to DataFrame:

- The data was flattened for easy processing.
- A Pandas DataFrame was created with the following columns:
 - filename
 - width
 - height
 - object_name
 - xmin
 - xmax
 - ymin
 - ymax
- The dataset contained 50,000 rows and 8 columns.
- Common objects identified: Persons (5,447), Cars (1,650).

Conversion to YOLO Format

For YOLO, bounding box coordinates were converted to the following format:

class_id center_x center_y width height

Formulas Used:

- Center X = $(xmin + xmax) / 2$ (normalized by image width)
- Center Y = $(ymin + ymax) / 2$ (normalized by image height)
- Width = $(xmax - xmin) / image\ width$
- Height = $(ymax - ymin) / image\ height$

Each image had a corresponding .txt file containing these labels.

Dataset Organization

The dataset was structured as follows:

```
data/
├── images/
│   ├── train/ (Training images)
│   └── test/  (Testing images)
└── labels/
    ├── train/ (Labels for training)
    └── test/  (Labels for testing)
```

Data Split:

- 80% of the images were assigned to the training set.
- 20% were used for testing.
- Labels were placed in corresponding folders.

Challenges Faced and Solutions

1. Labelling Compatibility Issue:
 - Labelling did not support the latest Python version.
 - Python 3.9.6 was installed, and venv was used to create a virtual environment.
2. Handling Multiple Objects in XML Files:
 - Extracted data was structured properly to include multiple objects from the same image.
3. Data Integrity:
 - Ensured that all images had corresponding annotation files.

Conclusion

The dataset was successfully prepared, ensuring that all images were annotated accurately. The extracted annotations were converted into both Pascal VOC and YOLO formats, and the dataset was split for training and testing. This structured dataset will be used for training the computer vision model.

Reference:

An AI assistant was utilized for paraphrasing, grammar, and sentence structure refinement..