**Data Labeling with CVAT**

Before the steps below were conducted, an account for CVAT was set up, and a project for testing image labeling was created.

<https://cvat.org/projects?page=1>

<https://blog.roboflow.com/cvat/>

|  | Task Set Up  Click on the blue button down the bottom, fill out the information and drag in an image from the computer.  Set up the different labels in this menu, all labeled images down the bottom as pictured here. |
| --- | --- |
|  | Job Data  Job has been created within the test project (using an image from work).  Once submit is clicked then the menu pictured is shown.  Not very clear to click on the job link at the bottom to edit, but good UI to see the information about that job/image to label |
|  | Drawing Screen  Pictured is the main menu for drawing on the image. |
|  | Labining Drawing  Able to use custom shapes and rectangles/polygons to draw on the image for each label, as shown here using a rectangle to highlight the hi-vis the person is wearing in this image. |
|  | Final Labels  Each box is shown in different colours. Allows editing of boxes, easy to view, and nicely laid out. |
|  | Another Image  Labeling for a face mask. |
|  | Exporting DataSet  It was easy to save and export the data set (with only one test image) to an obj.zip to use for model training.  CVAT allowed for downloading in multiple different model training formats which could be handy if we ever need to switch from YOLO.  The ease of this will make adding custom data to increase the accuracy of the model much easier. |

**OverView of CVAT:**

Very easy to navigate and use once the project was set up and the labels created (these steps were completed with the help of the tutorial).

This process was very easy and NOT time-consuming, will make labeling custom images for the model easy, and exporting of the data pretty simple.

**Data Labeling with DataGym**

<https://www.datagym.ai/>

|  | Downloaded the docker environment to run the data gym ai.  Cloned the data gym ai from the Github repo.  Not as clear to set up if getting multiple people in the team to use, takes a lot longer and requires more setting up initially  Have to change the DATAGYM\_DEACTIVATE-DBCHECKBLOCKER: false  To  DATAGYM\_DEACTIVATE-DBCHECKBLOCKER: It 'false'  Then navigated it to localhost:8080 it after installation It, it and the data gym UI was then available. |
| --- | --- |
|  | Home page |
|  | Creating labels  Able to set the shape of the label during the creation phase, however as found during later steps that this was not able to be edited/changed during the labeling phase of the images.  The creation of labels besides this was easy and straightforward. |
|  | Labeling Images  The image labeling UI was a little more user-friendly than CVAT, but as mentioned earlier above, not being able to change the labels was annoying.  It also was not clear as to how to submit the image, however clicking on a layer that was not being used and canceling it made the submit button appear, an extra step required which was not clear. |
|  | Released the tasks but not clear on what this does and how to download the data sets, menus screenshotted appeared.  Was not able to find out how to export to a YOLOv4 format after some searching. |

**OverView of DataGym AI**

CVAT better - easier to use, can change the shapes for the labels which is very helpful, and the setup required for CVAT is much easier compared to Data Gym if others in the team needed also to do data labeling.

The user interface looks visually appealing and is easy to navigate, however, it is not clear how some functions work. As this is a repo that is still in development, it is clear that some functionality is still to be added to the application to make it more user-friendly.

**Google Colab Data Labeling - LabelBox**

[**https://docs.labelbox.com/docs/create-an-api-key**](https://docs.labelbox.com/docs/create-an-api-key)

[**https://colab.research.google.com/github/Labelbox/labelbox-python/blob/develop/examples/basics/labels.ipynb#scrollTo=2969df72**](https://colab.research.google.com/github/Labelbox/labelbox-python/blob/develop/examples/basics/labels.ipynb#scrollTo=2969df72)

[**https://app.labelbox.com/projects/cl3gkr4gzfp1m087rdfknb95v/labels/activity**](https://app.labelbox.com/projects/cl3gkr4gzfp1m087rdfknb95v/labels/activity)

|  | **Creating API Key**  Created an account for LableBox and created an API as follows [**https://docs.labelbox.com/docs/create-an-api-key**](https://docs.labelbox.com/docs/create-an-api-key).  This API was created for the account and also for the purposes of connecting to google collab as seen below.  API was created for the test project I set up.  TO NOTE\* when creating the project to test, was able to set the type of media to be tagged and allowed for both images and videos - perfect for our project. |
| --- | --- |
|  | **Labeling**  Lables in the label box are created in the schema tab under the object title.  The labels are created and then can be used in the datasets when tagging an image.  For labeling, all the objects are done just with points, which allows for accurate tagging of images to the area wanting to be detected.  Overall very easy to label, and being able to set the labels beforehand is very helpful. |
|  | **After Saving Labeled Image**  Was happy to see a menu after saving the tagging of the two images in the dataset.  When we will need to tag much larger sets, then this menu will be very insightful and helpful. |
|  | **Google Collab**  The purpose of looking into LableBox was its ability to connect with google collab.  Was also able to connect with GitHub. Was easy to set up and follow the tutorial to connect with my project and give a review to a tagged image |
|  | **Google Collab - Reviewed Image**  Proof of connection with google collab, the image has been reviewed. |
|  | **Visualizing the image in the pipeline**  Using !pip commands to visualize one of the tagged images in google collab |
|  | **Exporting Lables**  Able to export the data file in the command line of google collab!  However, can only export with JSON which is a small downfall. |
|  | **Creating Model for Exporting DataSets**  To export the data model for model training outside of LabelBox, a model has to be created to export from, done in the SDK.  After some more reading, realized that you can just generate the report, and multiple different options are presented for the download, including the dataset. |

**OverView of LableBox**

Overall, I think that LableBoc will be the best fit for our team, with its ease of use and connectivity to google collab which is where the YOLO model is being trained.

The user interface is very easy to navigate, and the method of labeling mentioned above makes the labeling of people and the clothes they are wearing more accurate, and easy to label for our team. Being able to export in google collab is also a big bonus.

Overall, i think LabelBox is a great data labeling tool. It is compatible with Github and google collab which are both tools that we are using for our project. The user interface is very easy to navigate, and the method of labeling mentioned above makes the labeling of people and the clothes they are wearing more accurate, and easy to label for our team. Being able to export in google collab is also a big bonus.

However, after some time researching and trying to find options to export the labeled data sets to a YOLOv4 format, I was not able to uncover how to do this.

**Data Tool To Use: CVAT**

CVAT was very easy to create an account, log in, and label data without much setup and is very easy to navigate and use. The client also uses CVAT, and has cloud storage for this tool which was expressed during our client meeting (17/05/2022) that we would be able to use.

In the future, if a little more upskilling is put into LabelBox further into the project, then this could be quite helpful with the google collab and Github integration.