

Sistema:

Windows 11 + WSL2 Ubuntu 22.04.2 LTS

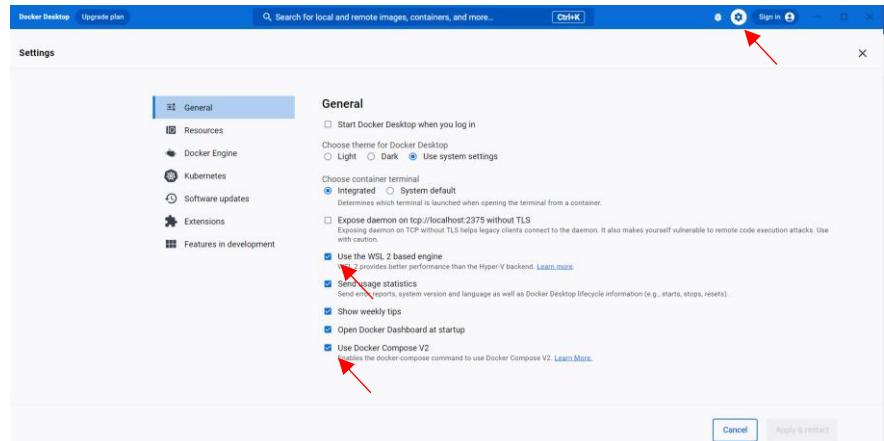
<https://learn.microsoft.com/es-es/windows/wsl/install>

1. Instal·lació

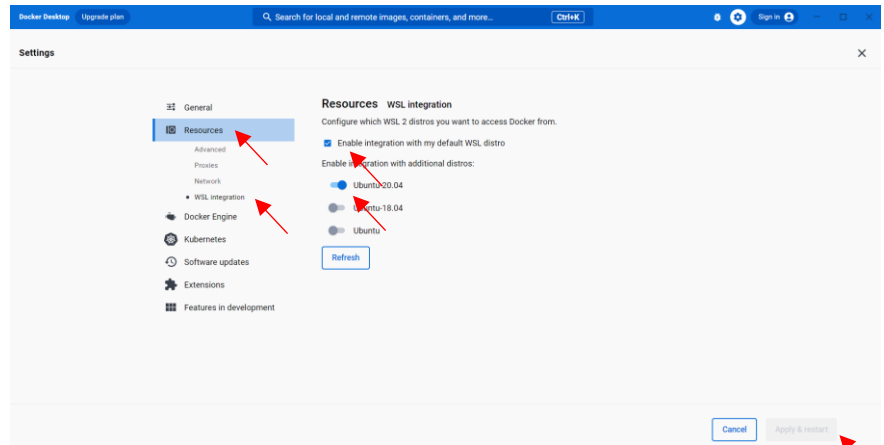
a. Instal·lar Docker Desktop (Windows)

b. Configurar Docker

i.



ii.



c. Descargar CVAT en el sistema del WSL2 (buscar icona Ubuntu al menú de Windows)

`git clone` <https://github.com/opencv/cvat>

Accedir al directori

`cd cvat`

- d. Un directori compartit per comoditat al tractar arxius grans

```
git config core.autocrlf false
```

```
docker volume create --name cvat_share --opt  
type=none --opt device=/mnt/share --opt o=bind
```

Crear l'arxiu "docker-compose.override.yml" amb:

```
services:  
  cvat_server:  
    volumes:  
      - cvat_share:/home/django/share:ro  
  cvat_worker_import:  
    volumes:  
      - cvat_share:/home/django/share:ro  
  
volumes:  
  cvat_share:  
    external: true
```

```
docker compose -f docker-compose.yml -f docker-  
compose.dev.yml -f docker-compose.override.yml up  
-d --build
```

(El pas anterior pot tardar molt, aproximadament 30 minuts amb un internet bo)

Registrar un super usuari

```
sudo docker exec -it cvat_server bash -ic 'python3  
~/manage.py createsuperuser'
```

i seguir les instruccions

Reiniciar l'ordinador o apagar i encendre l'ordinador

- e.

Accedir a CVAT en el Google Chrome a través de l'adreça localhost:8080

2. Preprocessament

- a. CVAT no funciona amb resolucions superiors a **3000 x 2000 px**. Redueixo de **4000 x 3000 px** a **2000 x 1500 px** (1 px de cada 2 px en ambdues direccions, aquest valor es reaprofità posteriorment al editar les anotacions) amb la comanda següent:

```
ffmpeg -y -i output_4_gt.mp4 -vf scale=2000:-2,setsar=1:1 -c:v libx264 "output_4_gt_ds.mp4"
```

Al input "output_4_gt.mp4" fes que l'**amplada** sigui 2000 px **mantenint la relació d'aspecte**, codifica-ho i **guarda-ho a** "output_4_gt_ds.mp4" (estic a la carpeta compartida "/mnt/share/"). Posteriorment ja es processaran les anotacions a la resolució inicial.

- b. El format d'anotacions de CVAT esta lleugerament canviat i la millor solució trobada requereix l'ús de classes en lloc de numero de track. Per poder importar anotacions a cada tasca (**cada vídeo es una tasca**) correctament, he creat els scripts `minimum_id.py` i `pred_to_cvat.py` que donat un arxiu MOT d'entrada genera un ".zip" importable des de CVAT (les anotacions estan adaptades) i un ".json" definint els labels:

```
python3 minimum_id.py input_1.txt output.txt
python3 prde_to_cvat.py input_2.txt output.zip
```

```
~/AntTracking$ python3 minimum_id.py DATA/output_4_gt.txt DATA/output_4_gt_min.txt
~/AntTracking$ python3 prde_to_cvat.py DATA/output_4_gt_min.txt gt.zip
```

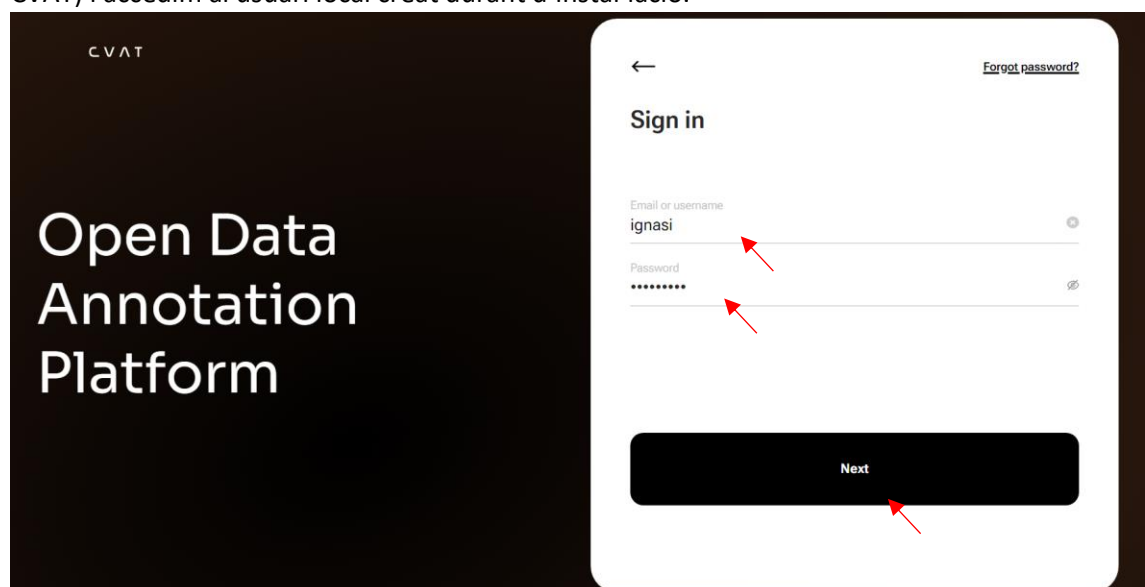
3. Instruccions nova tasca/vídeo en CVAT

a. Disposant dels resultats anteriors:

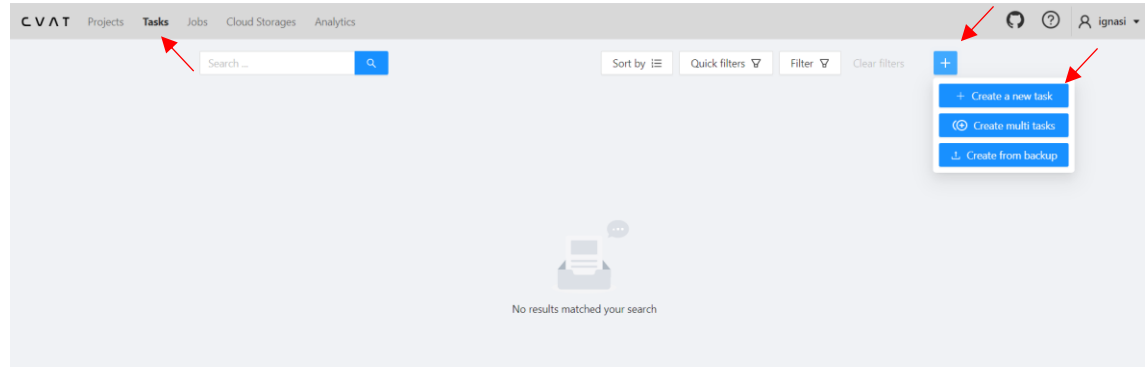
```
ignasi@DESKTOP-P9F9UTF:/mnt/share$ ls
output_4_gt_ds.mp4
ignasi@DESKTOP-P9F9UTF:/mnt/share$ tree -L 1 ~/AntTracking/
/home/ignasi/AntTracking/
├── DATA
├── LICENSE
├── OUTPUT
├── ant_detection.py
├── ant_detection_pca.py
├── associated_histograms.ipynb
├── associated_histograms_pca.ipynb
├── docopts
├── gt.json
├── gt.zip
├── hota_idf1.ipynb
├── metrics
├── minimum_id.py
├── models
├── ocsort_track.py
├── pca_tracks.py
├── plot_pca_directions.ipynb
├── plot_rectangles_video.py
├── pred_to_cvat.py
├── sort_inference.py
└── unassociated_histograms.ipynb

5 directories, 16 files
```

b. Accedim a localhost:8080 des del navegador Google Chrome (requeriment de CVAT) i accedim al usuari local creat durant d'instal·lació.



- c. En el menú de tasques, creem una nova:



- d. Configurarem la tasca:

- i. Utilitzem el nom del vídeo original com a nom de la tasca (deixem el nom del projecte buit):

The screenshot shows the 'Basic configuration' form for creating a new task. It has three main sections: 'Name', 'Project', and 'Labels'. The 'Name' field is filled with 'output_4_gt' and has a green checkmark. The 'Project' field is empty and has a placeholder text 'Select project'. The 'Labels' field is empty and has a placeholder text 'Select labels'.

- ii. Definim les etiquetes (Labels) com el ID de cada formiga; comencem important les IDs creades per el model, aquestes ja tenen el format desitjat en l'arxiu "gt.json". **Copiem el contingut del arxiu d'etiquetes "gt.json" i l'enganxem a "Labels": "Raw":**

The screenshot shows the 'Basic configuration' form for creating a new task. It has three main sections: 'Name', 'Project', and 'Labels'. The 'Name' field is filled with 'output_4_gt' and has a green checkmark. The 'Project' field is empty and has a placeholder text 'Select project'. The 'Labels' field is filled with JSON data from a file named 'gt.json'. The JSON data is displayed in a text area with a 'Raw' tab selected. Below the text area, there are 'Done' and 'Reset' buttons. A red arrow points to the 'Raw' tab, and another red arrow points to the 'Done' button.

- iii. Importem el vídeo de resolució meitat (2000x1500 px) de la carpeta compartida:

Basic configuration

* Name
output_4_gt

Project
Select project

* Labels
Raw dFP Constructor
[
 {
 "name": "1",
 "type": "rectangle",
 "attributes": []
 },
]

Done Reset

* Select files
My computer Connected file share Remote sources Cloud Storage

☒ root
☒ output_4_gt_ds.mp4

- iv. Seleccionem el format “MOT 1.1” per la tasca en les opcions avançades i creem la tasca:

Advanced configuration

* Sorting method
Lexicographical Natural Predefined Random
Specify how to sort images. It is not relevant for videos.

☐ Copy data into CVAT
If you have a low data transfer rate over the network you can copy data into CVAT to speed up work

☒ Use zip/video chunks ☒ Use cache

* Image quality
70 %
Overlap size Segment size

Start frame Stop frame Frame step

Chunk size

Dataset repository URL
e.g. https://github.com/user/repos [annotation/<anno_file_name>.zip]
Attach a repository to store annotations there

Choose format
MOT 1.1
COCO 1.0
COCO Keypoints 1.0
CVAT for video 1.1
CVAT for images 1.1
Datamaro 1.0
Datamaro 3D 1.0
LabelMe 3.0
Segmentation mask 1.1

Submit & Open Submit & Continue

Advanced configuration

* Sorting method
Lexicographical Natural Predefined Random
Specify how to sort images. It is not relevant for videos.

☐ Copy data into CVAT
If you have a low data transfer rate over the network you can copy data into CVAT to speed up work

☒ Use zip/video chunks ☒ Use cache

* Image quality
70 %
Overlap size Segment size

Start frame Stop frame Frame step

Chunk size

Dataset repository URL
e.g. https://github.com/user/repos [annotation/<anno_file_name>.zip]
Attach a repository to store annotations there

Choose format
MOT 1.1
COCO 1.0
COCO Keypoints 1.0
CVAT for video 1.1
CVAT for images 1.1
Datamaro 1.0
Datamaro 3D 1.0
LabelMe 3.0
Segmentation mask 1.1

☒ Use LFS (Large File Support)

Issue tracker

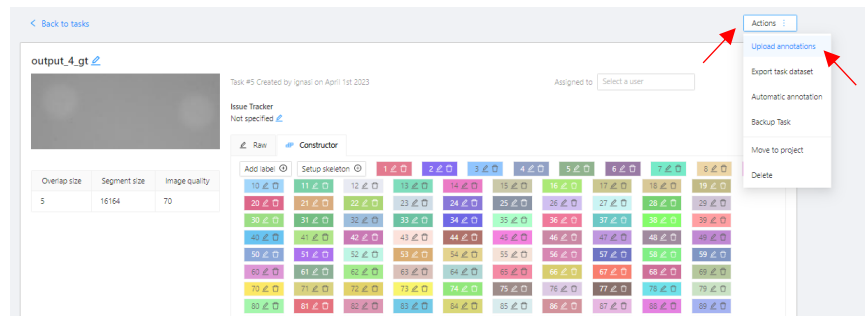
Attach issue tracker where the task is described

Source storage Target storage
Local Local

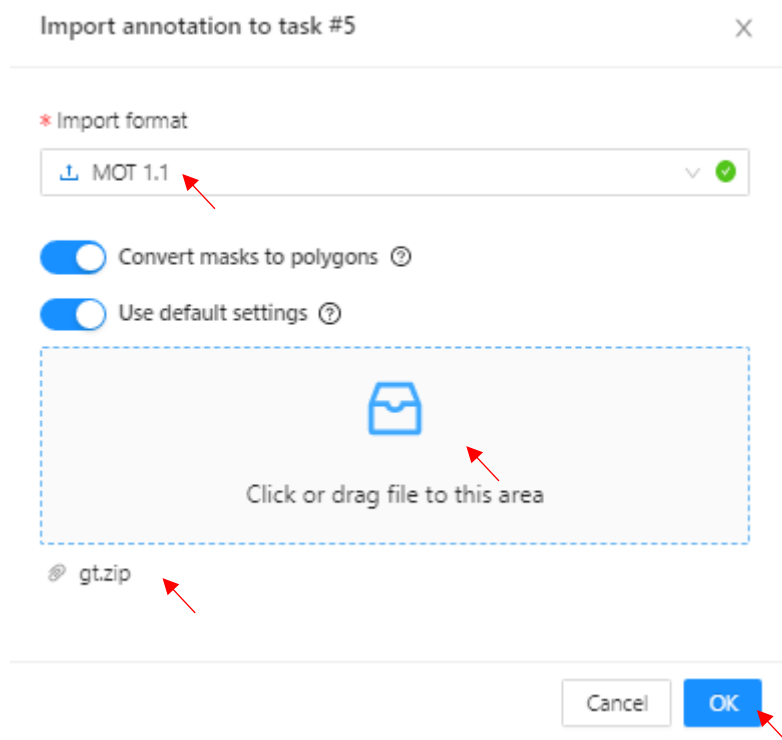
Submit & Open Submit & Continue

e. Importem les anotacions

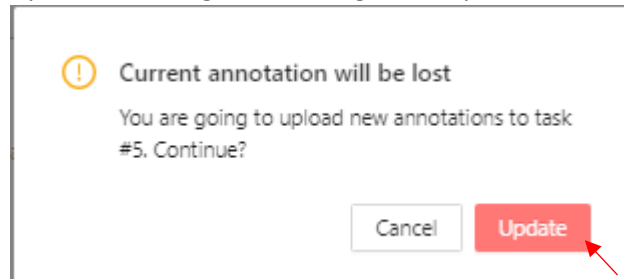
- i. Dins la tasca anem a “Actions”, “Update annotations” per carregar els resultats dels models:



- ii. Seleccionem el format “MOT 1.1” i l’arxiu generat prèviament “gt.zip”:



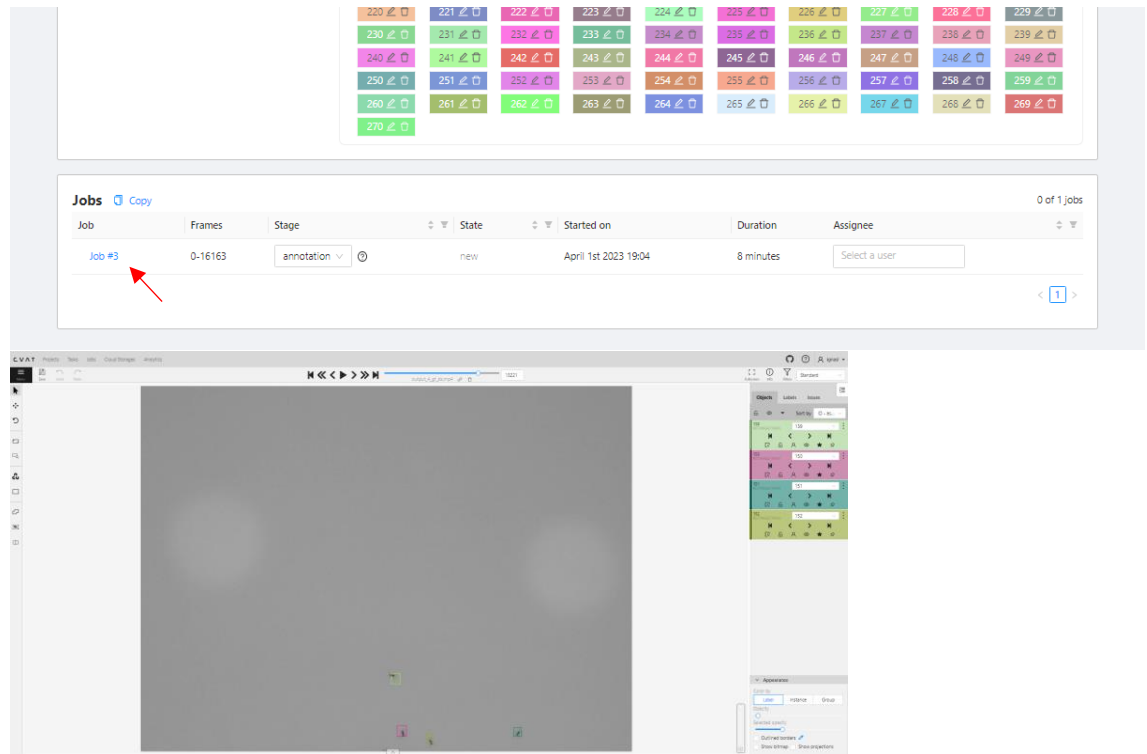
Apareixerà el següent missatge, l’acceptem:



- iii. Esperem al missatge “bullet” (no desapareix) que avisa sobre el final de la importació i actualitzem la pàgina (per assegurar de que tot està bé).

4. Instruccions sobre l' anotació:

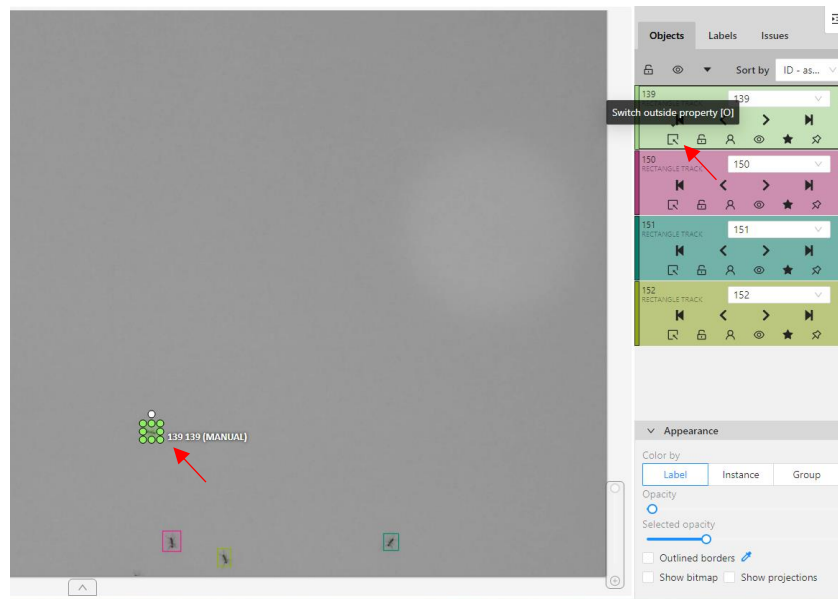
a. Accedir a la tasca:



b. Modificar regions existents o eliminar tracks sensors:

Funciona igual que redimensionar, moure (ratolí) o eliminar (boto suprimir del teclat) una imatge en un programa com PowerPoint. En general, no volem eliminar tracks sensors.

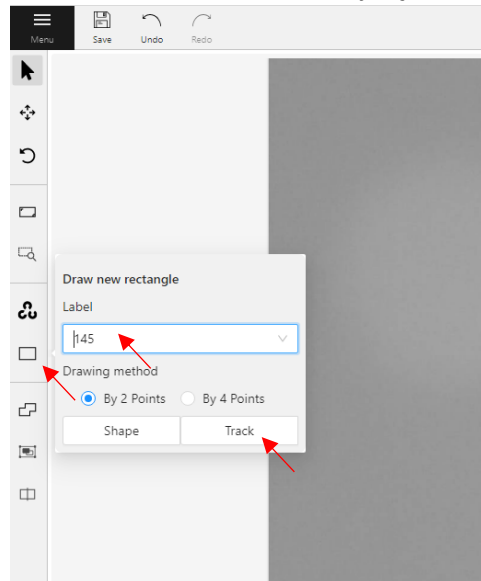
c. Finalitzar un track:



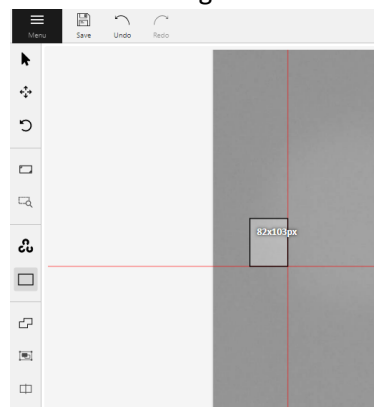
Al clicar la "O" o el boto senyalat, el track passa a un estat d'inactivitat "final": el track desapareix i no es pot tornar a fer aparèixer el mateix (es pot fer aparèixer un track nou de la mateixa classe que representa la mateixa

formiga).

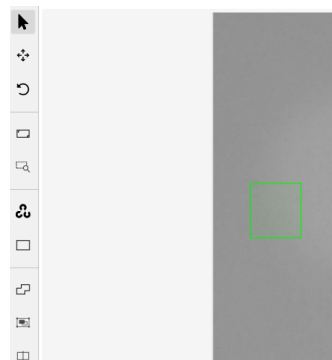
- d. **Crear un nou track (d'una ID existent):**
Podem buscar la ID/Label mitjançant text:



Clicant i arrossegant:

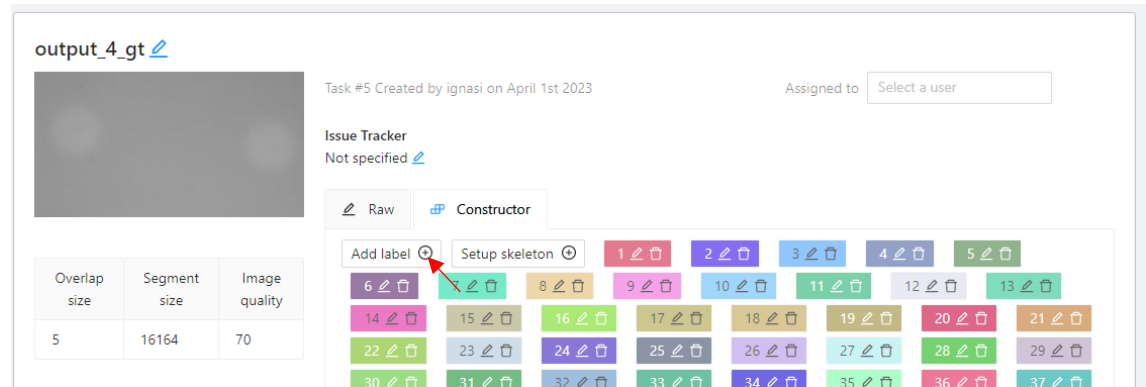


Tornant a clicar:

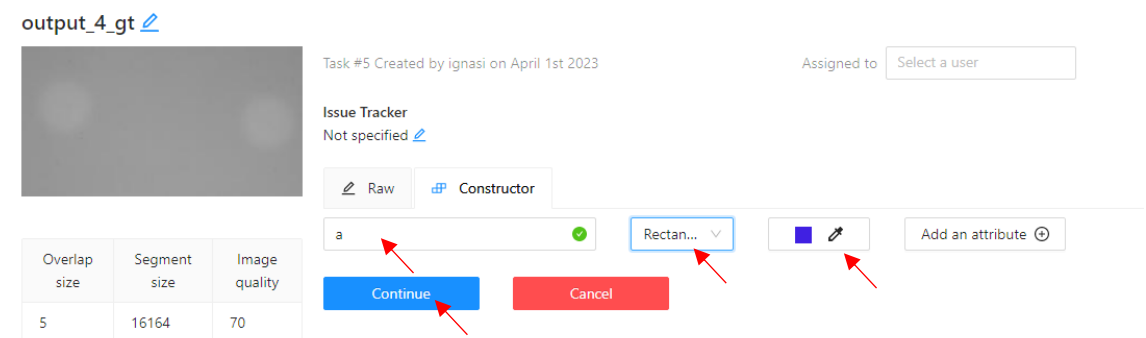


e. **Crear una Label nova per una formiga no detectada:**

Tornem al menú de la tasca (on s'ha fet la importació d'anotacions).



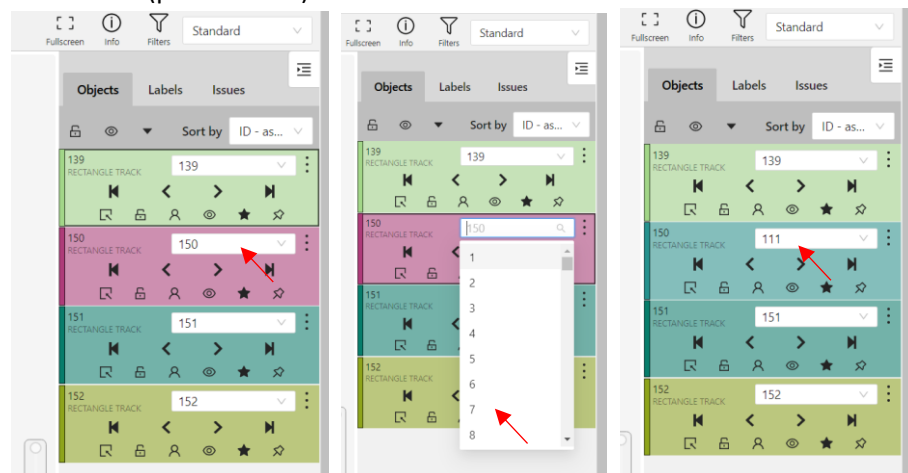
Escollim un nom, per assegurar que no hi ha col·lisió de noms, es pot usar lletres (a, b, c...) o alfanumèrics (a1, a2, a3...):



Posteriorment es pot usar la nova label al crear tracks nous.

f. **Fusió de tracks:**

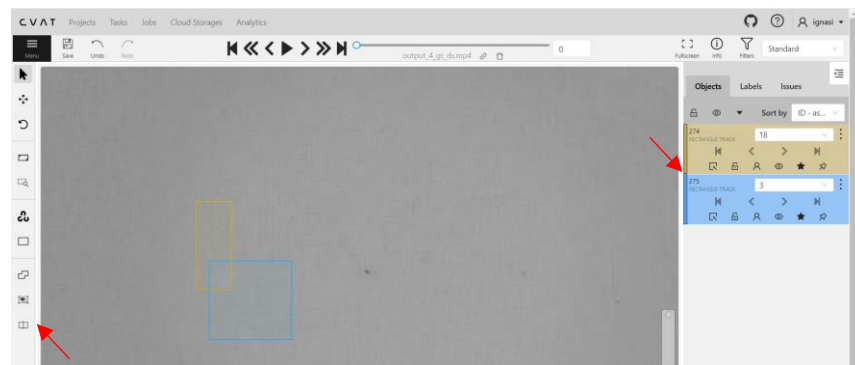
En cas de tracks perdut amb la formiga correctament perseguida, només cal canviar la classe del nou track a la del passat en un punt qualsevol i es canviarà tot el track (pasat i futur):



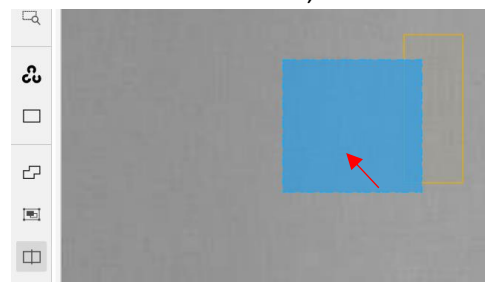
En aquest cas, també es pot buscar la ID escrivint.

g. **Arreglar intercanvi d'identitats:**

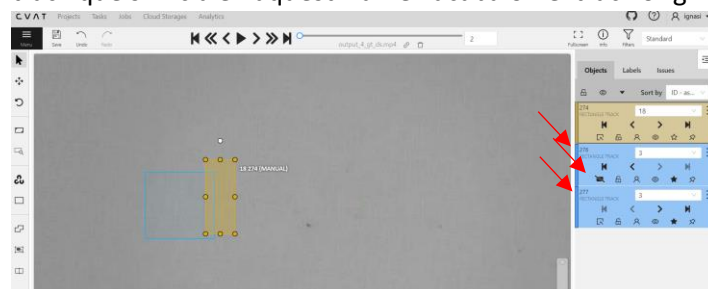
Primer dividim els tracks:



Al clicar la eina de dividir, i clicar al track a dividir:

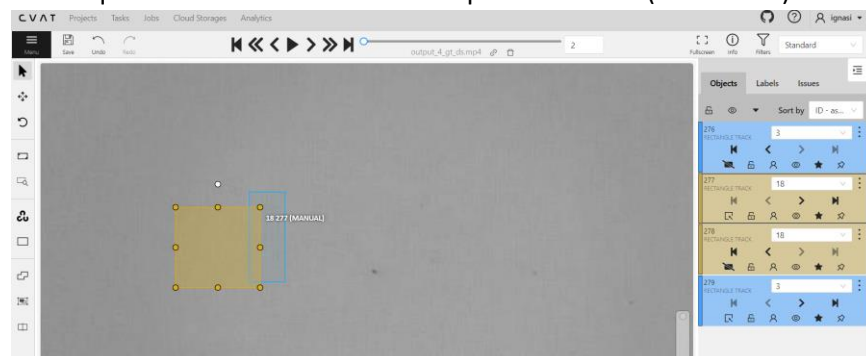


El track dividit es “duplica”; al menu, l'original queda finalitzat i es crea un track que s'inicia en aquest frame i acaba on el track original hauria acabat:



Ho fem pels 2 tracks que s'han confós.

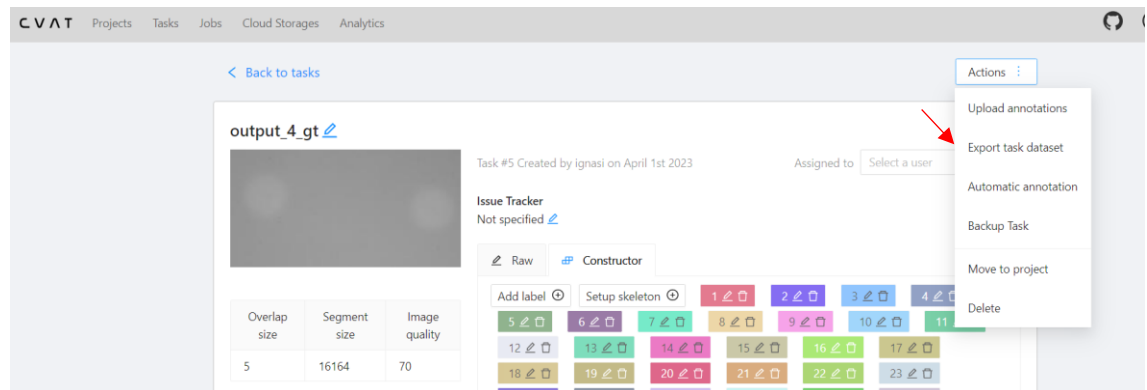
Modifiquem la classe dels nous tracks per la correcta (la del altre):



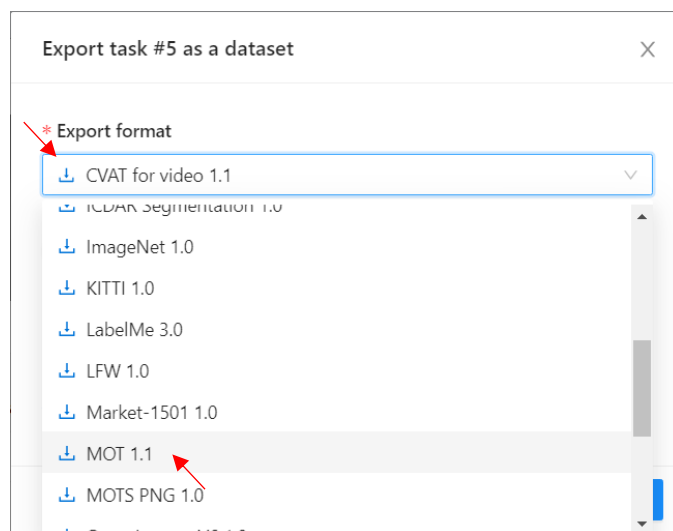
En cas de confusions més complexes entre varies formigues, la idea es la mateixa.

5. Exportació de resultats:

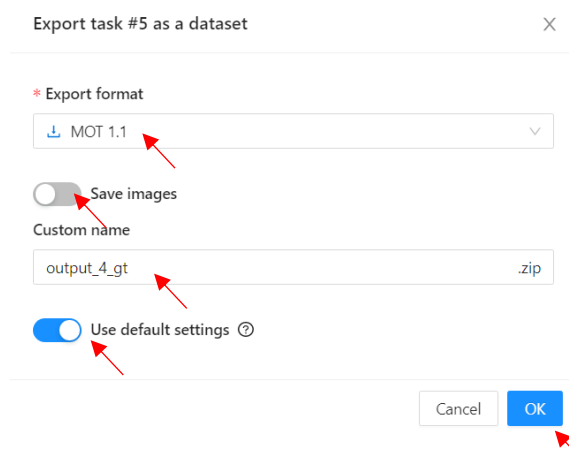
En la tasca, al menu “Actions”, clicar a “Export task dataset”:



Escollir la opció “MOT 1.1”



Guardar en el disc dur local (default Settings):



Finalment esperar a que estigui llest i es descarregui un zip amb les anotacions.

FAQs

1. CVAT no inicia bé o no reconeix l'usuari:

Amb el docker desktop de encès (en el cas de WSL2, el de Windows).

Des de la carpeta d'instal·lació del CVAT (des de WSL2), repetir les comandes:

```
docker volume create --name cvat_share --opt type=none -  
-opt device=/mnt/share --opt o=bind  
  
docker compose -f docker-compose.yml -f docker-  
compose.dev.yml -f docker-compose.override.yml up -d --  
build
```

2. Una formiga no ha estat detectada:

Es molt probable que s'hagi finalitzat un track anteriorment amb una identitat que posteriorment es reassignaria a la formiga no detectada.

Crear un nou track pot requerir molt de temps, val la pena comprovar els tracks finalitzats en els darrers 31 frames.

Exemples en vídeo

1. **Modificar regions existents i Arreglar oclusió**

<https://drive.google.com/file/d/1A0Y9yFOY6xPiw3OYJ5A94AmQozsQ9gA3/view?usp=sharing>

2. **Crear una Label nova i Arreglar intercanvi d'identitats**

<https://drive.google.com/file/d/19yytkZSj6btR62L6YreIMpN1mMZY2lwM/view?usp=sharing>