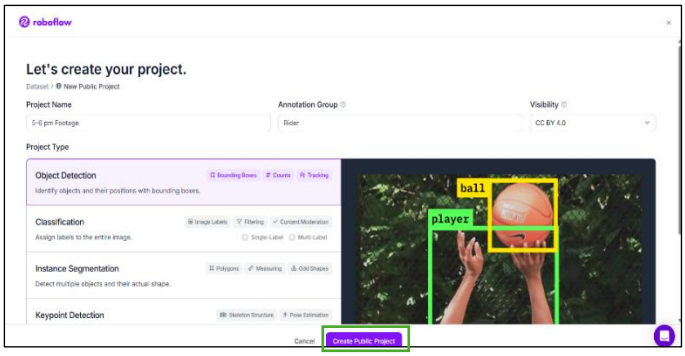
Group 5: Nilo, Daniella Kim N.
Peñaflor, Rowencell A.

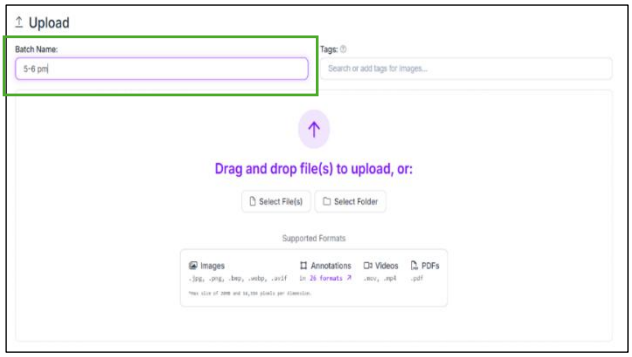
Course/Section: CSS182-3/FOPI01

M2-FA2: Data Pre-Processing

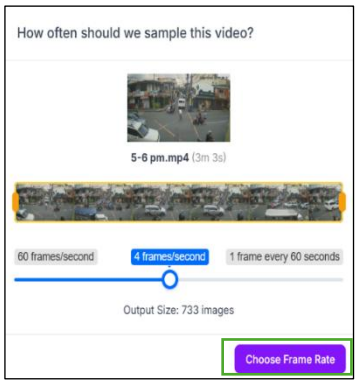
ANNOTATING 5-6 PM IMAGES

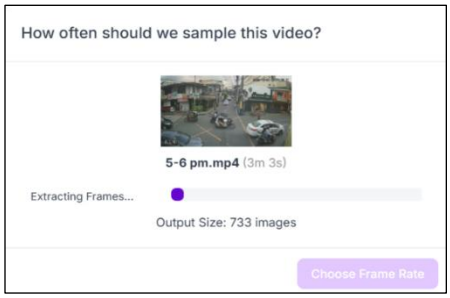
In this activity, the first step we took was to access the `RoboFlow` site using the link: `https://app.roboflow.com/dataset-altip`. In the `Dataset Workplace`, we created a new project by clicking the `New Project` button and named it `5-6 pm Footage` and `Rider` for the `Annotation Group` for the first video. Once done, we clicked the `Create Public Project` button, renamed the `Batch Name to 5-6 pm`, and dragged the first video in the `Upload page` to start processing the video.

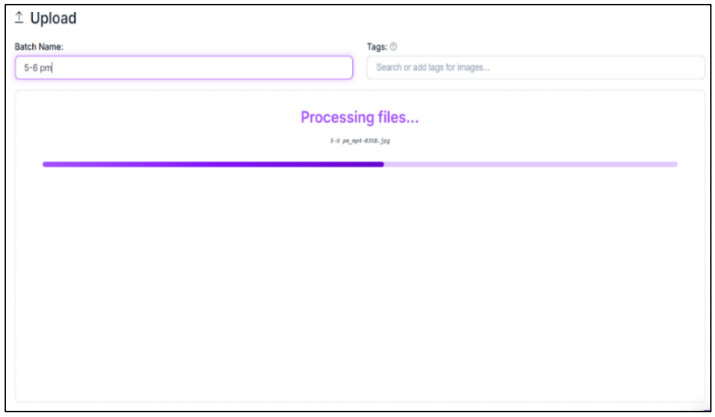


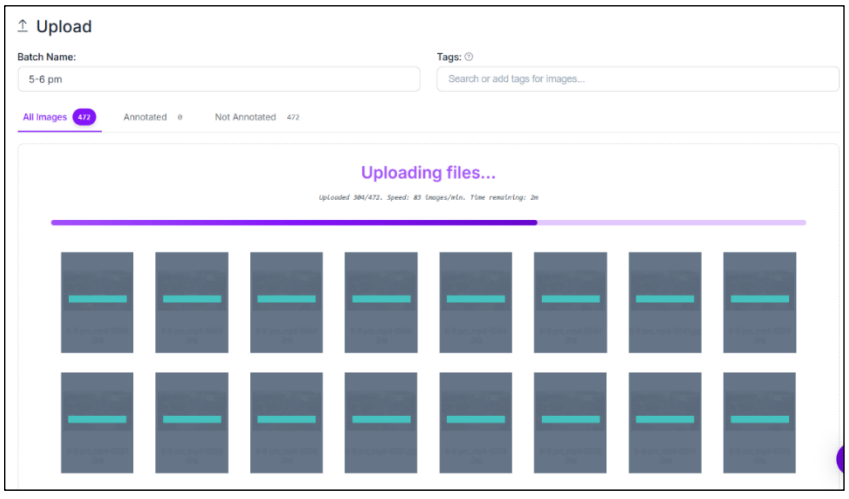


After the scanning of the file is finished, we chose a frame rate of `4 frames/second` to create an `Output Size of 733 images`. Upon confirming, the extraction of frames started, and the file was processed. Subsequently, the files were uploaded, and the images were reduced to `472 images` because of frame duplication.

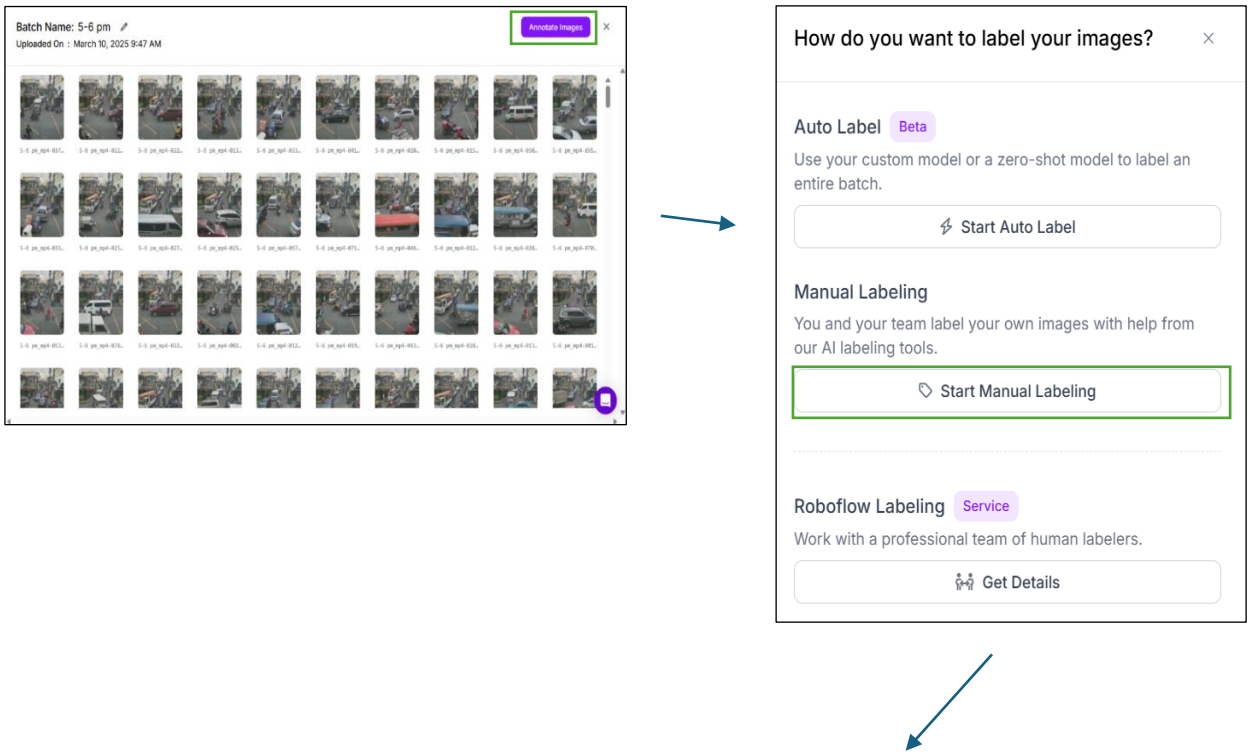


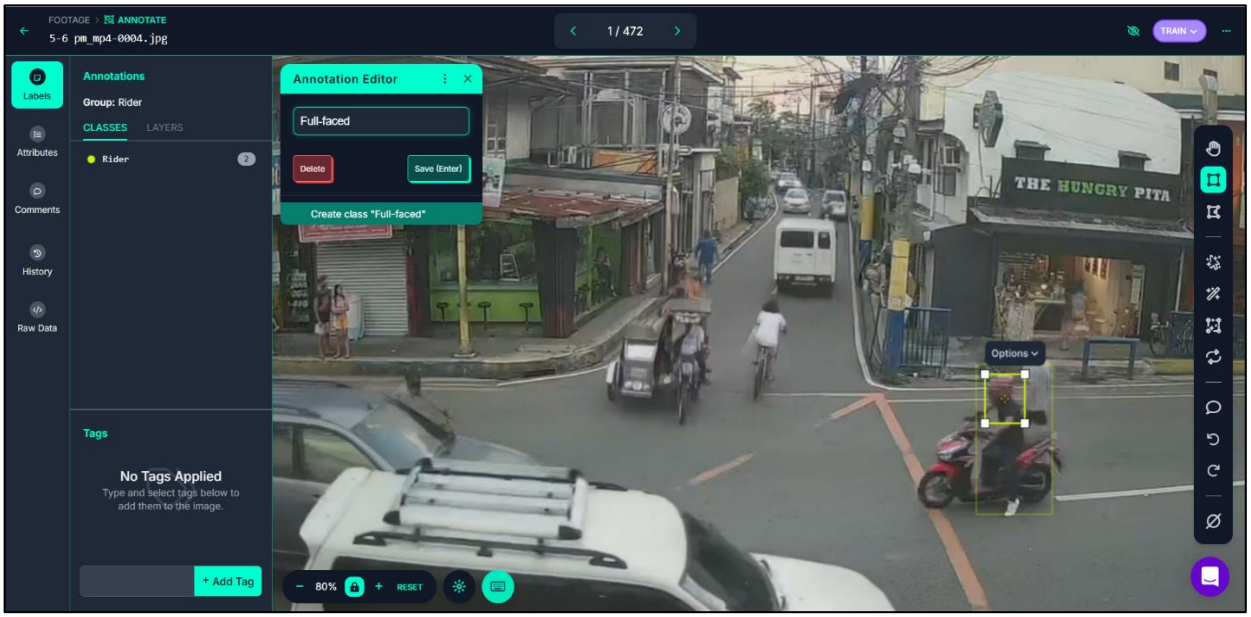






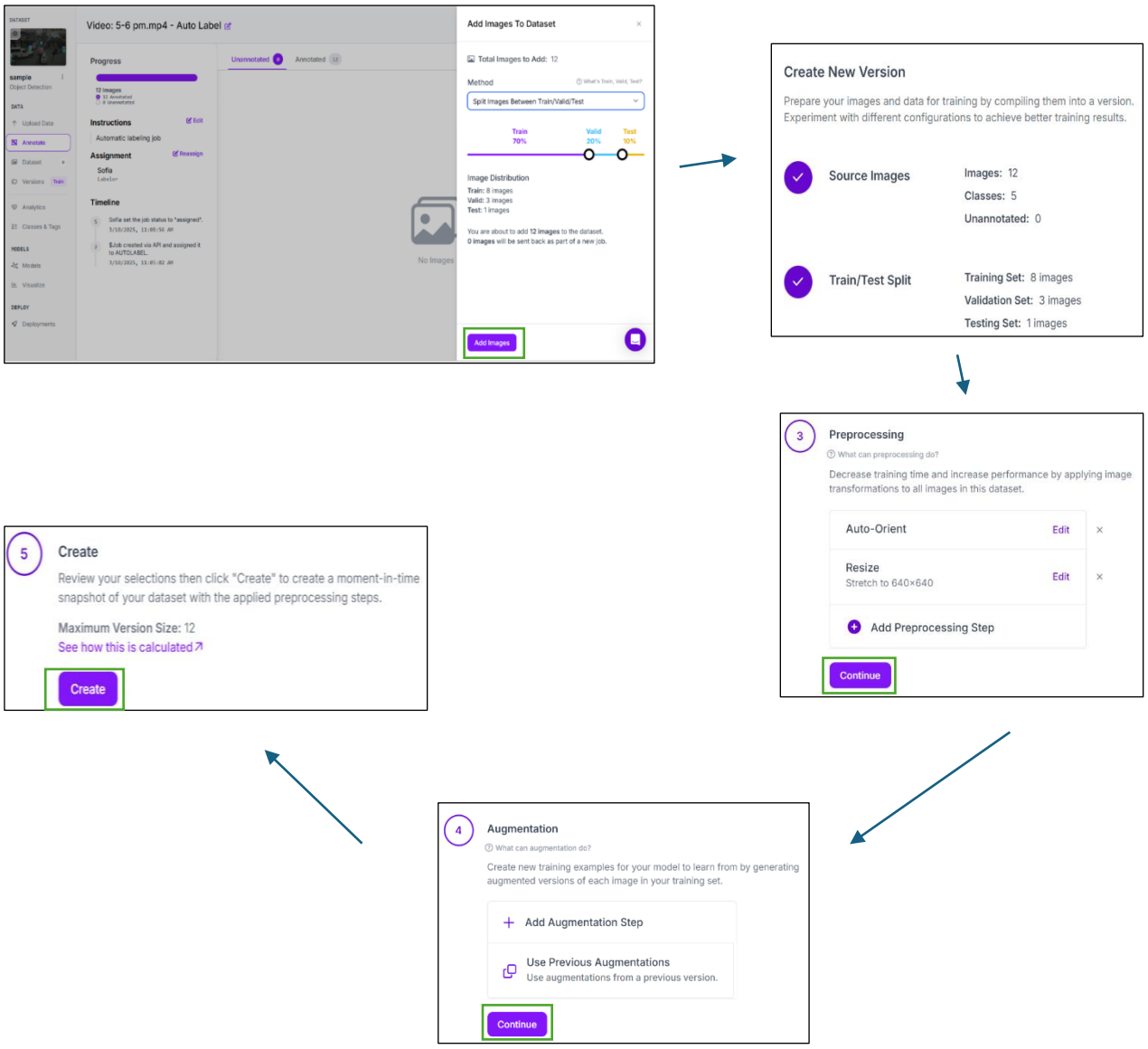
Once done, we started annotating the images by manually labeling them. We created the 5 classes: `Rider, Full-Faced, Half-Faced, Invalid, and Not Wearing Helmet` to classify the motorcycle riders and the type of helmet they are wearing in the image.

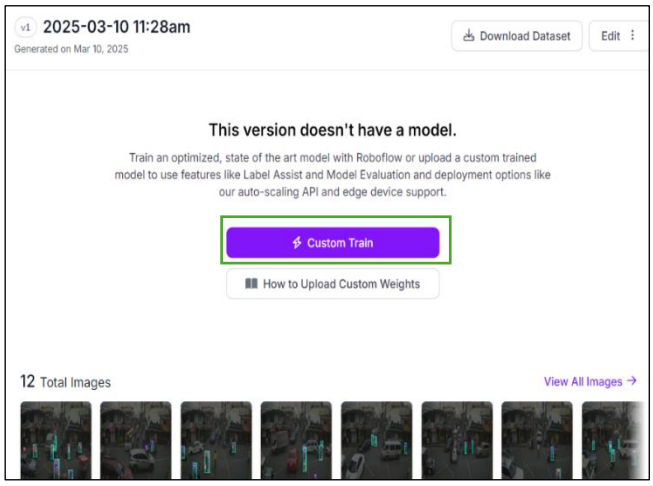


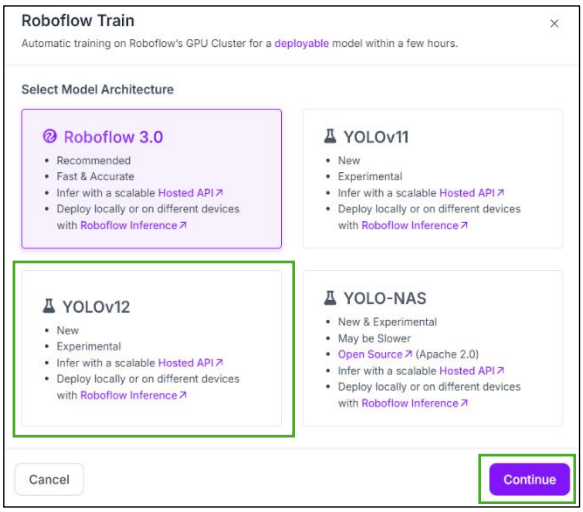


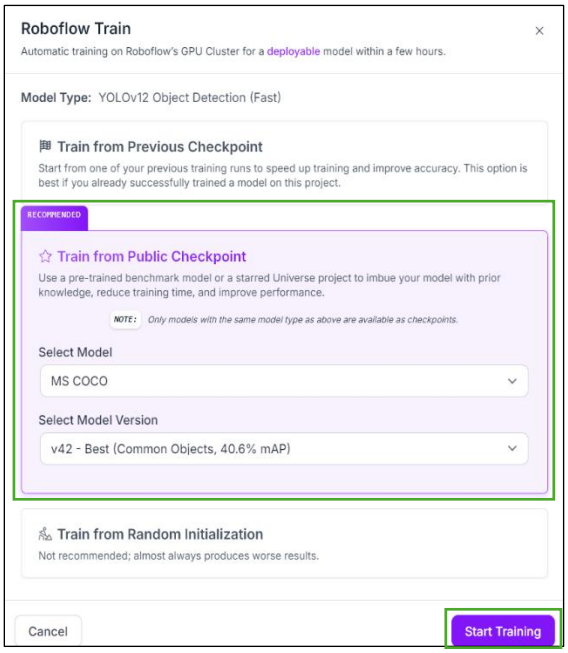
CREATING AND APPLYING LABEL ASSIST

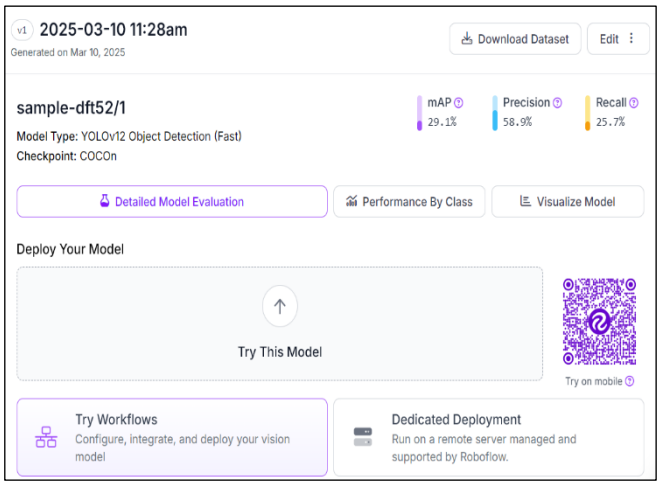
To accelerate the annotating process, we tried the `Label Assist` feature, but before using the feature, we first created a new dataset containing 12 images from the video, annotated it, and trained it using `YOLOv12`. Afterwards, we applied the `Label Assist` feature on our original dataset and selected the first version of the project sample as our model. Then we selected all the classes in the dataset and used it for labeling. The `Label Assist` automatically detects and annotates the objects in the images, but we checked them and properly fixed their position and classes before moving on to the next image.

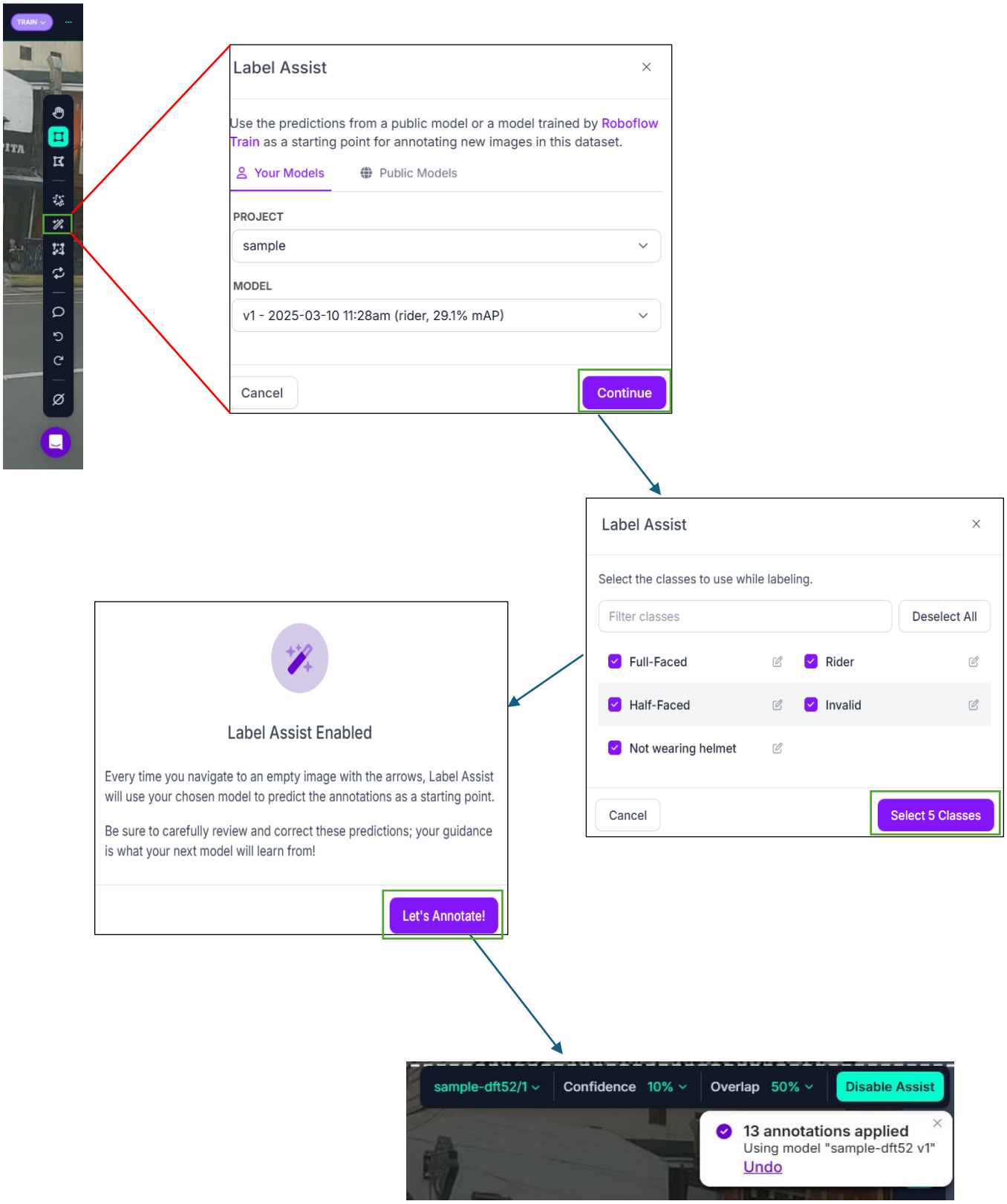



v1
2025-03-10 11:28am
Generated on Mar 10, 2025
Download Dataset
Edit
This version doesn't have a model.
Train an optimized, state of the art model with Roboflow or upload a custom trained model to use features like Label Assist and Model Evaluation and deployment options like our auto-scaling API and edge device support.
Custom Train
How to Upload Custom Weights
12 Total Images
View All Images


Roboflow Train
Automatic training on Roboflow's GPU Cluster for a deployable model within a few hours.
Select Model Architecture
Roboflow 3.0
Recommended
Fast & Accurate
Infer with a scalable Hosted API
Deploy locally or on different devices with Roboflow Inference
YOLOv11
New
Experimental
Infer with a scalable Hosted API
Deploy locally or on different devices with Roboflow Inference
YOLOv12
New
Experimental
Infer with a scalable Hosted API
Deploy locally or on different devices with Roboflow Inference
YOLO-NAS
New & Experimental
May be Slower
Open Source (Apache 2.0)
Infer with a scalable Hosted API
Deploy locally or on different devices with Roboflow Inference
Cancel
Continue

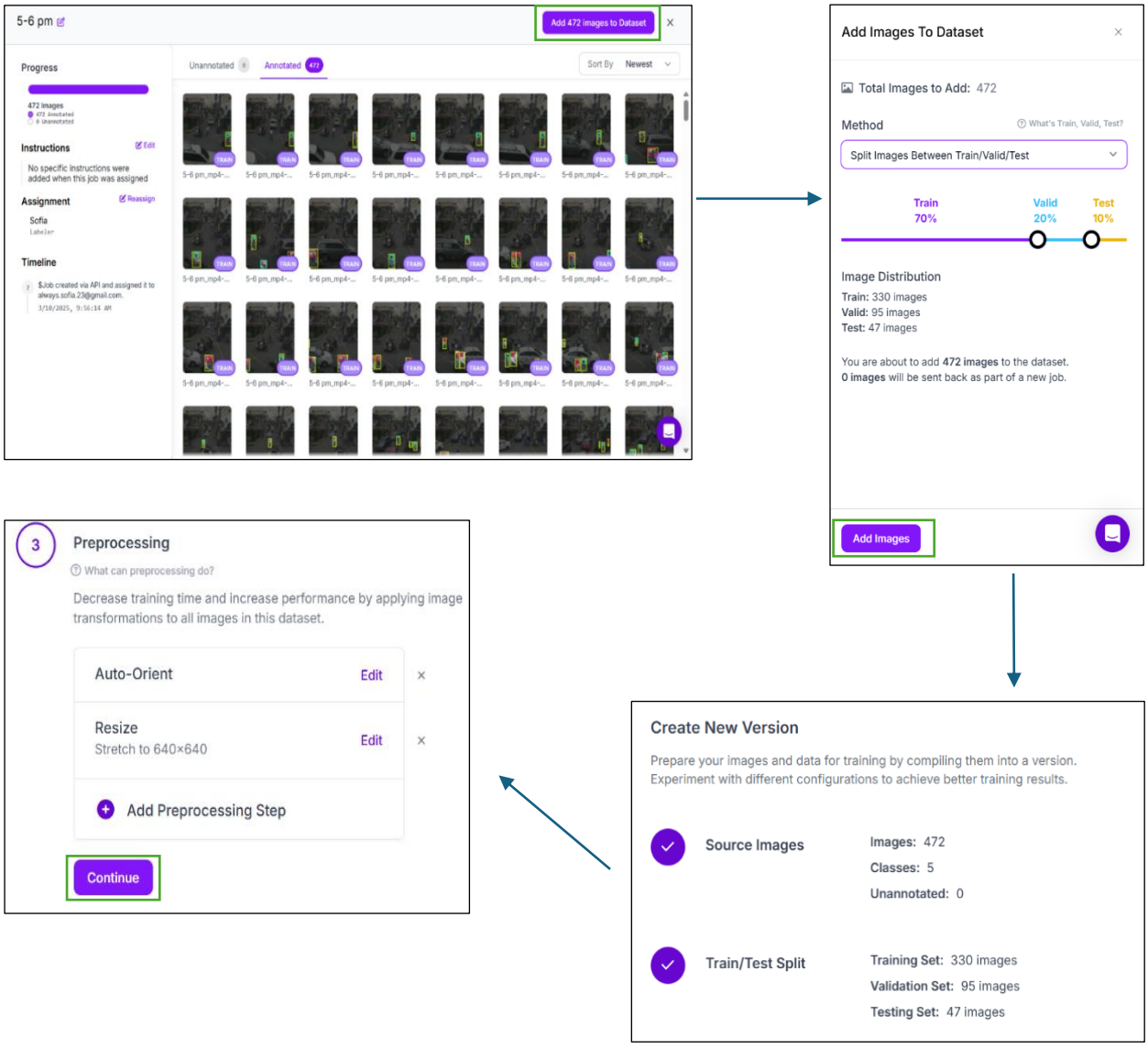

Roboflow Train
Automatic training on Roboflow's GPU Cluster for a deployable model within a few hours.
Model Type: YOLOv12 Object Detection (Fast)
Train from Previous Checkpoint
Start from one of your previous training runs to speed up training and improve accuracy. This option is best if you already successfully trained a model on this project.
RECOMMENDED
Train from Public Checkpoint
Use a pre-trained benchmark model or a starred Universe project to imbue your model with prior knowledge, reduce training time, and improve performance.
NOTE: Only models with the same model type as above are available as checkpoints.
Select Model
MS COCO
Select Model Version
v42 - Best (Common Objects, 40.6% mAP)
Train from Random Initialization
Not recommended; almost always produces worse results.
Cancel
Start Training

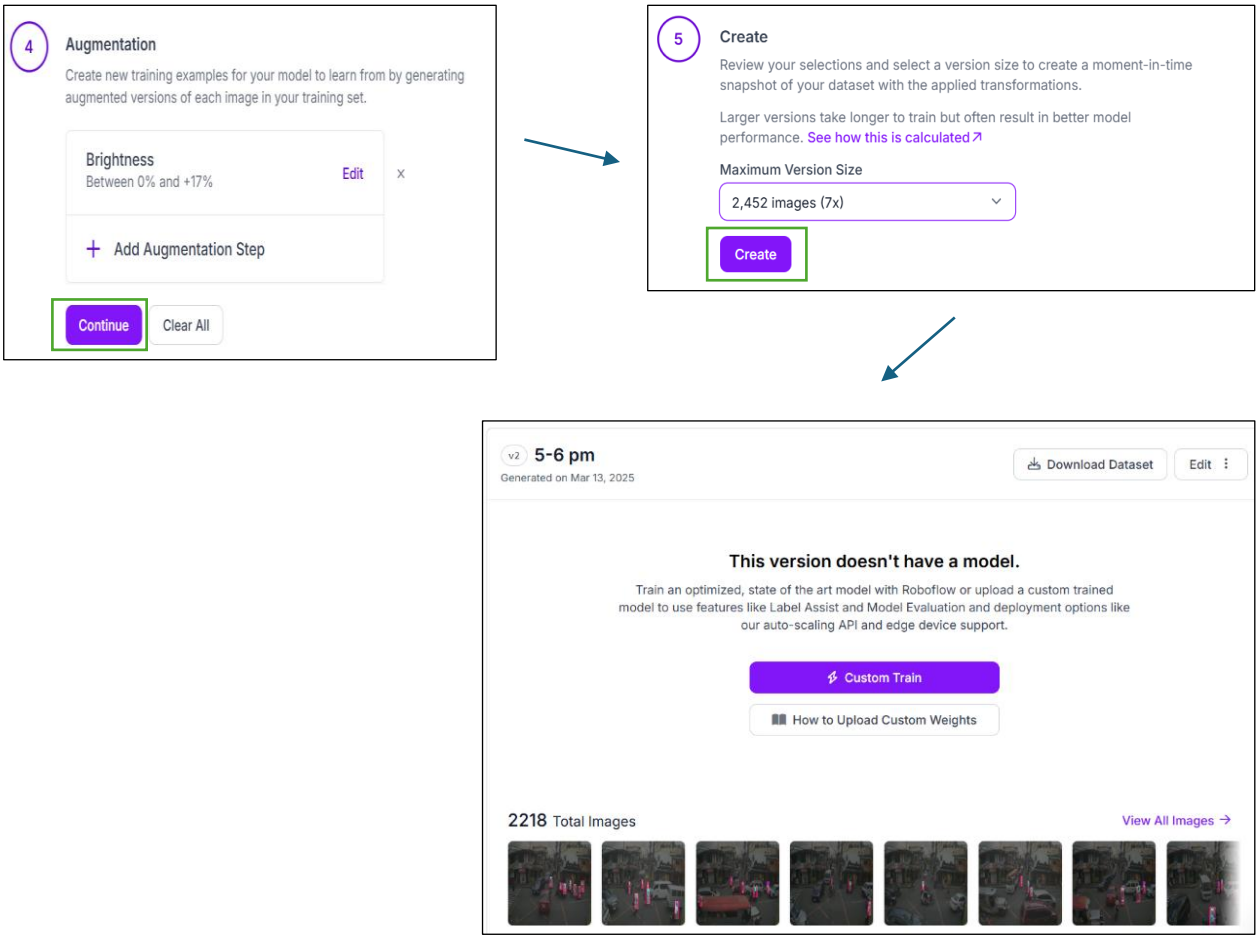

v1
2025-03-10 11:28am
Generated on Mar 10, 2025
Download Dataset
Edit
sample-dft52/1
Model Type: YOLOv12 Object Detection (Fast)
Checkpoint: COCOn
mAP 29.1%
Precision 58.9%
Recall 25.7%
Detailed Model Evaluation
Performance By Class
Visualize Model
Deploy Your Model
Try This Model
Try on mobile
Try Workflows
Configure, integrate, and deploy your vision model
Dedicated Deployment
Run on a remote server managed and supported by Roboflow.


Label Assist
Use the predictions from a public model or a model trained by Roboflow Train as a starting point for annotating new images in this dataset.
Your Models
Public Models
PROJECT
sample
MODEL
v1 - 2025-03-10 11:28am (rider, 29.1% mAP)
Cancel
Continue
Label Assist
Select the classes to use while labeling.
Filter classes
Deselect All
Full-Faced
Rider
Half-Faced
Invalid
Not wearing helmet
Cancel
Select 5 Classes
Label Assist Enabled
Every time you navigate to an empty image with the arrows, Label Assist will use your chosen model to predict the annotations as a starting point.
Be sure to carefully review and correct these predictions; your guidance is what your next model will learn from!
Let's Annotate!
sample-dft52/1
Confidence 10%
Overlap 50%
Disable Assist
13 annotations applied
Using model "sample-dft52 v1"
Undo

CREATING THE 5-6 PM DATASET

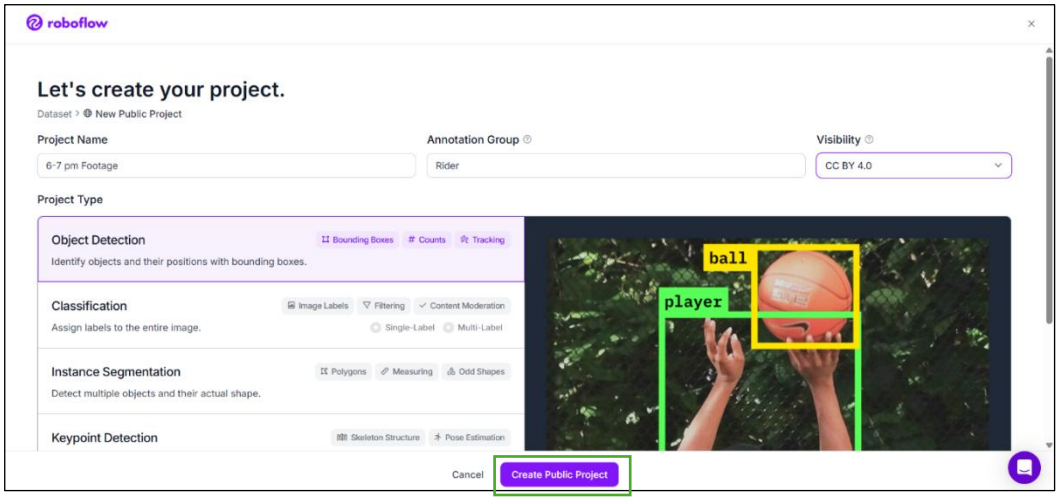
After we finished annotating the images, we added them to the dataset by clicking the `Add 472 Images to the Dataset` button and split the images between `Train, Valid, and Test`. Next, we `add preprocessing techniques` such as `Auto-Orient` and `Resize` to apply image transformation and increase their performance, and `Brightness` as for `augmentation`. We selected the `Maximum Version Size` of `2,452 images` and clicked the `Create` button to start generating the dataset. The finished dataset contains `2,218 images`.

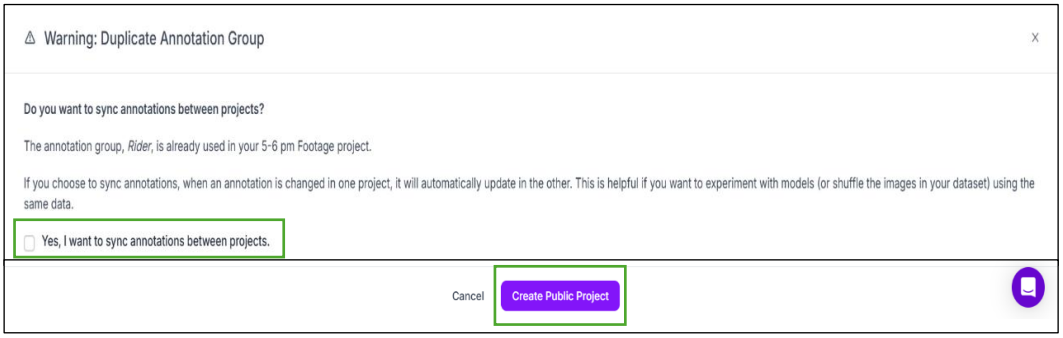


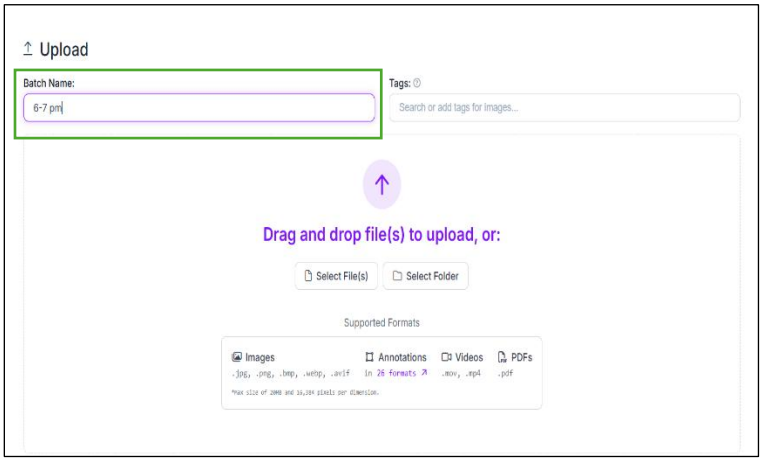


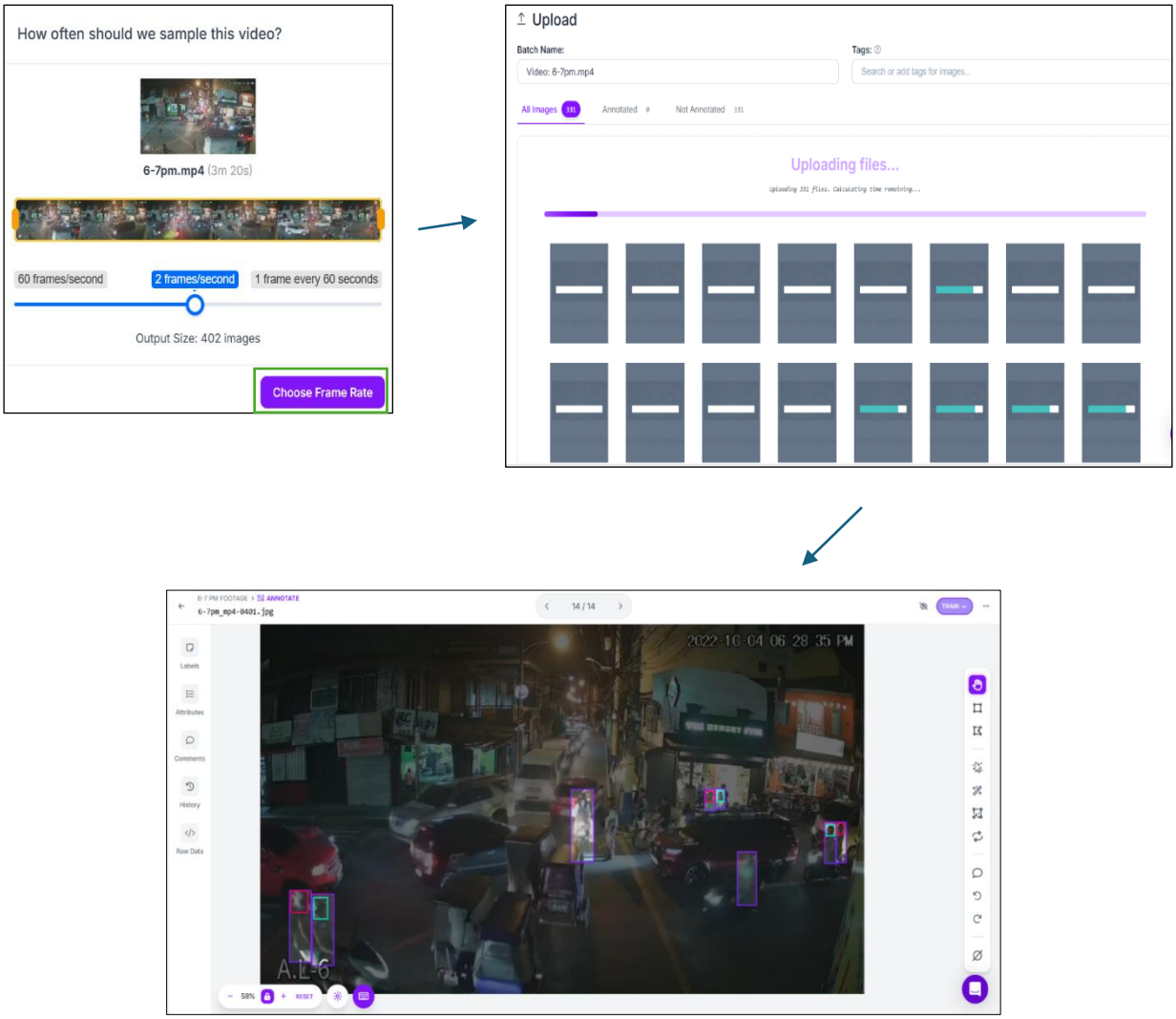
ANNONATING 6-7 PM IMAGES

After creating the `5-6 pm dataset`, we moved on to annotating the `6-7 pm images`. For this dataset, we created a `New Project` and named it `6-7 pm Footage` with `Rider` as its `Annotation Group`. We did not sync the classes with the previous dataset, so the previous dataset would not be affected if there are changes in labels within this dataset. Next, we named the `Batch Name` as `6-7 pm` and dragged the second video to start processing it. Once the scanning is done, we chose a `frame rate of 2 frames/seconds` to create an `Output Size of 402 images`, and after the extraction were reduced to `331 images`. Then, we started annotating the images with the help of `Label Assist`.

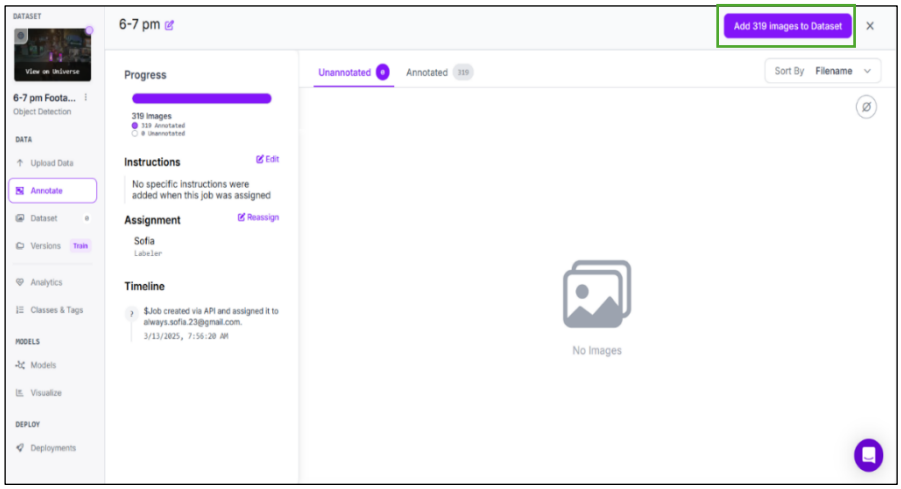

roboflow
Let's create your project.
Dataset > New Public Project
Project Name
6-7 pm Footage
Annotation Group
Rider
Visibility
CC BY 4.0
Project Type
Object Detection
Bounding Boxes
Counts
Tracking
Identify objects and their positions with bounding boxes.
Classification
Image Labels
Filtering
Content Moderation
Assign labels to the entire image.
Single-Label
Multi-Label
Instance Segmentation
Polygons
Measuring
Odd Shapes
Detect multiple objects and their actual shape.
Keypoint Detection
Skeleton Structure
Pose Estimation
ball
player
Cancel
Create Public Project

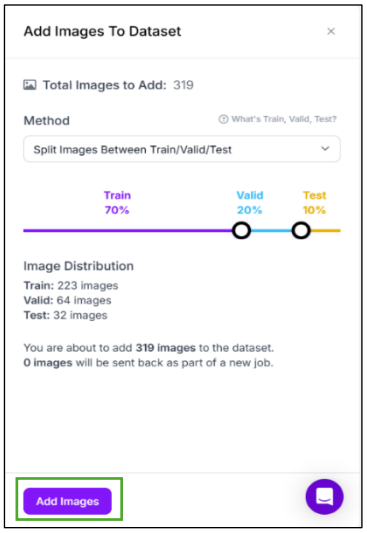

Warning: Duplicate Annotation Group
Do you want to sync annotations between projects?
The annotation group, Rider, is already used in your 5-6 pm Footage project.
If you choose to sync annotations, when an annotation is changed in one project, it will automatically update in the other. This is helpful if you want to experiment with models (or shuffle the images in your dataset) using the same data.
Yes, I want to sync annotations between projects.
Cancel
Create Public Project

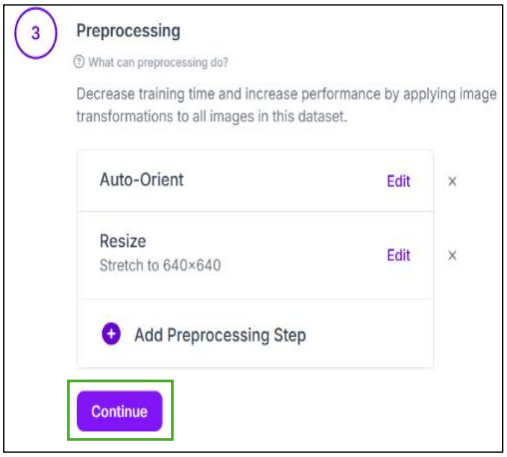

Upload
Batch Name:
6-7 pm
Tags:
Search or add tags for images...
Drag and drop file(s) to upload, or:
Select File(s)
Select Folder
Supported Formats
Images
Annotations
Videos
PDFs

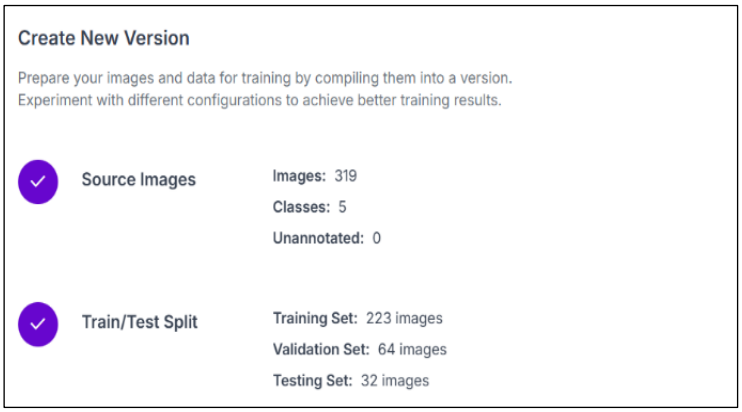


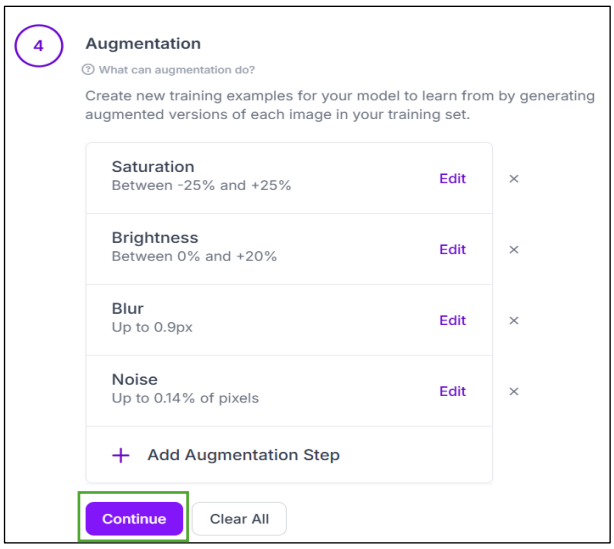
After annotating the images, similar to the previous dataset, we add them to the dataset by clicking the `Add 319 images to the Dataset` button and split the images between `Train, Valid, and Test`. We used the same `pre-processing techniques`, but for the `augmentation`, we used `Saturation, Brightness, Blur, and Noise`. Then, we selected the `Maximum Version Size of 2,326 images` and clicked the `Create` button to start generating the dataset. The finished dataset contains `2,326 images`.

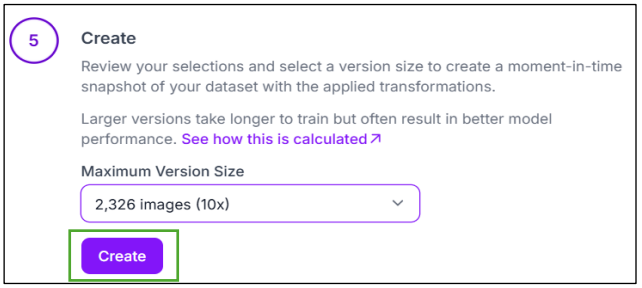

6-7 pm
Add 319 images to Dataset
Progress
319 images
319 Annotated
0 Unannotated
Instructions
Edit
No specific instructions were added when this job was assigned
Assignment
Reassign
Sofia
Labeler
Timeline
$Job created via API and assigned it to always.sofia.23@gmail.com.
3/13/2025, 7:56:20 AM
Unannotated 0
Annotated 319
Sort By Filename
No Images

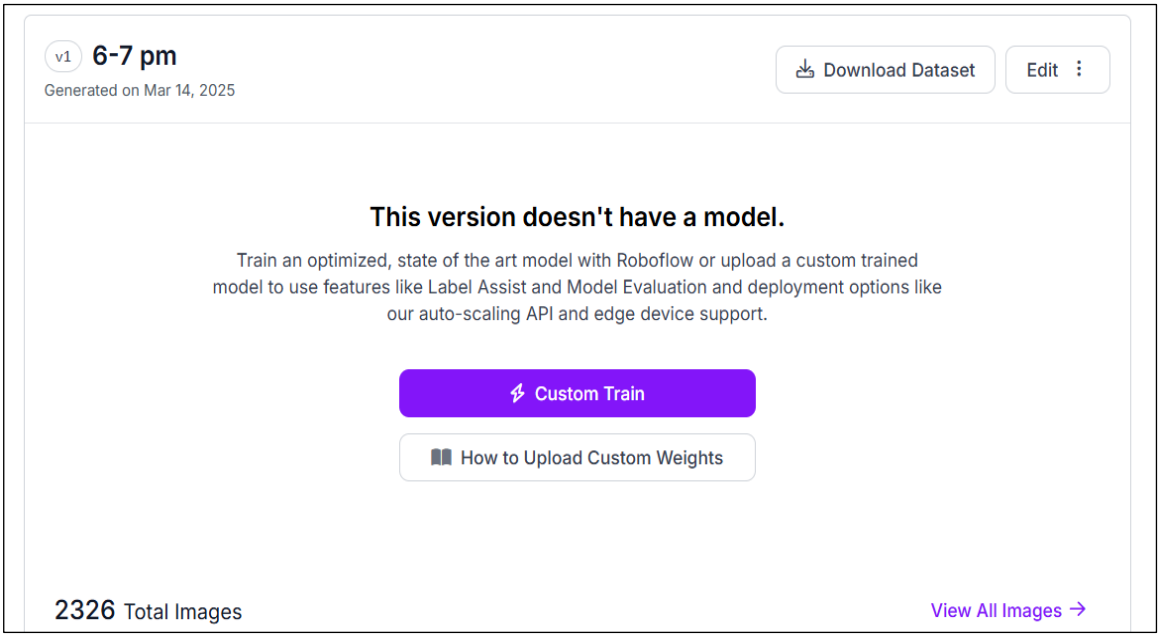

Add Images To Dataset
Total Images to Add: 319
Method
What's Train, Valid, Test?
Split Images Between Train/Valid/Test
Train 70%
Valid 20%
Test 10%
Image Distribution
Train: 223 images
Valid: 64 images
Test: 32 images
You are about to add 319 images to the dataset.
0 images will be sent back as part of a new job.
Add Images


3
Preprocessing
What can preprocessing do?
Decrease training time and increase performance by applying image transformations to all images in this dataset.
Auto-Orient
Edit
Resize
Stretch to 640×640
Edit
Add Preprocessing Step
Continue


Create New Version
Prepare your images and data for training by compiling them into a version.
Experiment with different configurations to achieve better training results.
Source Images
Images: 319
Classes: 5
Unannotated: 0
Train/Test Split
Training Set: 223 images
Validation Set: 64 images
Testing Set: 32 images

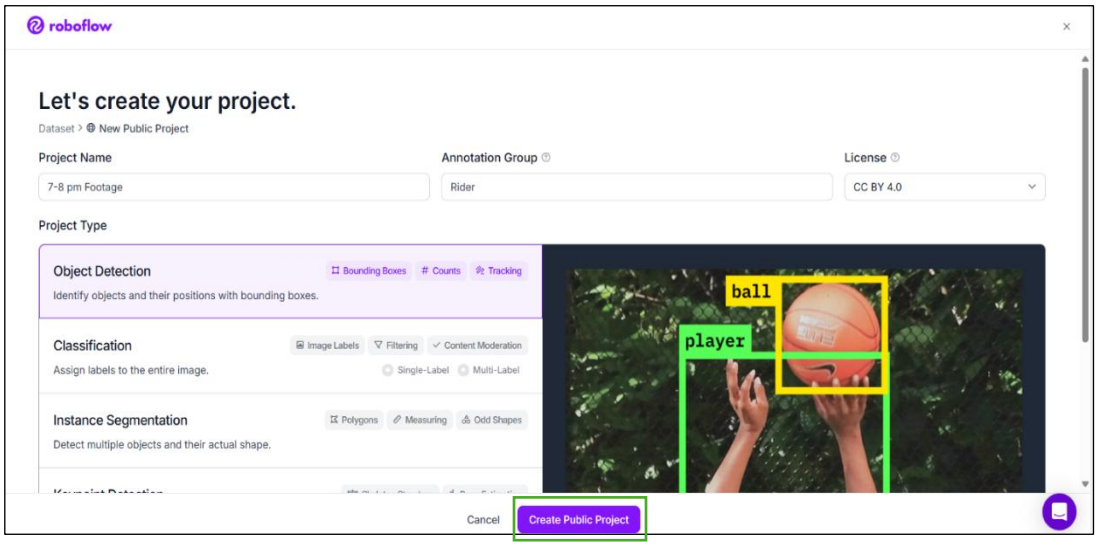

4
Augmentation
What can augmentation do?
Create new training examples for your model to learn from by generating augmented versions of each image in your training set.
Saturation
Between -25% and +25%
Edit
Brightness
Between 0% and +20%
Edit
Blur
Up to 0.9px
Edit
Noise
Up to 0.14% of pixels
Edit
Add Augmentation Step
Continue
Clear All

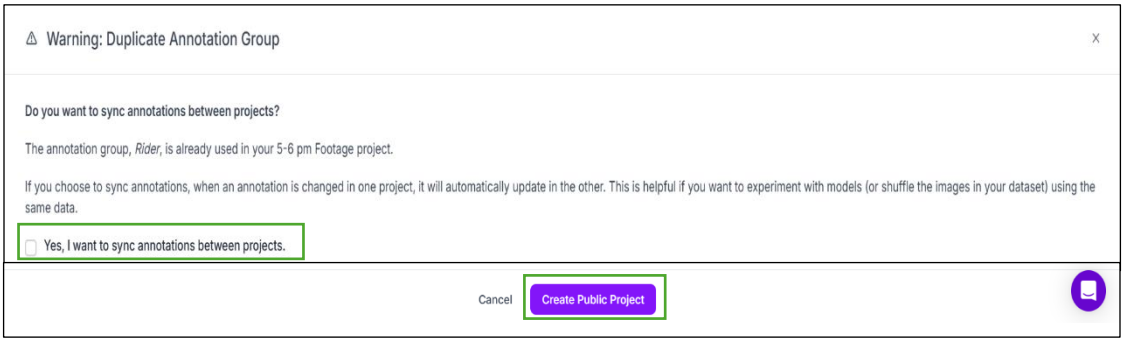

5
Create
Review your selections and select a version size to create a moment-in-time snapshot of your dataset with the applied transformations.
Larger versions take longer to train but often result in better model performance. See how this is calculated
Maximum Version Size
2,326 images (10x)
Create

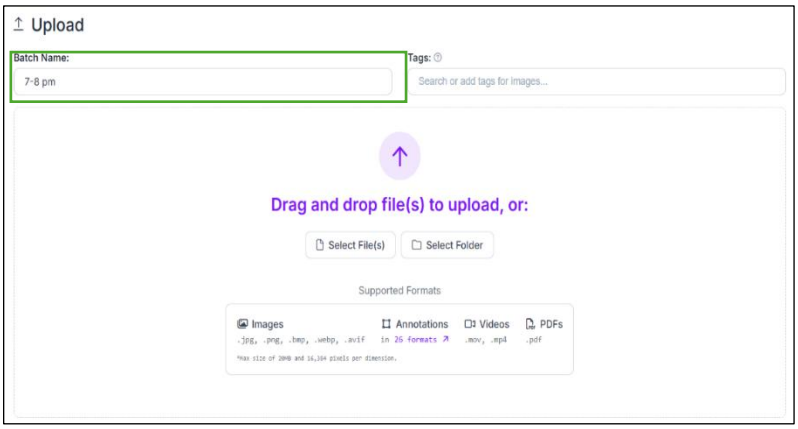


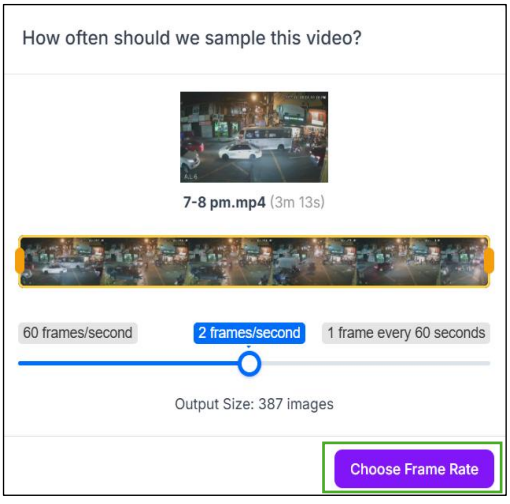
ANNONATING 7-8 PM IMAGES

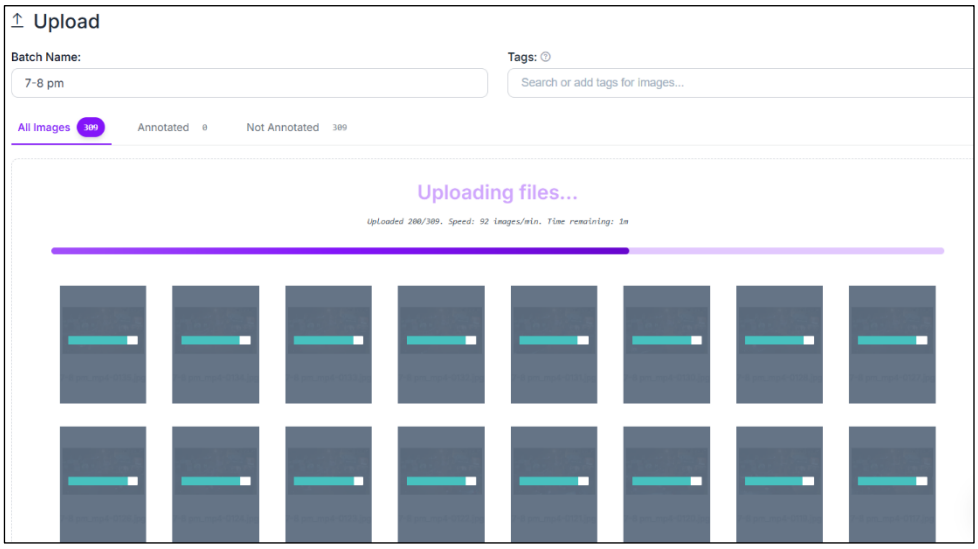
Finally, we moved on to annotating `7-8 PM images`. For this dataset, just like the `6-7 PM` images, we created a `New Project` and named it `7-8 pm Footage with Rider as its Annotation Group`. Once again, we did not sync the classes with the previous dataset, so the previous dataset would not be affected if there are changes in labels within this dataset. We named the `Batch Name as 7-8 pm` and dragged the third video to start processing it. Like the previous process, we waited until the scanning of the file is done and chose a `frame rate of 2 frames/seconds` to create an `Output Size of 387 images`. The extracted image was reduced to `309 images` and then we started annotating them with the help of `Label Assist`.

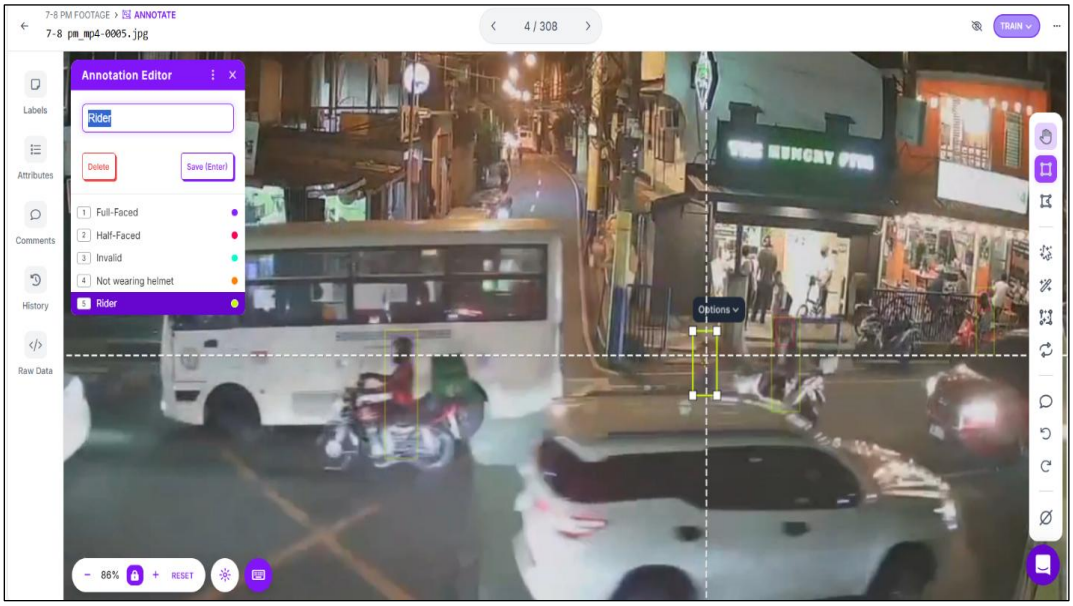

roboflow
Let's create your project.
Dataset > New Public Project
Project Name
7-8 pm Footage
Annotation Group
Rider
License
CC BY 4.0
Project Type
Object Detection
Bounding Boxes
Counts
Tracking
Identify objects and their positions with bounding boxes.
Classification
Image Labels
Filtering
Content Moderation
Assign labels to the entire image.
Single-Label
Multi-Label
Instance Segmentation
Polygons
Measuring
Odd Shapes
Detect multiple objects and their actual shape.
ball
player
Cancel
Create Public Project


Warning: Duplicate Annotation Group
Do you want to sync annotations between projects?
The annotation group, Rider, is already used in your 5-6 pm Footage project.
If you choose to sync annotations, when an annotation is changed in one project, it will automatically update in the other. This is helpful if you want to experiment with models (or shuffle the images in your dataset) using the same data.
Yes, I want to sync annotations between projects.
Cancel
Create Public Project

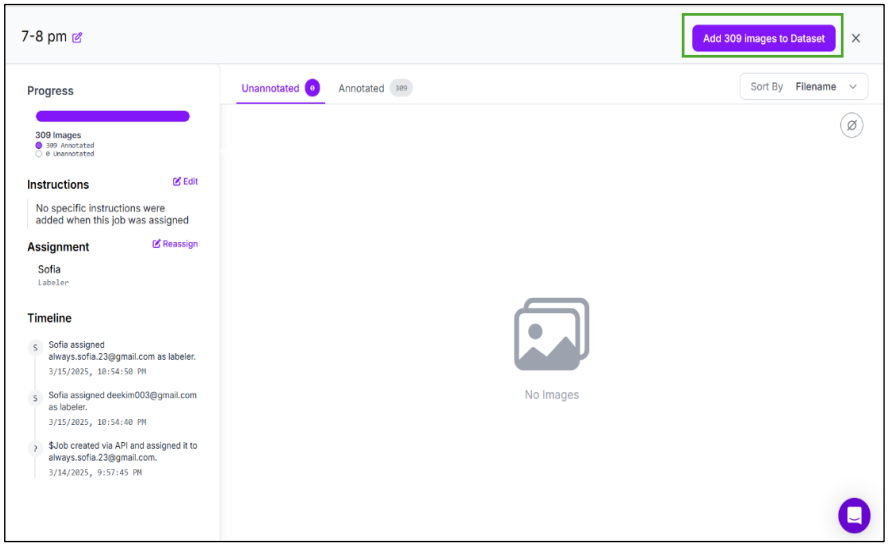

Upload
Batch Name:
7-8 pm
Tags:
Search or add tags for images...
Drag and drop file(s) to upload, or:
Select File(s)
Select Folder
Supported Formats
Images
Annotations
Videos
PDFs

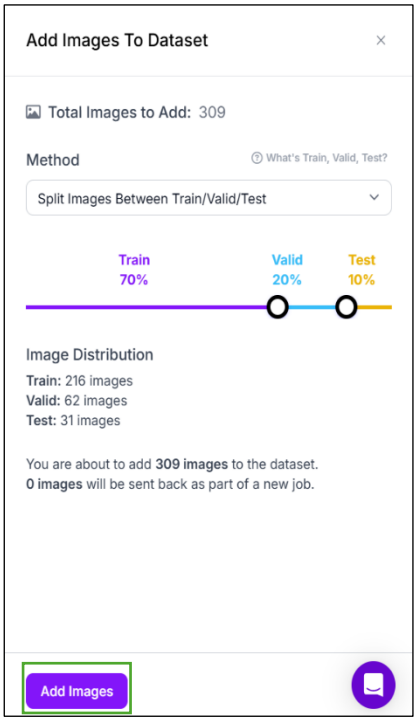

How often should we sample this video?
7-8 pm.mp4 (3m 13s)
60 frames/second
2 frames/second
1 frame every 60 seconds
Output Size: 387 images
Choose Frame Rate

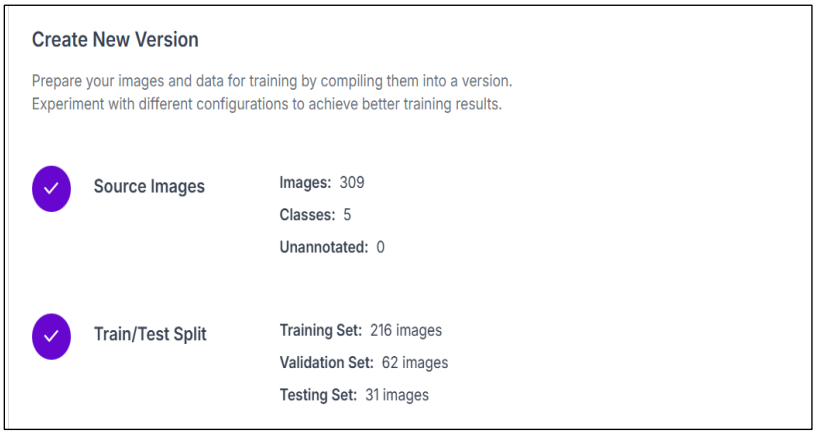


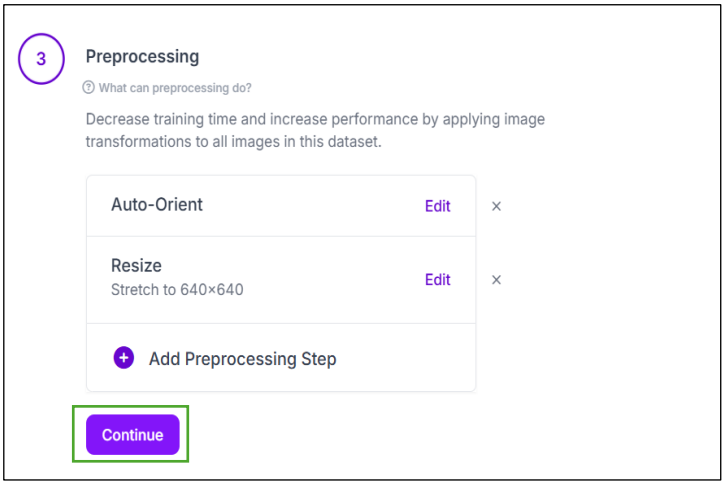


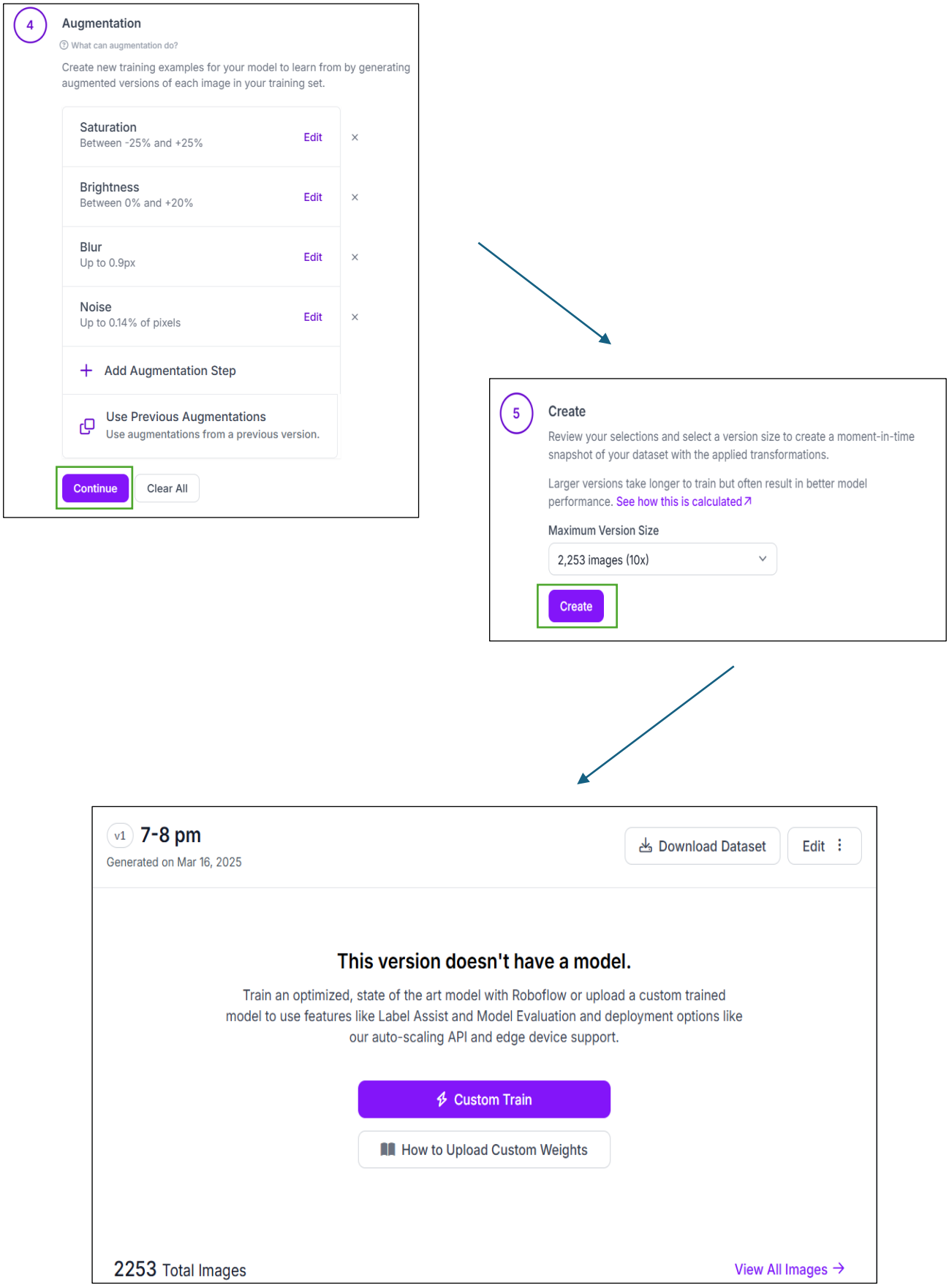
After annotating the images, we add them to the dataset by clicking the `Add 309 images to the Dataset` button and split the images between `Train, Valid, and Test.` We used the same `pre-processing techniques` and `augmentation` options. Then, we selected the `Maximum Version Size of 2,253 images` and clicked the `Create` button to start generating the dataset. The finished dataset contains `2,253 images.`


7-8 pm
Add 309 images to Dataset
Progress
309 Images
309 Annotated
0 Unannotated
Instructions
Edit
No specific instructions were added when this job was assigned
Assignment
Reassign
Sofia
Labeler
Timeline
Sofia assigned always.sofia.23@gmail.com as labeler.
3/15/2025, 10:54:50 PM
Sofia assigned deekim003@gmail.com as labeler.
3/15/2025, 10:54:40 PM
$Job created via API and assigned it to always.sofia.23@gmail.com.
3/14/2025, 9:57:45 PM
Unannotated 0
Annotated 309
Sort By Filename
No Images


Add Images To Dataset
Total Images to Add: 309
Method
What's Train, Valid, Test?
Split Images Between Train/Valid/Test
Train 70%
Valid 20%
Test 10%
Image Distribution
Train: 216 images
Valid: 62 images
Test: 31 images
You are about to add 309 images to the dataset.
0 images will be sent back as part of a new job.
Add Images


Create New Version
Prepare your images and data for training by compiling them into a version.
Experiment with different configurations to achieve better training results.
Source Images
Images: 309
Classes: 5
Unannotated: 0
Train/Test Split
Training Set: 216 images
Validation Set: 62 images
Testing Set: 31 images


3
Preprocessing
What can preprocessing do?
Decrease training time and increase performance by applying image transformations to all images in this dataset.
Auto-Orient
Edit
Resize
Stretch to 640×640
Edit
Add Preprocessing Step
Continue


4
Augmentation
What can augmentation do?
Create new training examples for your model to learn from by generating augmented versions of each image in your training set.
Saturation
Between -25% and +25%
Edit
Brightness
Between 0% and +20%
Edit
Blur
Up to 0.9px
Edit
Noise
Up to 0.14% of pixels
Edit
Add Augmentation Step
Use Previous Augmentations
Use augmentations from a previous version.
Continue
Clear All
5
Create
Review your selections and select a version size to create a moment-in-time snapshot of your dataset with the applied transformations.
Larger versions take longer to train but often result in better model performance. See how this is calculated
Maximum Version Size
2,253 images (10x)
Create
v1
7-8 pm
Generated on Mar 16, 2025
Download Dataset
Edit
This version doesn't have a model.
Train an optimized, state of the art model with Roboflow or upload a custom trained model to use features like Label Assist and Model Evaluation and deployment options like our auto-scaling API and edge device support.
Custom Train
How to Upload Custom Weights
2253 Total Images
View All Images