Scottish Government
Riaghaltas na h-Alba

A New Application for Automated Video Identification of Marine Species (AVIMS)

Scottish Government Marine Directorate

Scottish Marine and Freshwater Science Vol 14 No 7

M. Mackiewicz, G. French and M. Fisher

Annex 2 – AVIMS Tutorial Slides

A New Application for Automated Video Identification of Marine Species (AVIMS)

Annex 2 – AVIMS Tutorial Slides

Scottish Marine and Freshwater Science Vol 14 No 7

M. Mackiewicz, G. French and M. Fisher

Published by the Marine Directorate of the Scottish Government

ISBN:
ISSN: 2043-7722
DOI: 10.7489/12473-1

The Marine Directorate of the Scottish Government is responsible for the integrated management of Scotland's seas. Scottish Marine and Freshwater Science is a series of reports that publishes the results of scientific research and monitoring carried out by the Marine Directorate. It also publishes the results of marine and freshwater scientific work that has been carried out for the Marine Directorate under external commission. These reports are not subject to formal external peer-review.

This document is an Annex to the Main Report which presents the results of scientific work commissioned by the Marine Directorate and carried out at the School of Computing Sciences, University of East Anglia.

The full report should be quoted as:
Mackiewicz, M., French, G. and Fisher, M. (2023). A new application for automated video identification of marine species (AVIMS). Scottish Marine and Freshwater Science Vol 14 No 7, 18pp. DOI: 10.7489/12473-1

© Crown copyright 2023

You may re-use this information (excluding logos and images) free of charge in any format or medium, under the terms of the Open Government Licence. To view this licence, visit:
http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/
or email: psi@nationalarchives.gsi.gov.uk.

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

A New Application For Automated Video Identification Of Marine Species (AVIMS)

Annex 2 – AVIMS Tutorial Slides

M. Mackiewicz, G. French and M. Fisher

School of Computing Sciences, University of East Anglia,
Norwich Research Park, Norwich. NR4 7TJ

Abstract

In the course of its environmental monitoring activities, the Scottish Government's Marine Directorate collects a large amount of underwater video to, for example, obtain information on the numbers of fish in rivers or on species living on the seabed. Manual analysis of this footage is laborious and costly, but Machine Learning algorithms can now be used to automate such image analysis. The Marine Directorate commissioned the University of East Anglia to develop a web-based application to allow staff to create, train and execute machine learning-based (semi-)automated analysis of video footage without a need to interact with the underlying computer code. The application was tested using three diverse sets of video footage and found to be usable by staff without computer science or coding experience. The tool was able to detect and count sea pens in footage from towed underwater vehicles, salmon smolts at sea in underwater footage from towed fishing gear and adult salmon and sea trout in footage from underwater or overhead cameras at fixed locations on rivers. Improving the accuracy of the models at detecting and counting organisms of interest will require the use of larger annotated datasets in further training of the algorithms, but the current application provides a basis for further developing these.

Contents

Abstract 1
Contents 2
Abbreviations 3
Overview 5
AVIMS 7
Machine Learning Intro 11
Using AVIMS 25
- Upload footage 30
- Select frames for labelling 39
- Labelling schema 50
- Annotate objects in training images 58
- Building a training set 65
- Train an object detection model 78
- Analyse videos 86
- Circular workflow 101

Contact 106

Abbreviations

AVIMS	Automated Video Identification of Marine Species
CNN	Convoluted Neural Network
GPU	Graphics Processing Unit
GUI	Graphical User Interface
mAP	Mean Average Precision
OS	Operating System

AVIMS tutorial

Geoff French

University of East Anglia

September 2021

Overview

Overview

- AVIMS
- Machine learning intro
- Using AVIMS
  - Step 1a: Upload footage
  - Step 1b: Select frames for annotation
  - Step 2: Labelling schema: what objects/creates are we interested in?
  - Step 3: Annotate objects in training images
  - Step 4: Build a training dataset
  - Step 5: Train an object detection model
  - Step 6: Analyse videos
  - Circular work-flow

AVIMS

AVIMS

- AVIMS is an automated video analysis tool for detecting and counting various creatures or objects of interest in videos

AVIMS

- Detect then track
  - A creature or object of interest is detected in a video frame
  - If it is detected in a nearby location in subsequent frames it can be tracked
  - Tracking objects across frames supports quantification
  - The objects can be classified, e.g. by species, etc.

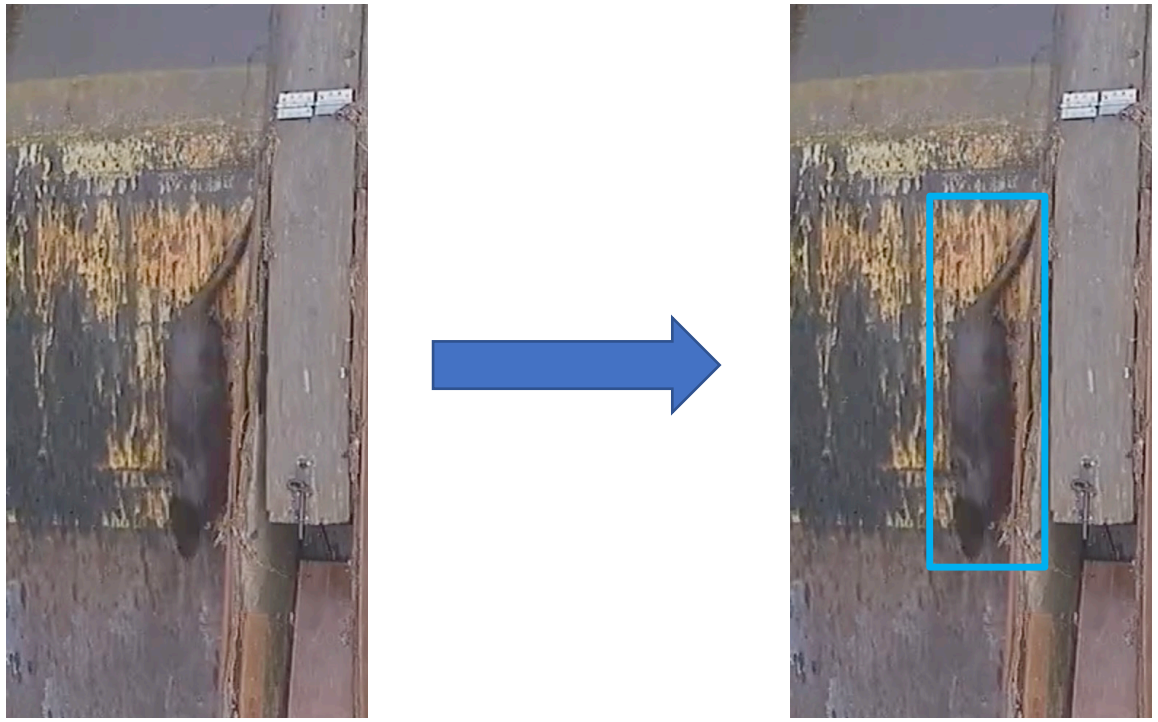
AVIMS

- AVIMS uses machine learning
- This requires the user (you ☺) to manually annotate images, identifying objects if interest so that the system can be taught what to look for
- Once trained, the system can be let loose on videos you want analysed

Machine learning intro

Machine learning intro

- A machine learning *model* is trained so that given *input* samples, it will predict correct *outputs*



Training an ML model

- Construct a dataset consisting of:
  - Input samples
  - Manually prepared ground truths; one for each input
- These are handed to the training software that trains the model to predict the outputs given the inputs
- Then we set the model loose on samples, e.g. images/video that we want analyzed!
  - Well, that's the theory anyway…

Machine learning – generalization

- Generalization refers to the ability of the model to generate correct predictions given input samples it hasn't seen before
- Machine learning models often *overfit* their training data
  - They often generate very good predictions for samples in the training set
  - Performance on as-of-yet unseen samples is usually lower
  - It would be advantageous to know how well our model will perform on unseen samples…

Machine learning – train/test split

- We take a dataset of samples with corresponding ground truths and split it into a *training set* and *test set*
- The model is trained using the *training set*
- The samples in the *test set* are held out
- After training, the performance of the model is measured using the *test set*
- In theory, this should give you an idea of how well it will perform in practice…

Machine learning – train/test split

- In practice, how well *test set* accuracy matches the accuracy obtained in the real world depends on how the *test set* was chosen!
- You need to ensure that the distribution of the samples in the *test set* matches that of the samples the model will be applied to in practice
- If the samples in the *test set* are too similar to those in the *training set* we will get a poor estimate of performance

Machine learning – train/test split

- How should we split our data? It depends on the use case.
- We will discuss three options
  - Split by site / camera
    - E.g. use site A, B, C, D, E, F for training, G & H for test
  - Split by video
    - E.g. use videos (all from the same camera) 1, 2, 3, 4 for training, 5 & 6 for test
  - Split by image/frame
    - E.g. select frames from within a video and randomly split between train and test

Machine learning – train/test split – split by camera

TRAIN

All images are from different cameras

TEST

Machine learning – train/test split – split by camera

- Measures the ability of the model to generalize to new sites/viewpoints
  - Probably what is wanted in many cases where you want to conduct surveys at new sites and use a model you trained earlier
- Very challenging from the perspective of the model!
- The training set will need to have images from a number of different cameras/viewpoints in order to have sufficient variety so that the viewpoints in the test set have training samples that are similar enough

Machine learning – train/test split – split by video

TRAIN

All images are from the same camera, but from different videos

TEST

Machine learning – train/test split – split by video

- Less challenging as the model does not have to generalize to new viewpoints
- Sometimes you only have one camera or differentiating between cameras doesn’t make sense, e.g. when the camera is moving in an ocean floor survey
- If you can’t get enough sites/cameras to get good enough performance, try annotating some frames from videos that are getting poor results

Machine learning – train/test split – split by frame/image

TRAIN

All images are from the same video, but from different but close-by frames

TEST

Machine learning – train/test split – split by frame/image

- Least challenging
  - Model performance will be good on close-by or similar frames, but poor on different videos or frames with different content, so watch out!
- If you have 1 long video and you want to annotate say 40 frames and have the model automatically analyse the rest

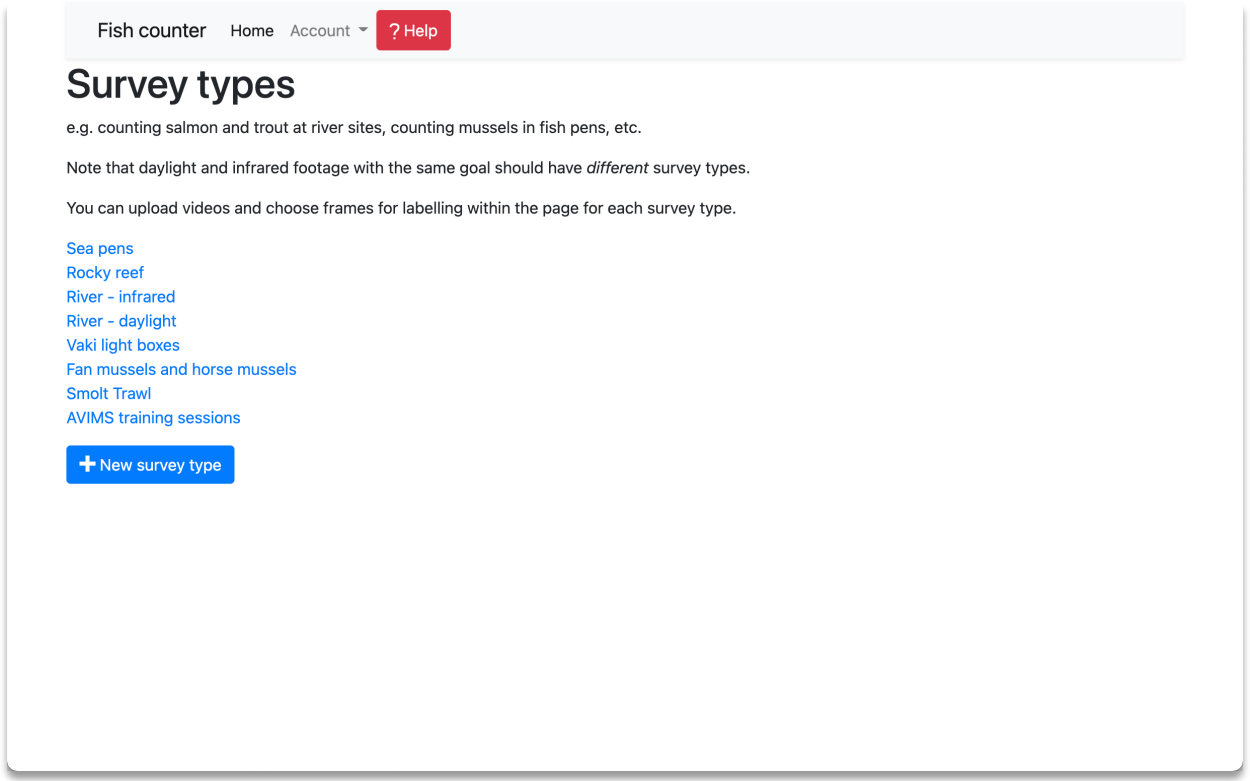
Machine learning – cyclic workflow

- In practice, we would:
  1. Select samples for our dataset
  2. Manually annotate them
  3. Split into train/test
  4. Train a model
  5. Evaluate its performance on novel samples/images/frames/videos/etc.
  6. If there are frames on which the model performs unacceptably poorly
     a. Add them to the dataset
     b. Go back to step 2 and manually annotate them and repeat the cycle until performance is good enough

Using AVIMS

Choose your survey type

- After logging in you need to choose the survey type you want to work on:



Choose your survey type

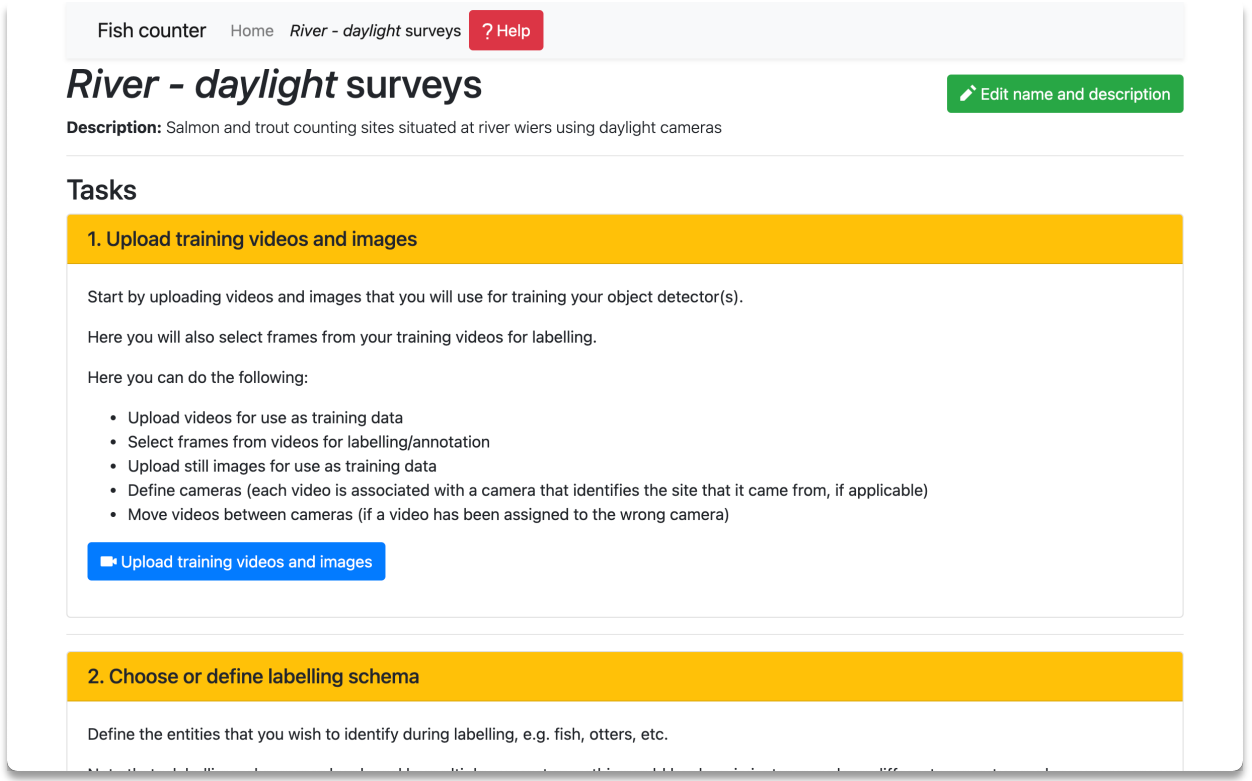
- A survey type:
  - Focuses on detecting a particular range of species or creatures of interest in common conditions/types of location
    - E.g. ‘counting salmon and otters at river weir sites’ should have a different survey type from the one used to ‘count sea pens in underwater footage’
  - Contains footage with a common ‘appearance’, separating daylight from night-time infrared, underwater from above water
    - E.g. while sharing the common goal of counting salmon and otters at river weir sites, put the daylight and night-time infrared footage in different survey types
    - If we had underwater cameras at river weir sites too, they should go in a different survey type

Choose your survey type

- When to create a new survey type:
  - Create a new survey type when there isn't one that shares the same goals or the footage has a very different appearance than one of the existing ones.
  - Don't create different survey types for e.g. different years. Split this by different 'cameras' (more on this later)
- Why?
  - All of the data associated with a survey type will be used to train an object detection machine learning (ML) model. Best performance will be obtained by having this model focus on a specific problem and type of *visual appearance*, so group footage by these criteria

Choose your survey type

- After choosing or creating a survey type you should see this:

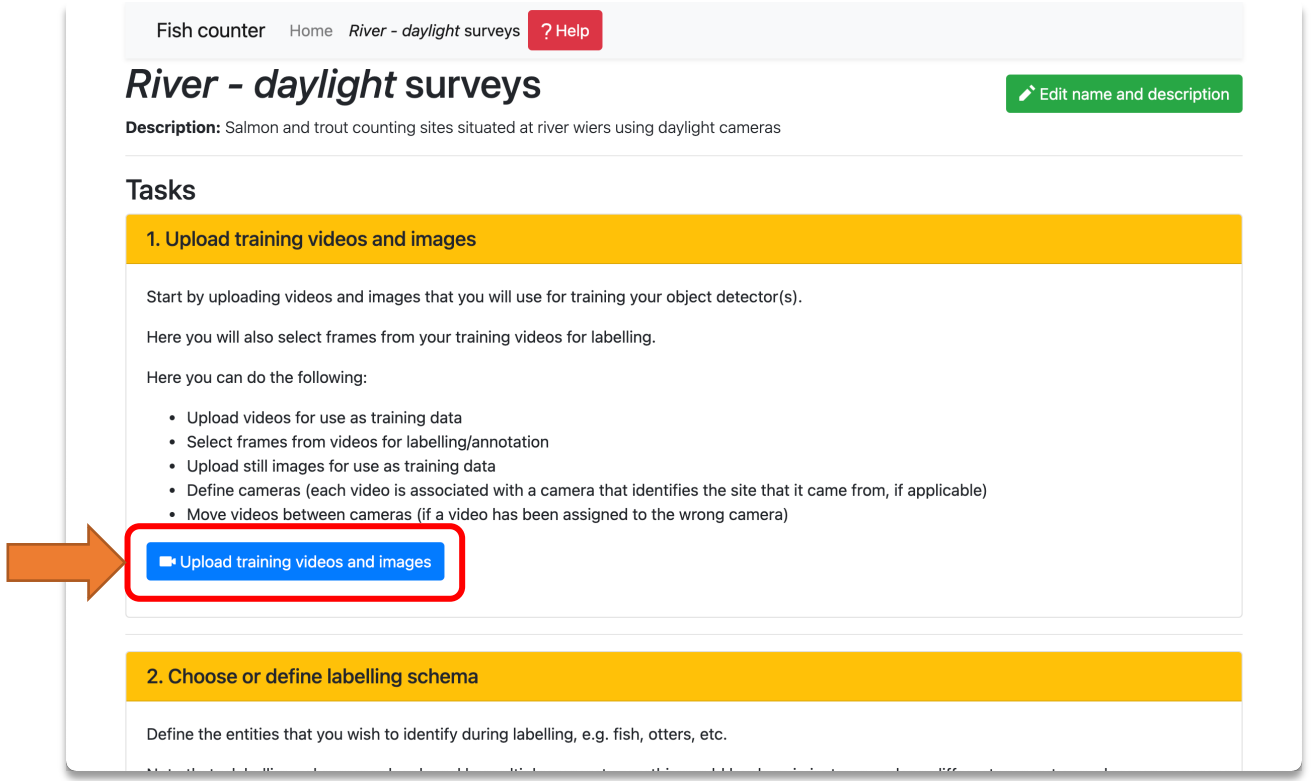


Using AVIMS

1.a. Upload footage

1.a. Upload footage

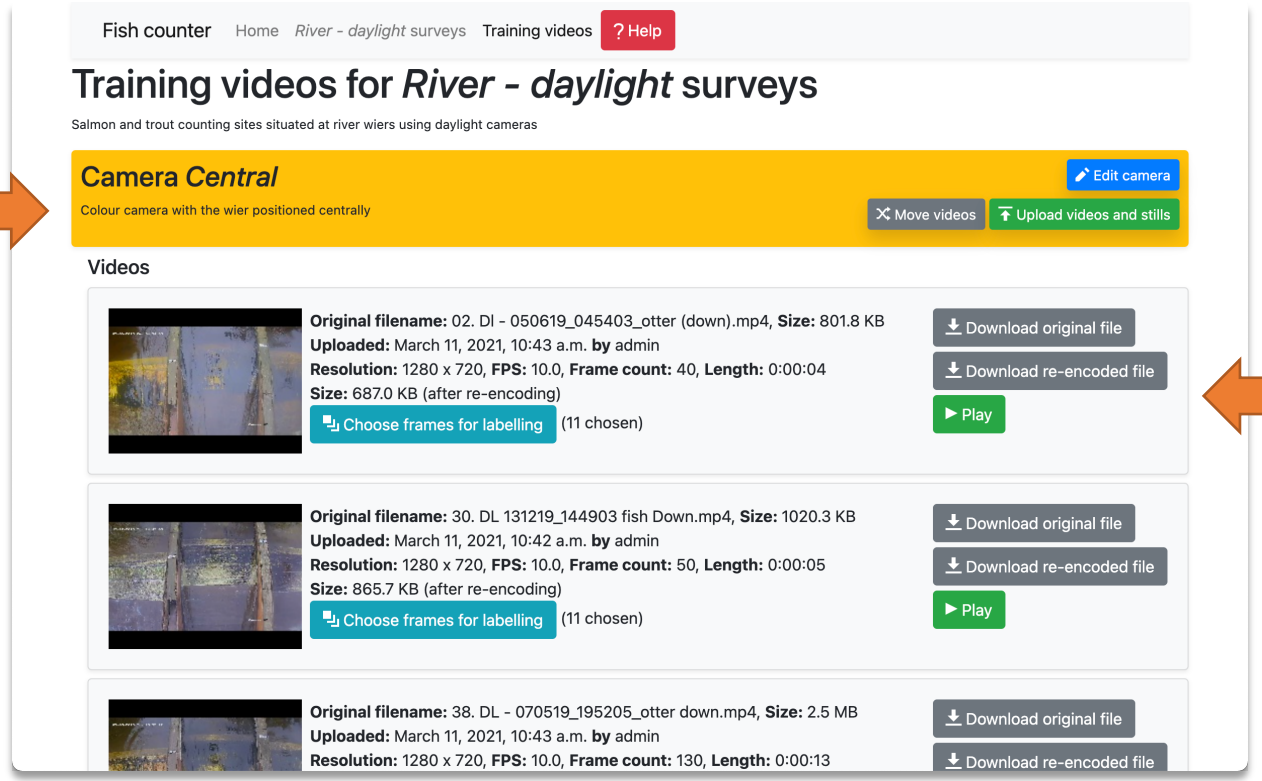
- You need to upload some footage that will be used to train the ML model. From the survey type home, click:



1.a. Upload footage

- The training videos page:

Footage divided by camera



Previously uploaded videos

1.a. Upload footage

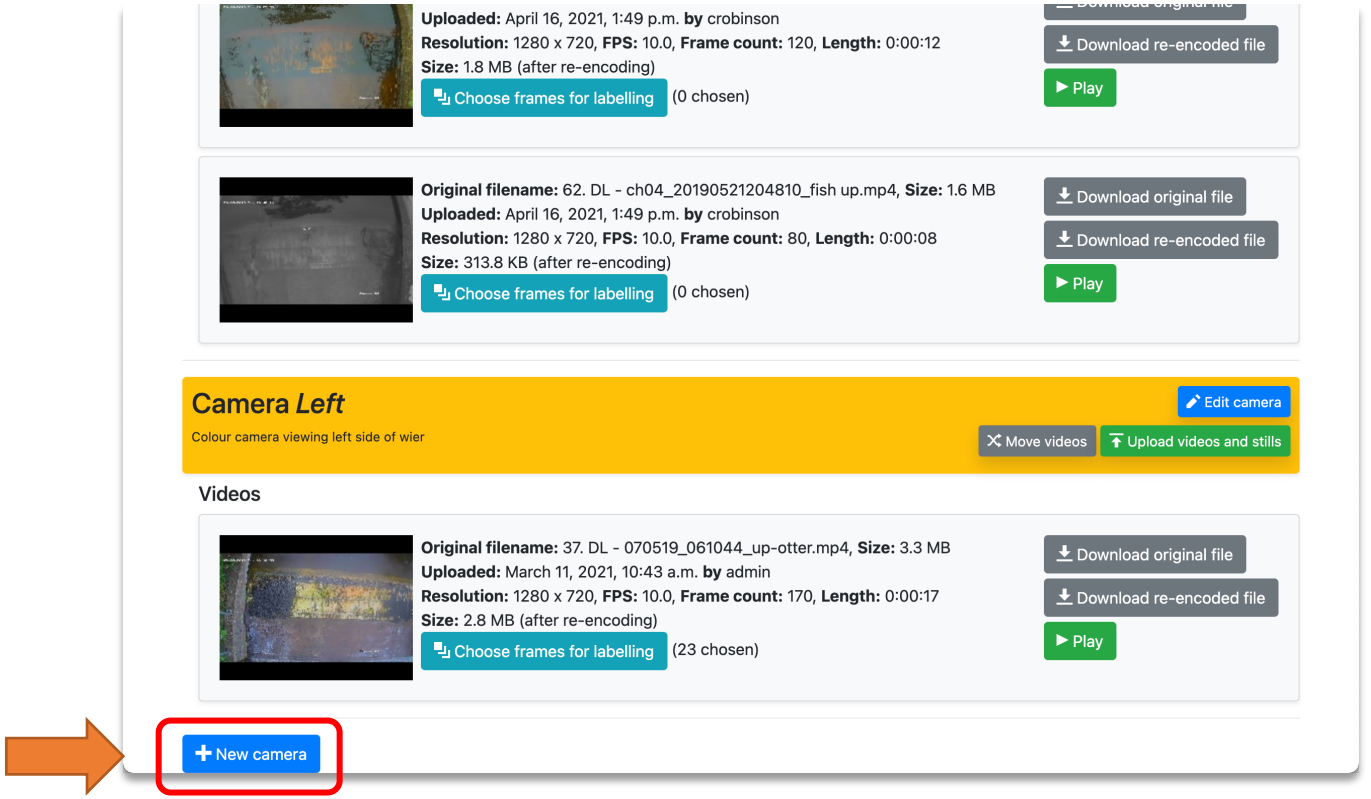
- In the ML intro we mentioned that footage from multiple sites or cameras can be assigned to be in the *training set* or *test set* depending on where it came from
- In this section of the site, you can create a *camera* for each camera/location/site form which footage was obtained

1.a. Upload footage

- In seabed surveys conducted using an underwater camera suspended below a vessel, it probably doesn't make sense to divide footage by the *specific device* that was used to record it
- It probably makes more sense to split by location, time of year, or year
- So make a different *camera* for each of these sources of footage.
  - Apologies for the incorrect terminology here; we selected these terms early in the project and they stuck.

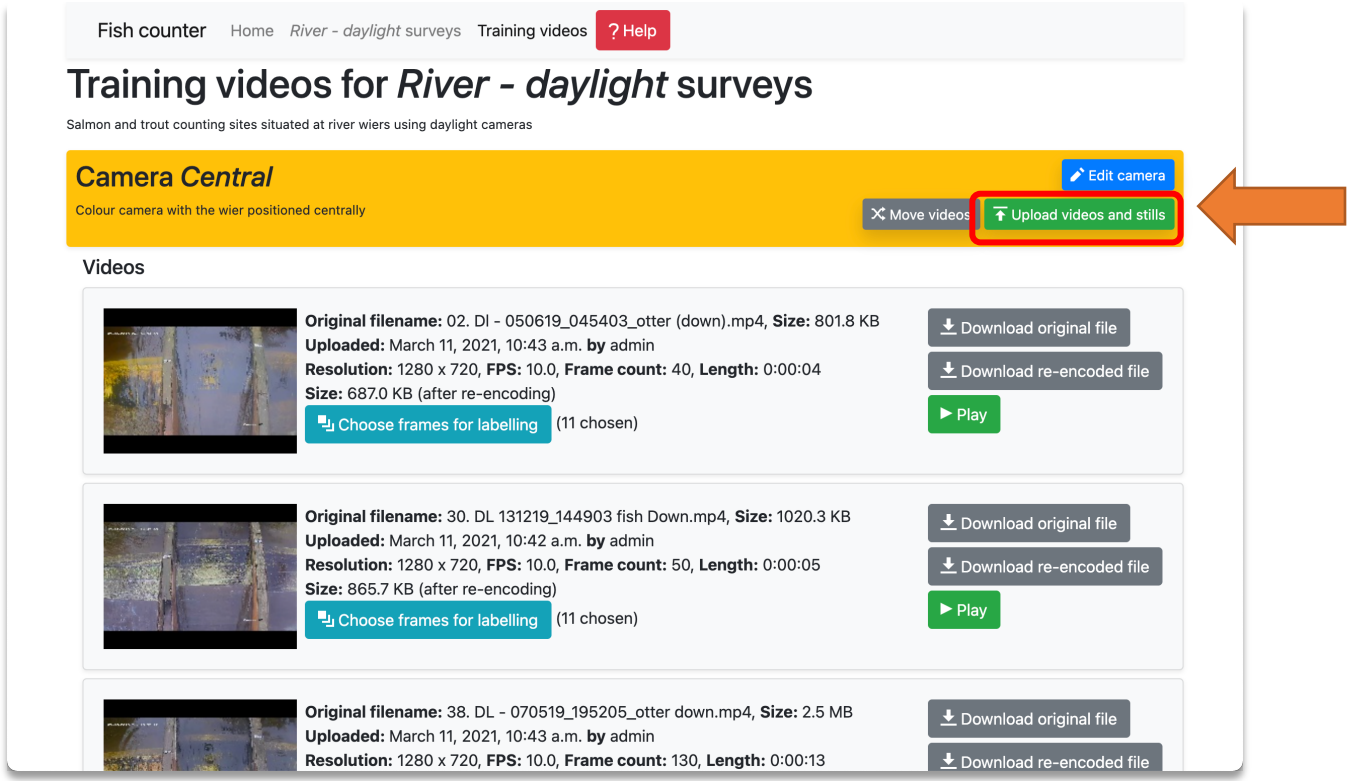
1.a. Upload footage

- If you need to create a new camera, scroll to the bottom of the upload page and click the *new camera* button:



1.a. Upload footage

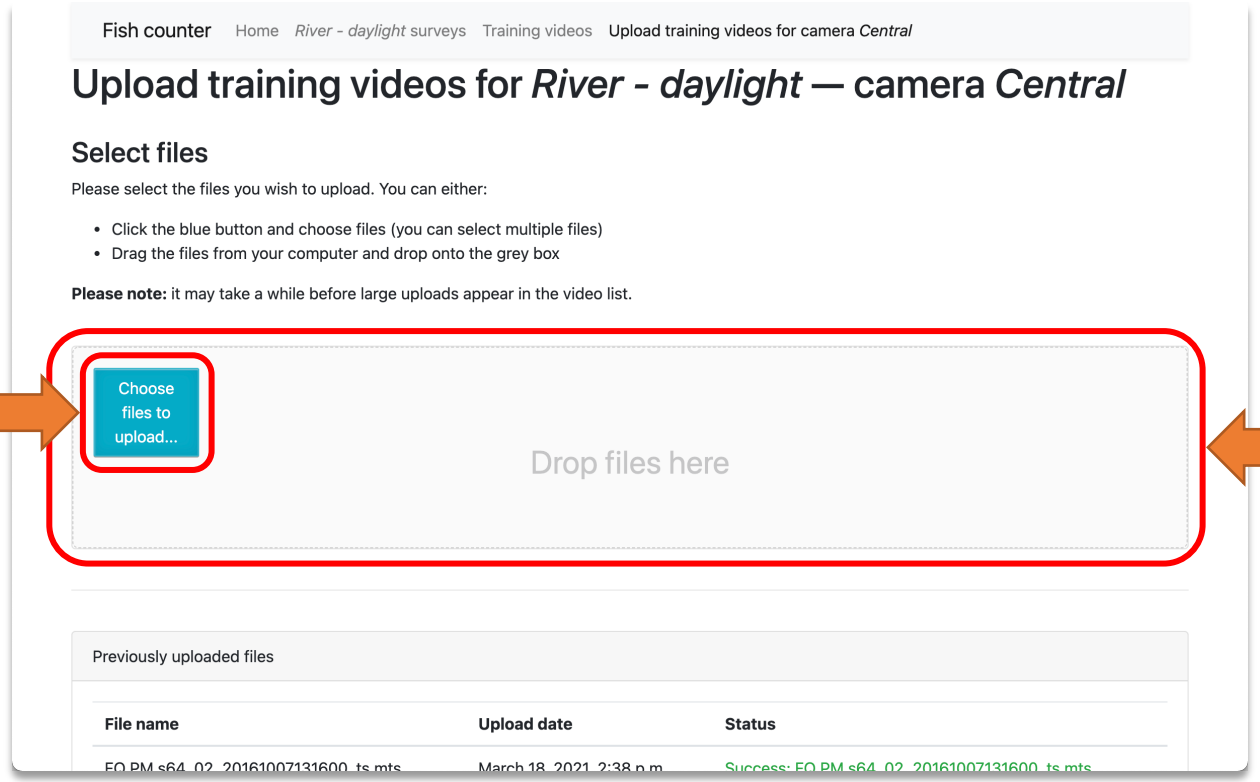
- Choose a camera and click the *upload videos and stills* button in its' yellow header



1.a. Upload footage

- Clicking *upload videos and stills* button takes you here. You can select files from your computer to upload. Choose either video files, or still images

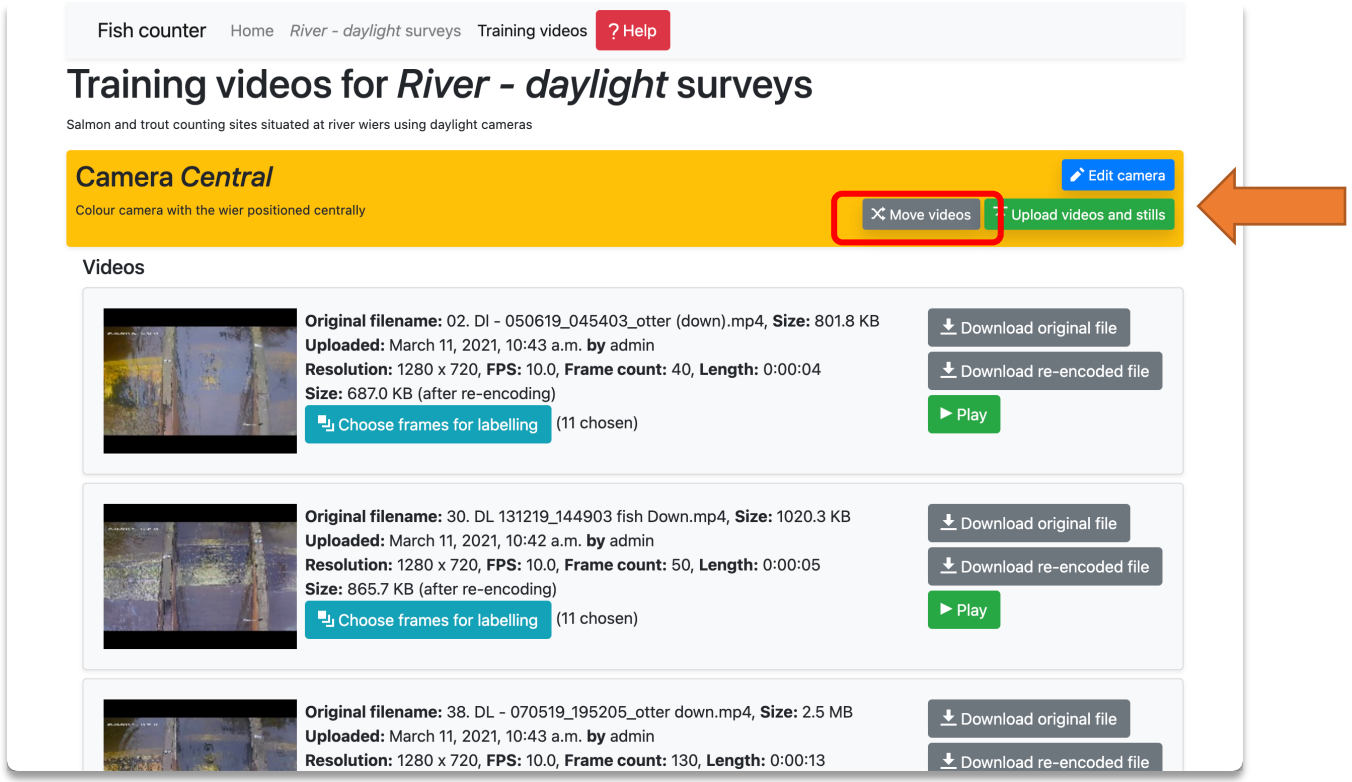
Click here to open a dialog box for locating files on your computer



Or, drag and drop files from your computer to this area

1.a. Upload footage

- *NOTE*: if footage has been assigned/uploaded to the wrong camera; the *move videos* button will allow you to change cameras afterwards:

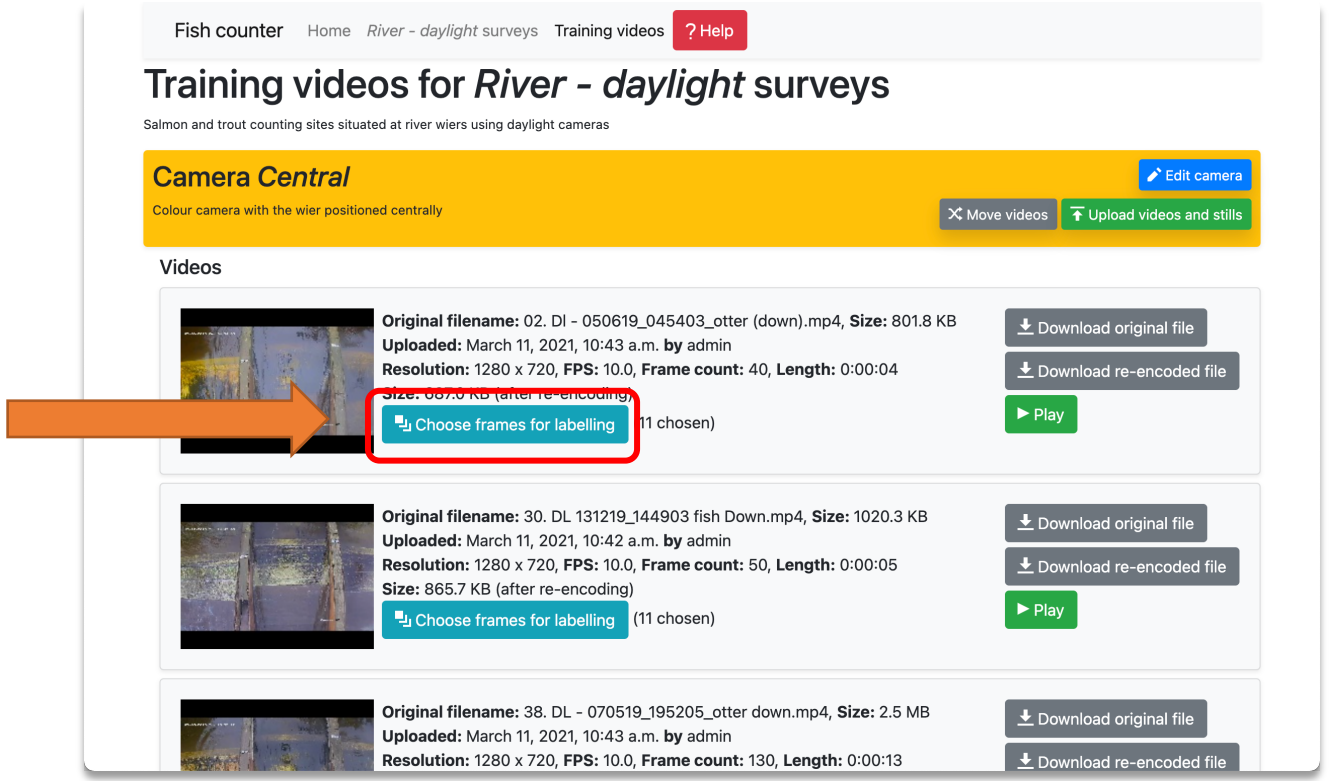


Using AVIMS

1.b. Select frames for labelling

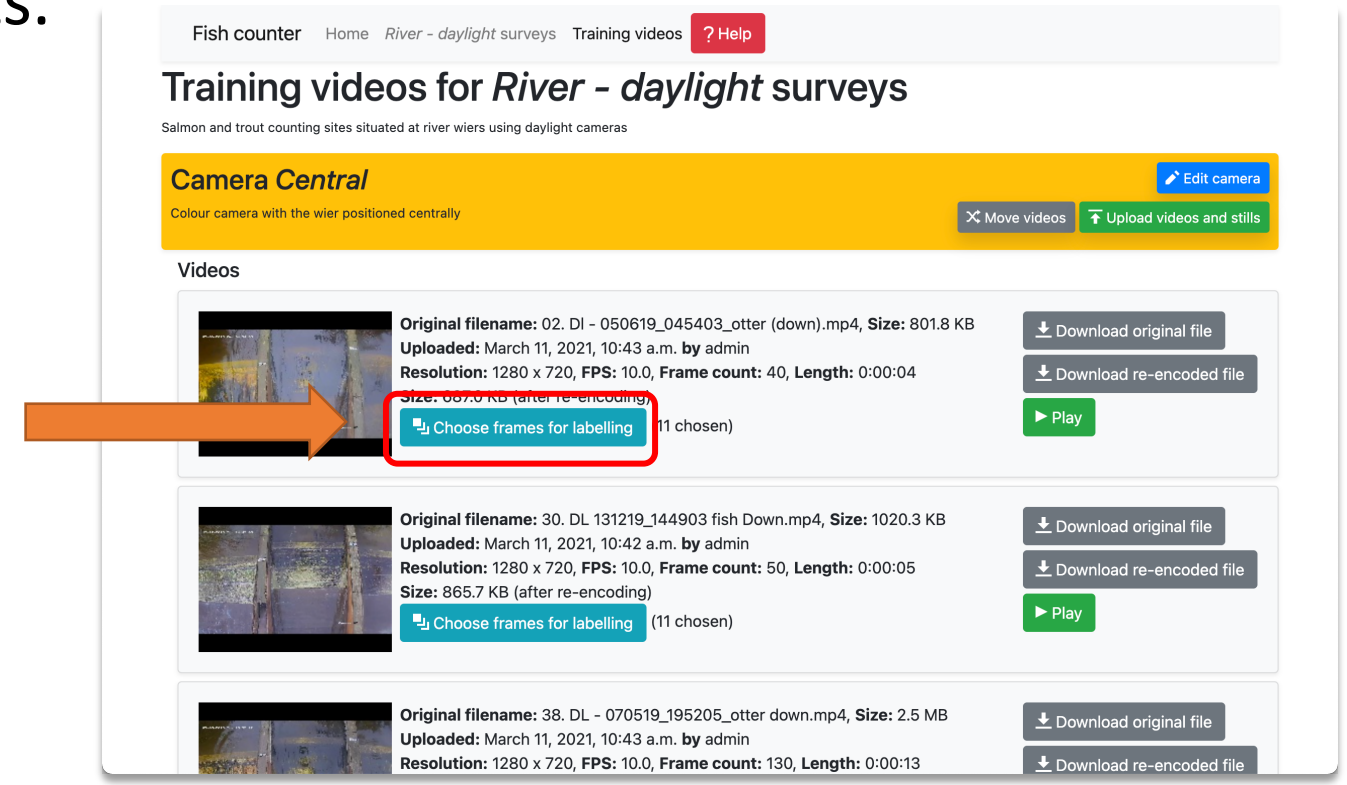
1.b. Select frames for labelling

- From the training videos page, click the *choose frames for labelling* button adjacent to the video you want to select frames from:



1.b. Select frames for labelling

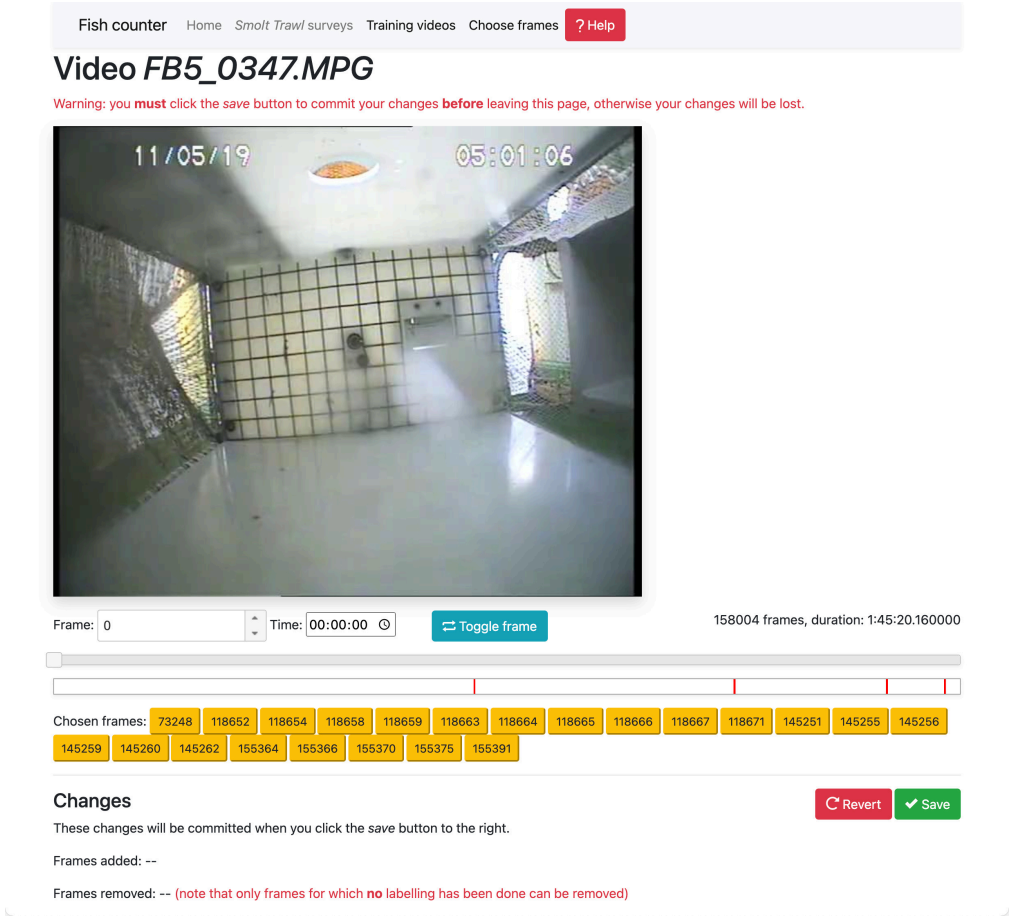
- Note the *(11 chosen)* to the right of the button; the number of frames chosen from this video. Choose frames from a variety of videos for best results.



1.b. Select frames for labelling

Use the *frame chooser* to choose frames for labelling in step 3.

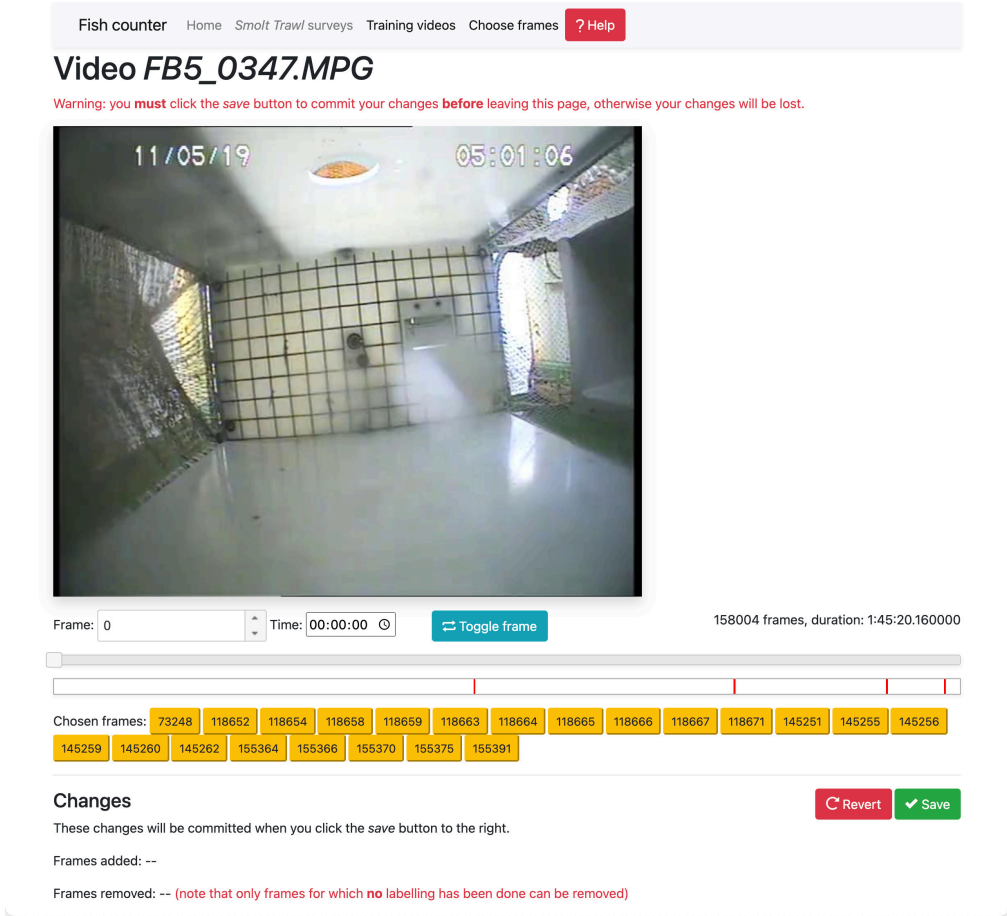
Pick frames that are as representative as possible of the footage you intend to analyze.



1.b. Select frames for labelling

Rules of thumb:

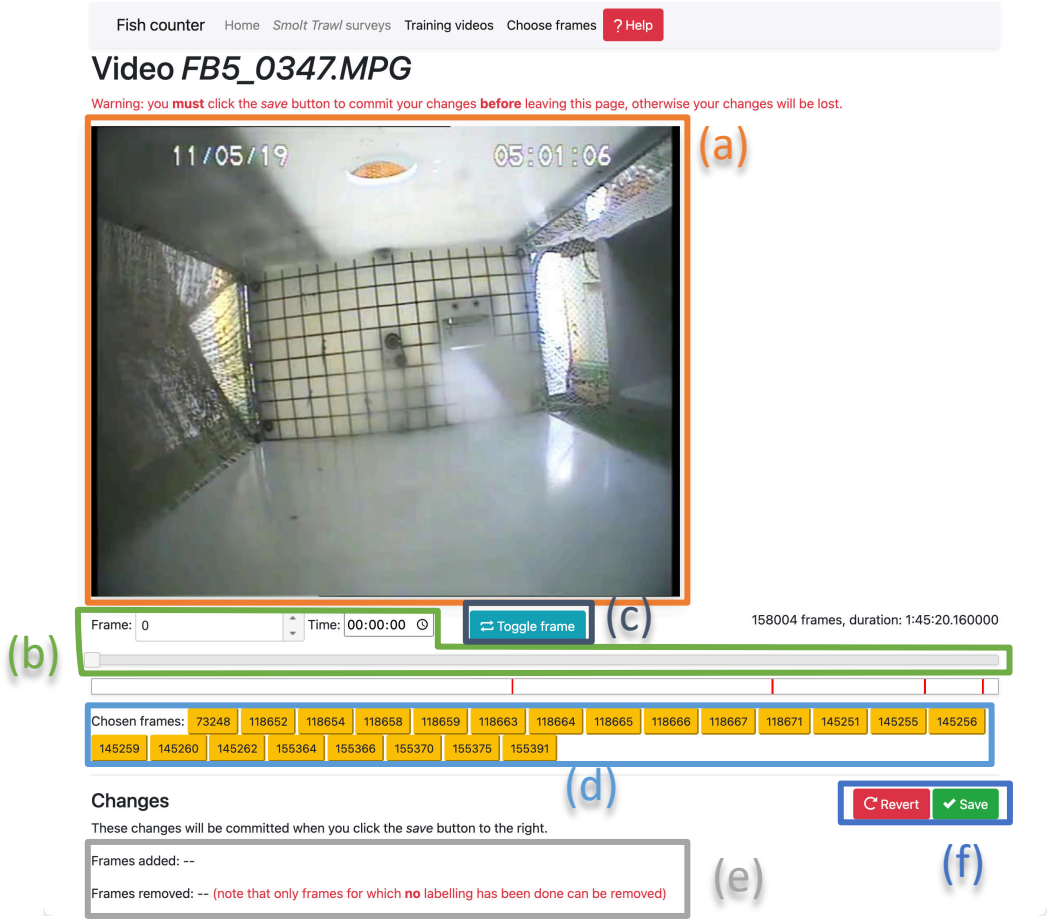
- Choose frames such that as many examples of each kind of creature/objects of interest are visible
- Don’t pick too many of the same individual creature/onject
- Try to choose frames in which the appearances (pose, size, angle, colour, etc.) vary as much as possible



1.b. Select frames for labelling

Frame chooser

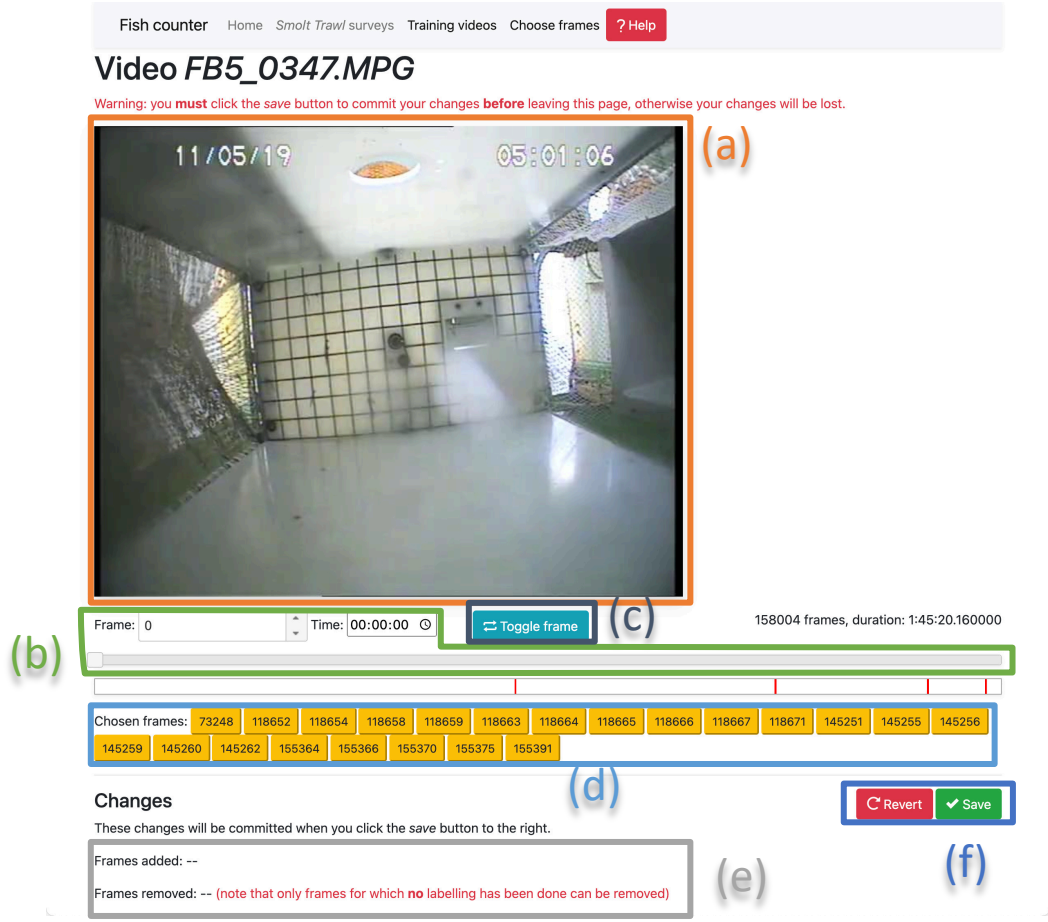
a. Image of current frame
b. Change current frame
c. Add/remove current frame from the list in (d)
d. The list of selected frames
e. The list of changes you have made by selecting/unselecting
f. Buttons to save or cancel/revert



1.b. Select frames for labelling

Work-flow:

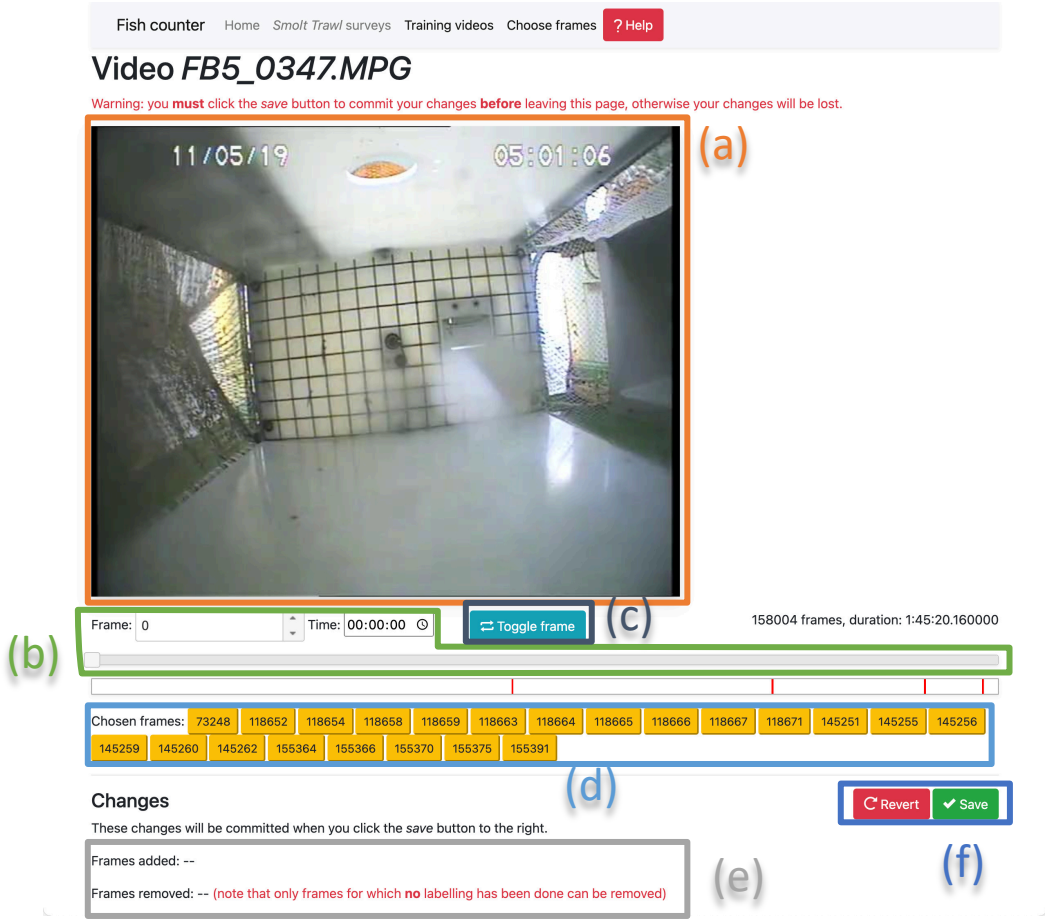
- **NOTE**: you must click the *save* button to save your changes once you are done; navigating back to the previous page will result in your changes to the selected frames being **lost**!



1.b. Select frames for labelling

Work-flow:

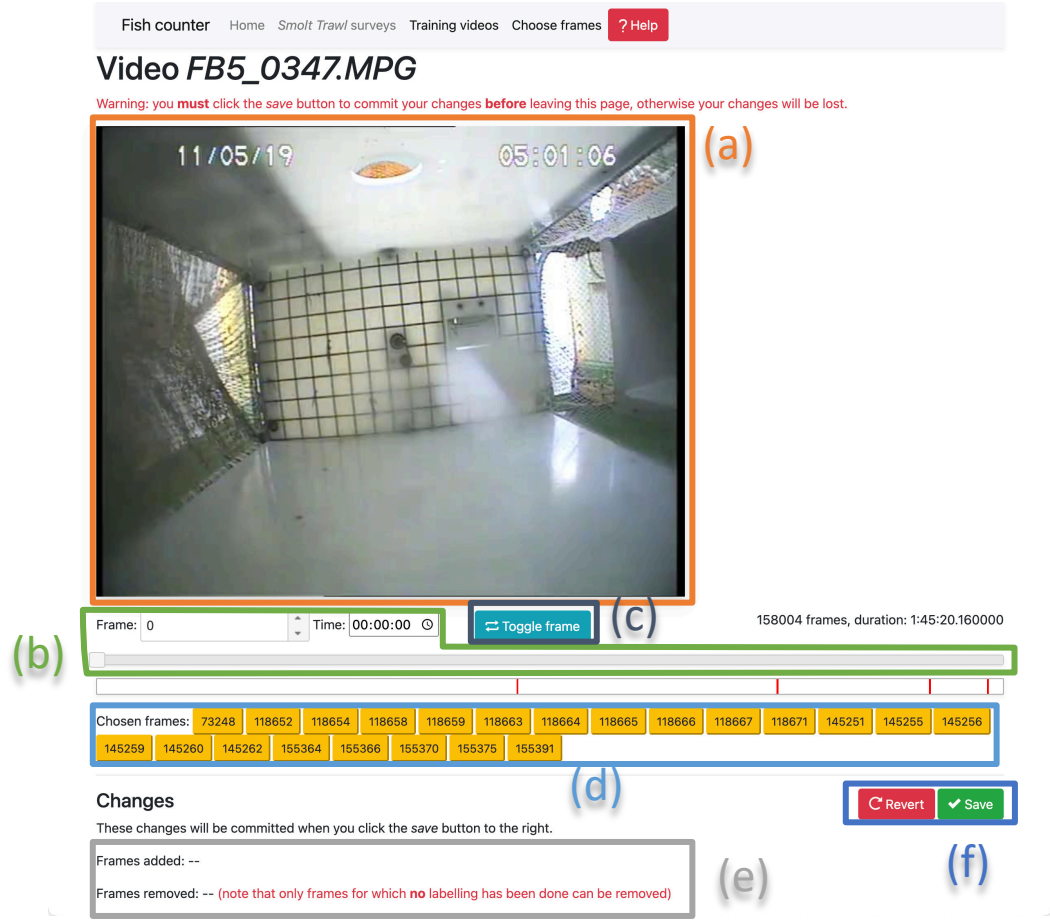
- Use the slider bar in (b) to sweep through the video and find creatures of interest
  - If the video is long, play it in a video player, find the time at which you see see something you want to annotate then enter the time into the time entry box.
  - Alternatively use the frame entry box



1.b. Select frames for labelling

Work-flow:

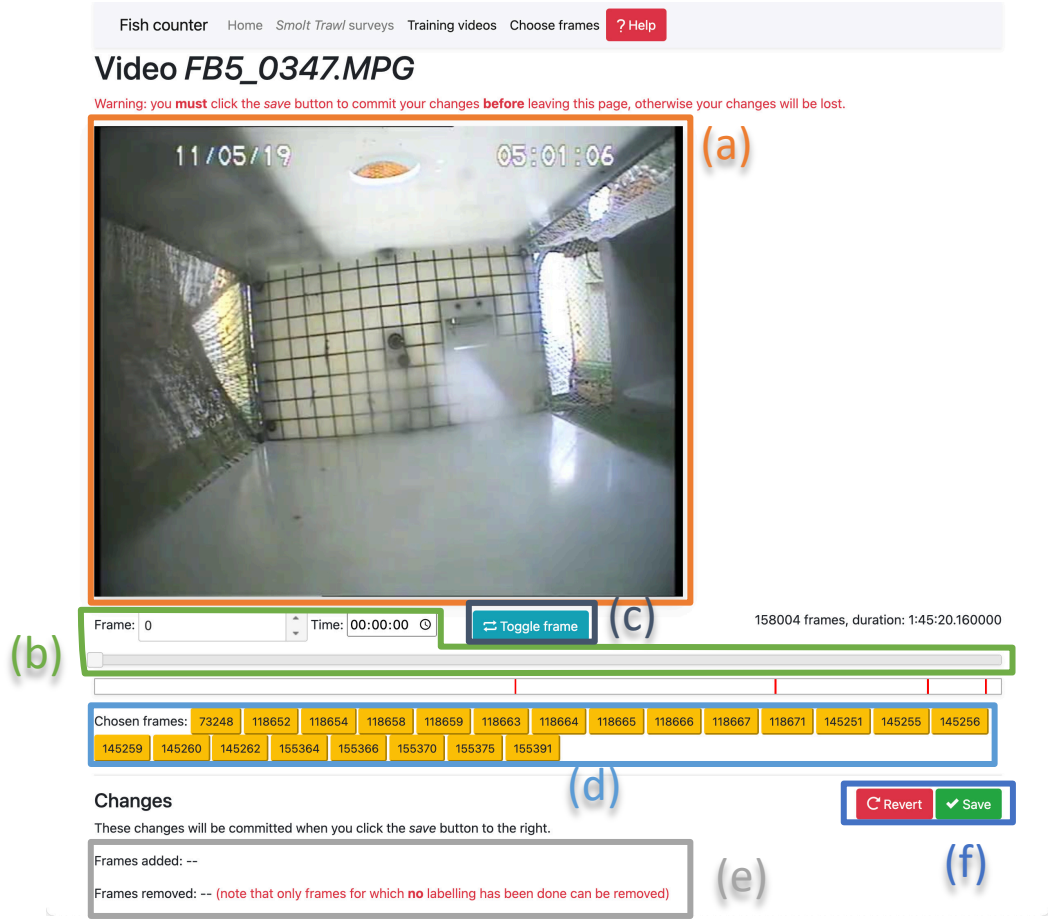
- Click the *toggle frame* button (c)
- A red mark will appear in the bar below the slider
- The frame will be added to the list of chosen frames in (d)
- It will also be listed next to *frames added* in (e)



1.b. Select frames for labelling

Work-flow:

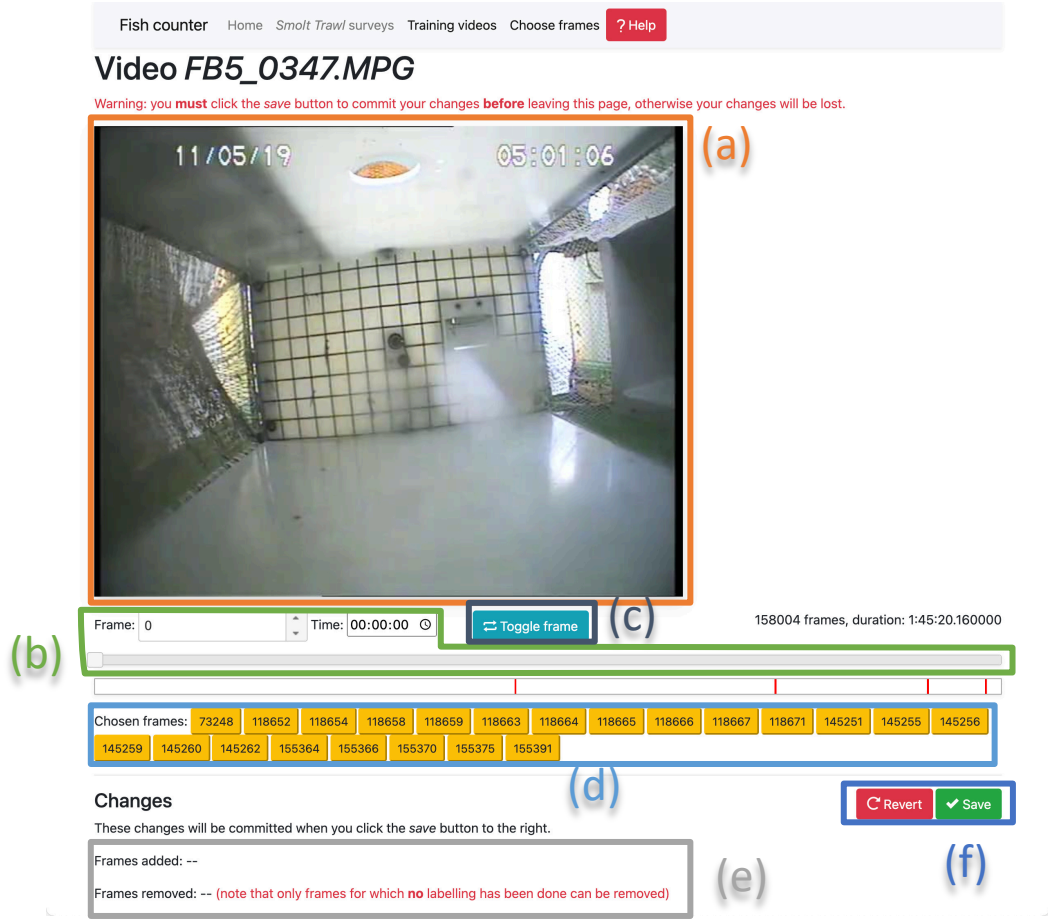
- Frames that you *unselect* will be added to the list after *frames removed* in (e)



1.b. Select frames for labelling

Work-flow:

- If you are happy with your changes to the list of selected frames, click the *save* button in (f)
- If you wish to discard your changes, click *revert*

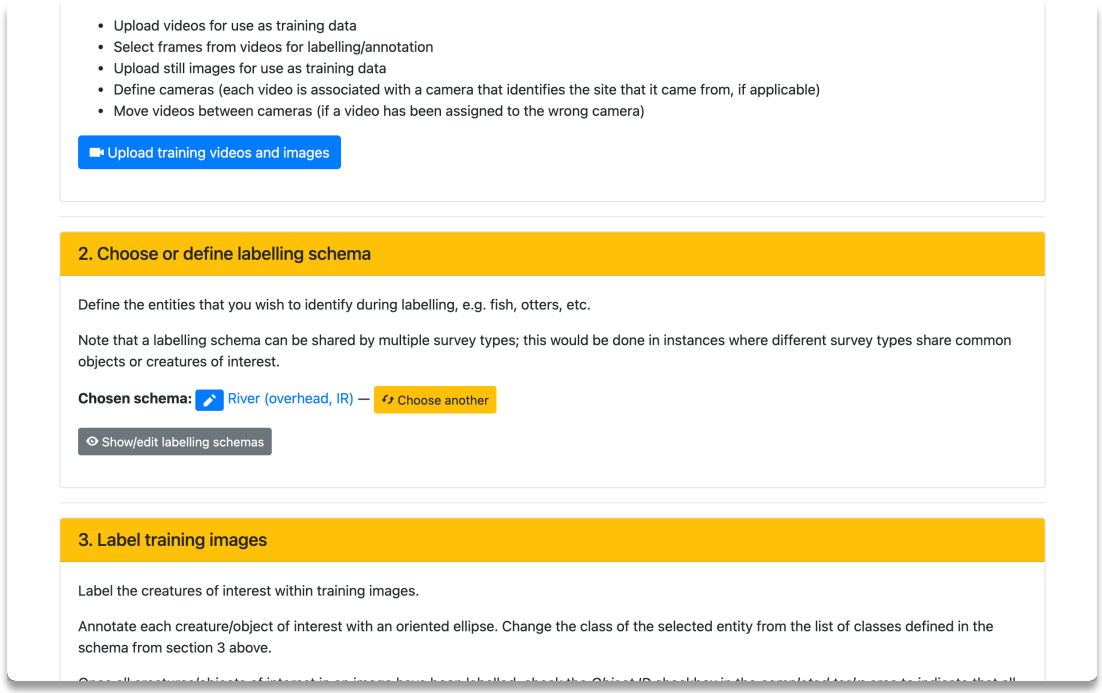


Using AVIMS

2. Labelling schema: what objects/creates are we interested in?

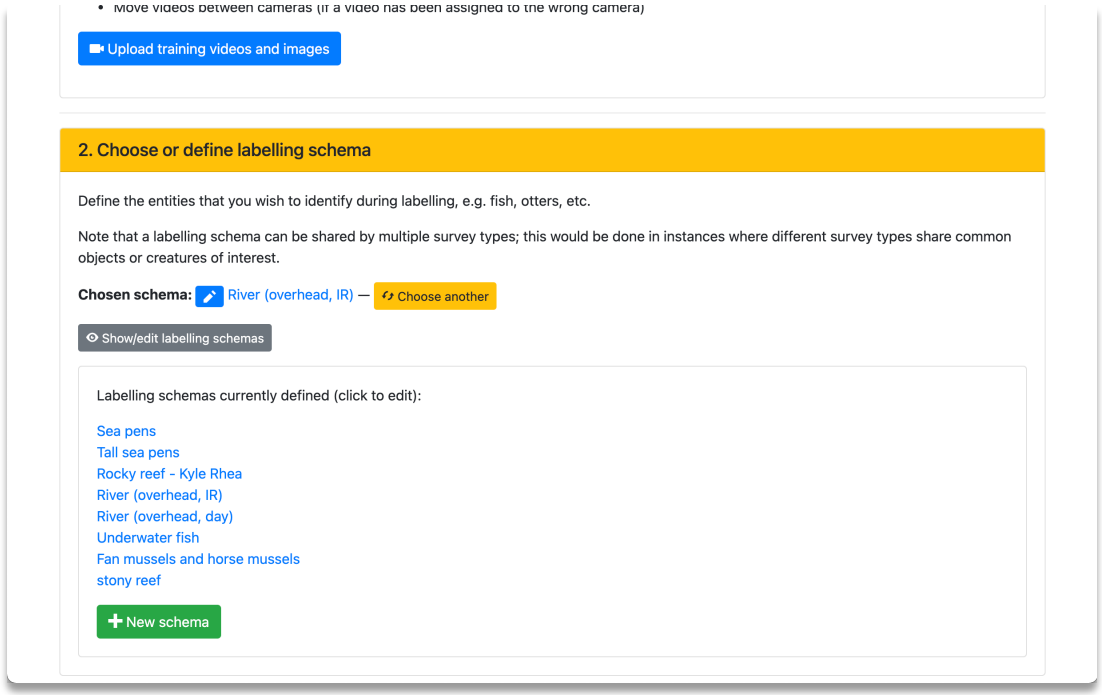
2. Labelling schema

- Section 2 of the survey page lets you change and edit schemas
- You can edit the schema chosen for the survey type by clicking the blue pen button.
- You can switch schemas by clicking the *choose another* button



2. Labelling schema

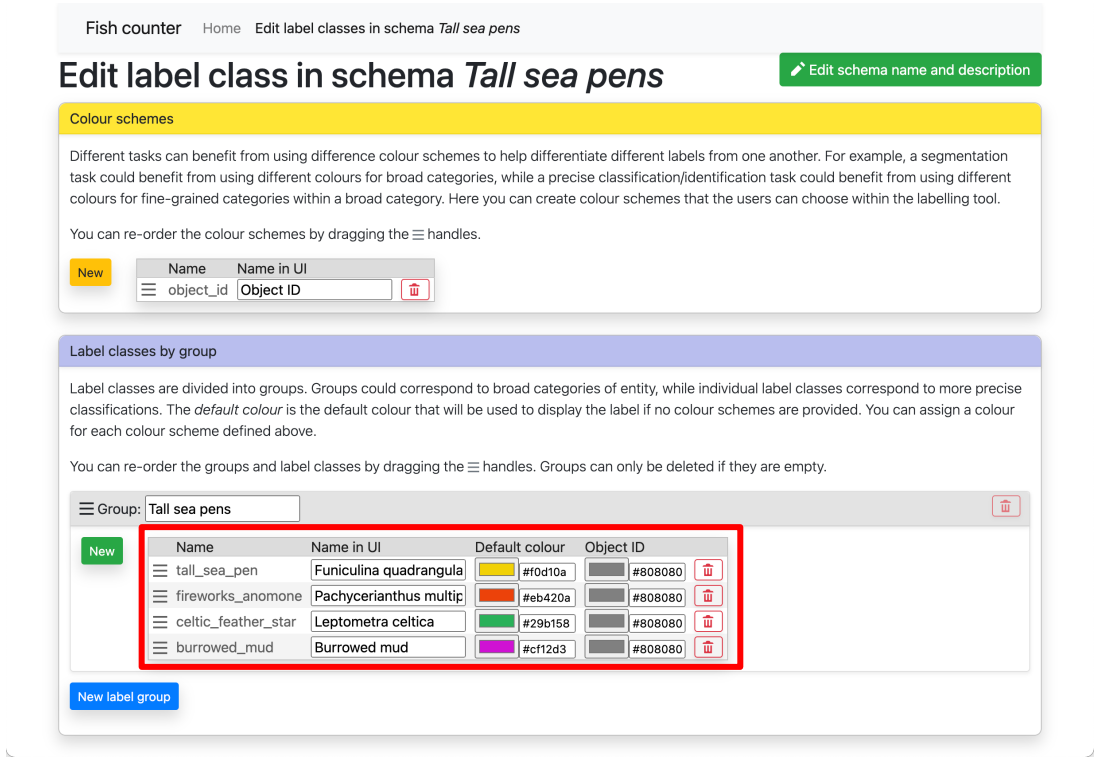
- Clicking the *show/edit labelling schemas* button shows all available schemas, along with the *new schema* button for creating new ones
- If an existing schema covers the all of creature/object types you are interested in, you can use it
  - Multiple survey types can share a single schema



2. Labelling schema

The schema editor

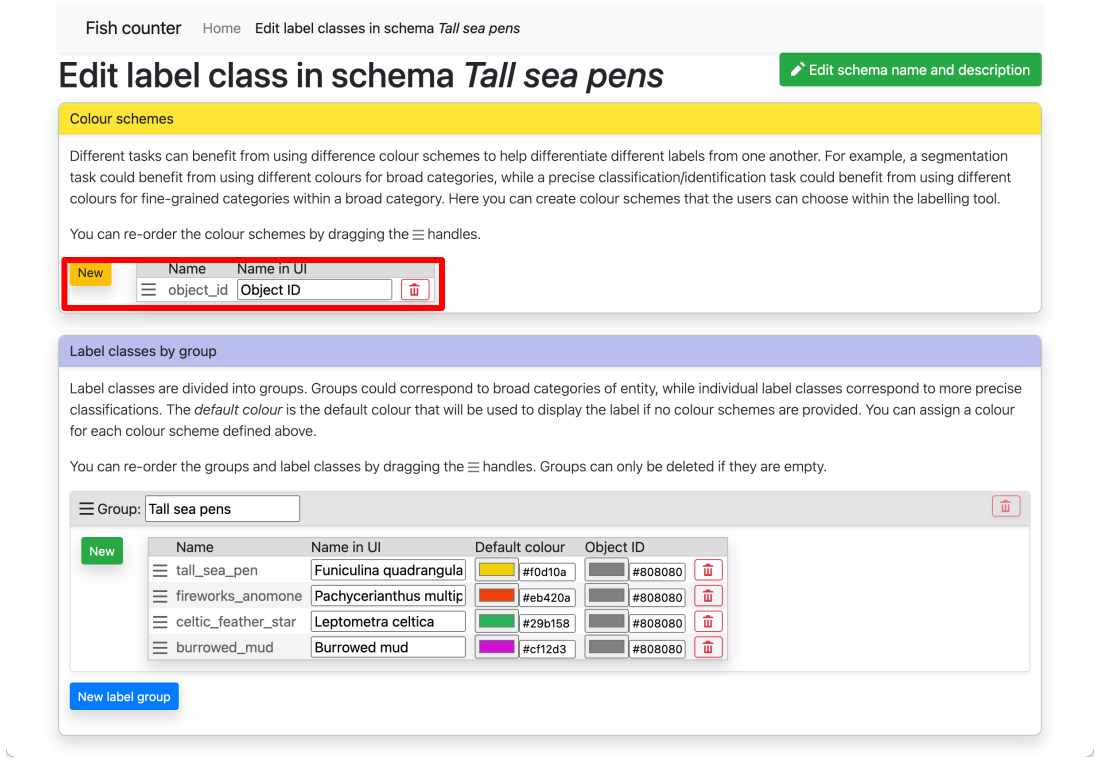
- The list of types of creature/object defined by the schema are listed as *label classes* in the grey table towards the bottom of the image
- You can add as many label classes as you need to a schema



2. Labelling schema

The schema editor

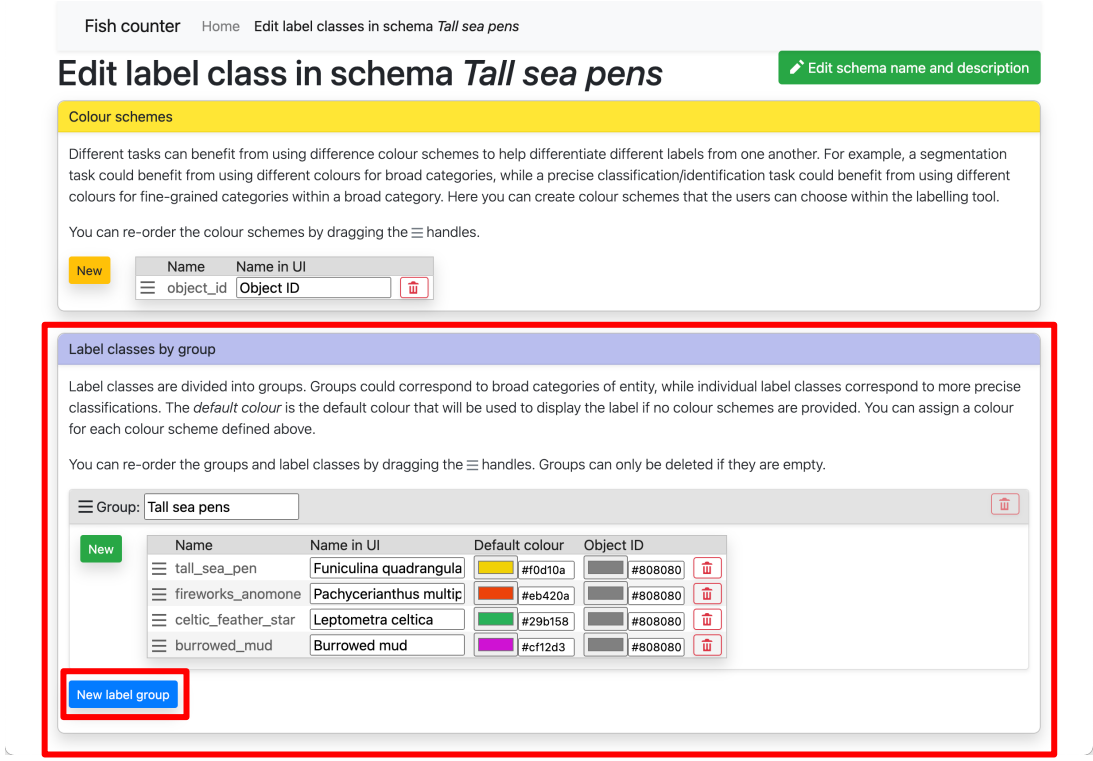
- In the *colour schemes* section, check if there is an entry in the table with the named *object_id* (as seen in the image)
  - If not, click the yellow *New* button and a form will appear. Enter *object_id* for the *name* field and *Object ID* for the *name in UI* field. Then click the green *Create* button that will have appeared in the form.



2. Labelling schema

The schema editor

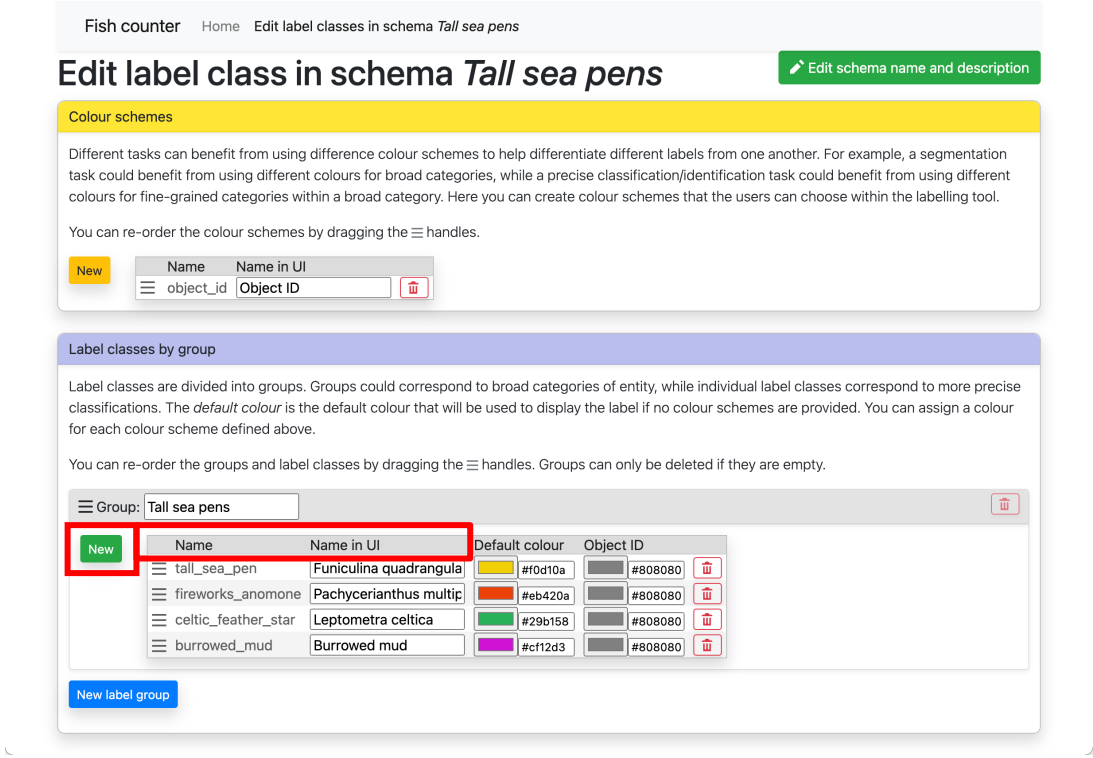
- Mostly you will be adding new label classes and groups in the *label classes by group* section with the lilac header.
  - If you don't see a grey *group* box (like the one titled *Tall sea pens* in the image), click the blue *New label group* button
    - In the form that appears, enter a name and click *create*



2. Labelling schema

The schema editor

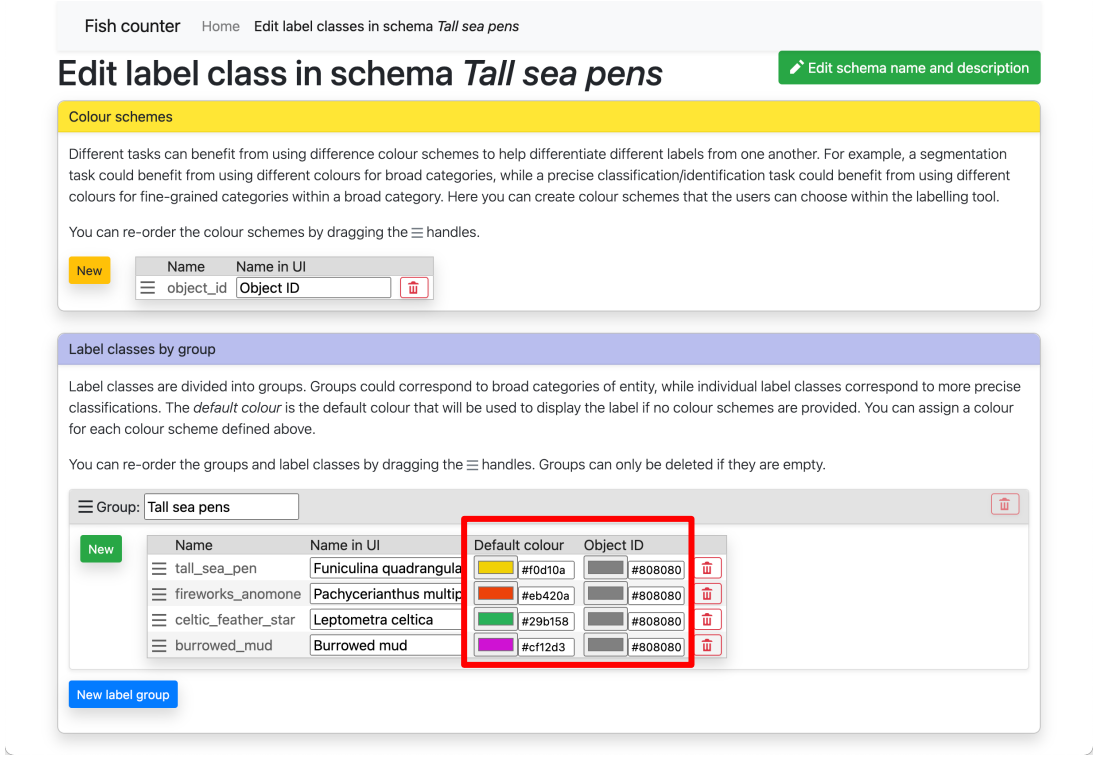
- To add a label class to a group, click the green *New* button. A form will appear. Fill in the *name* and *name in UI* fields and click the *create* button.
  - You can use these fields how you wish
  - *Name in UI* is what you will see in your list of choices in the labelling tool that we will discuss later



2. Labelling schema

The schema editor

- The colours in the *default colour* and *Object ID* colour columns in the label classes table define the colours that will be used to distinguish different types of object in the labelling tool and for identifying object types in the videos resulting from analysis
  - Choose colours that are not too similar to one another

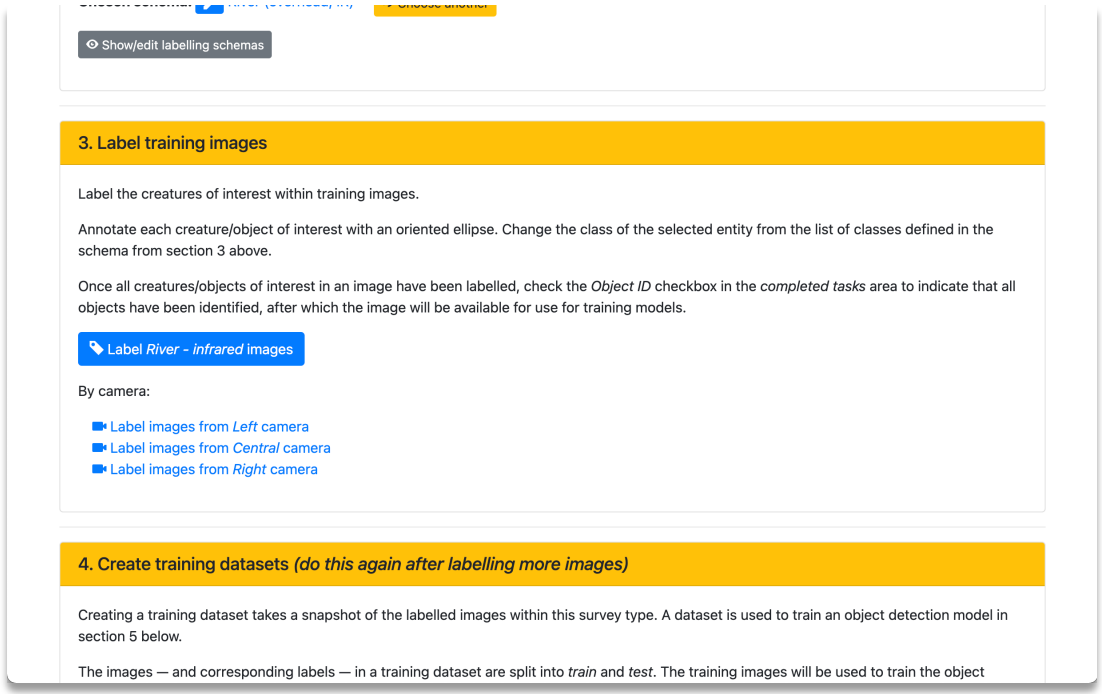


Using AVIMS

3. Annotate objects in training images

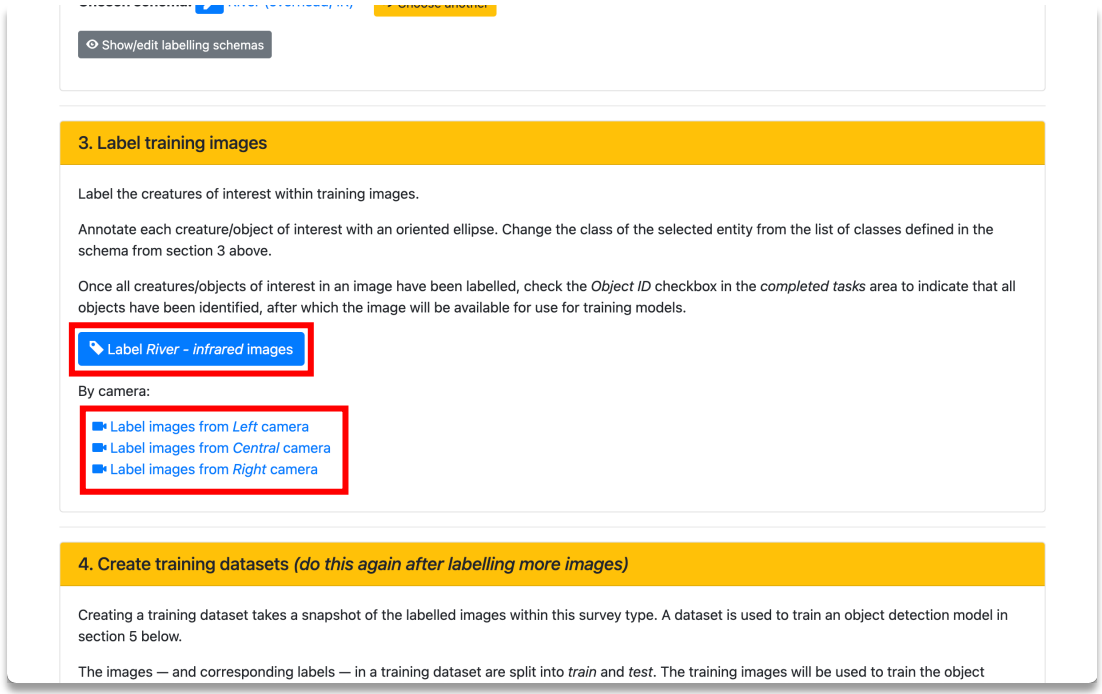
3. Annotate training images

- Section 3 of the survey page has buttons to take you to the annotation part of the site
- Section 3 will be grey and not usable if:
  - You have not uploaded and training videos
  - Or you have not selected any frames for labelling



3. Annotate training images

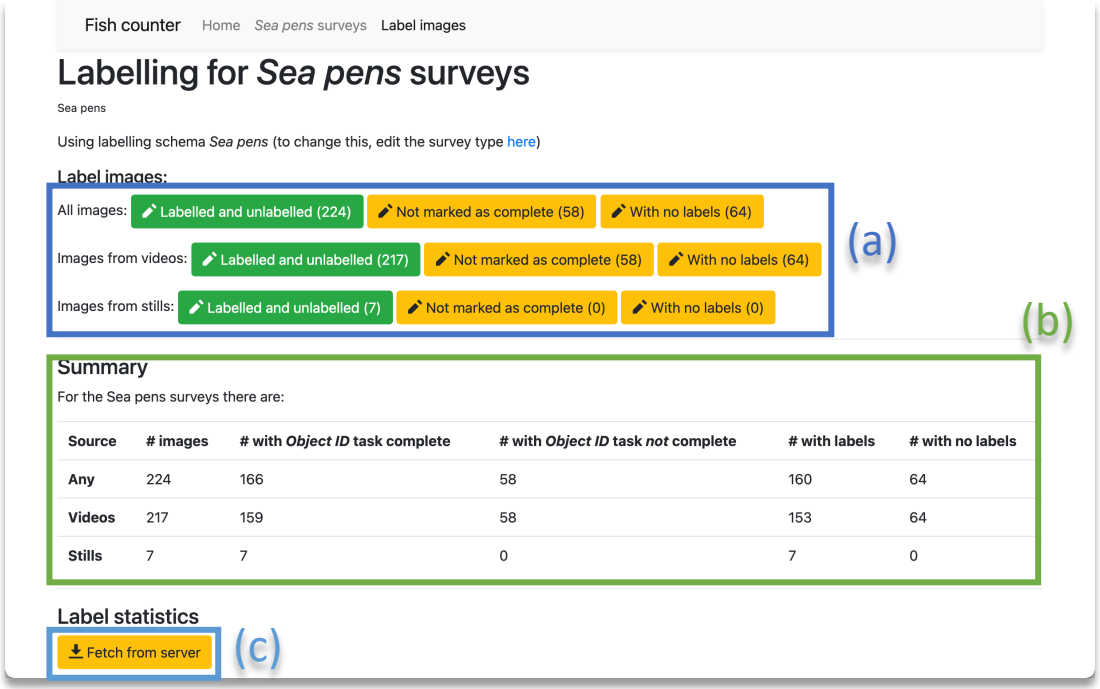
- You can choose to focus on all training images associated with the survey type (the blue button) or focus on images from each camera (the links below)
  - Clicking either the button or one of the links will take you to...



3. Annotate training images

The labelling page

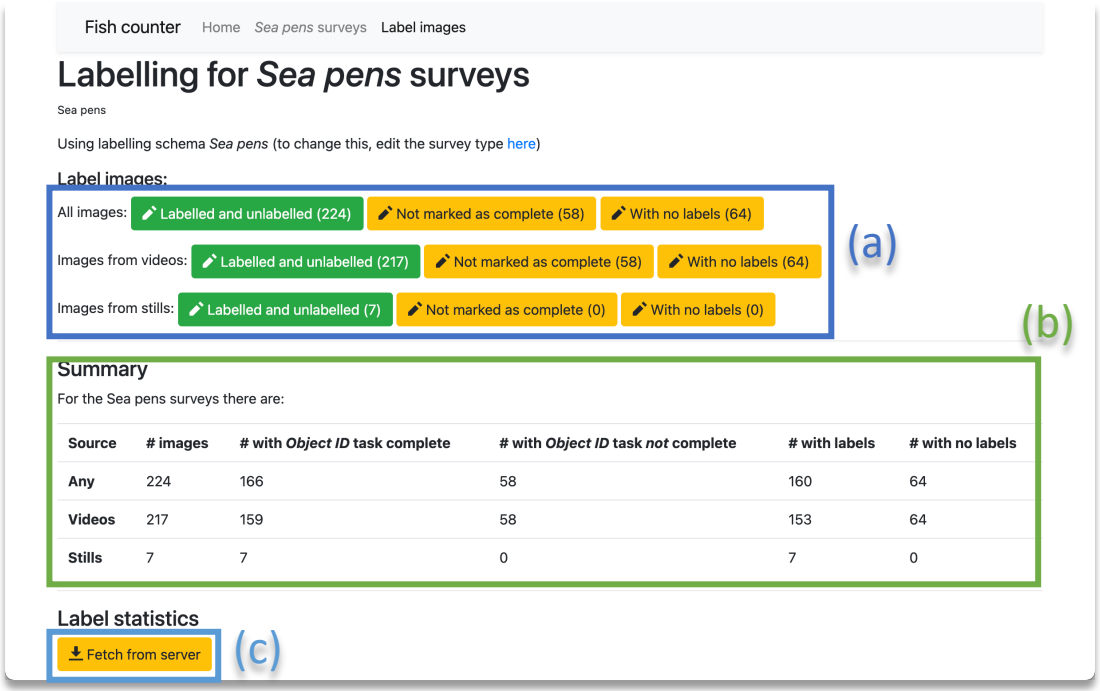
- The buttons in (a) allow you to choose a subset of the images to label. You can choose images for which labelling has not been completed, or images for which no labelling has been done at all, or all images. You can also choose to focus on images from videos or stills.



3. Annotate training images

The labelling page

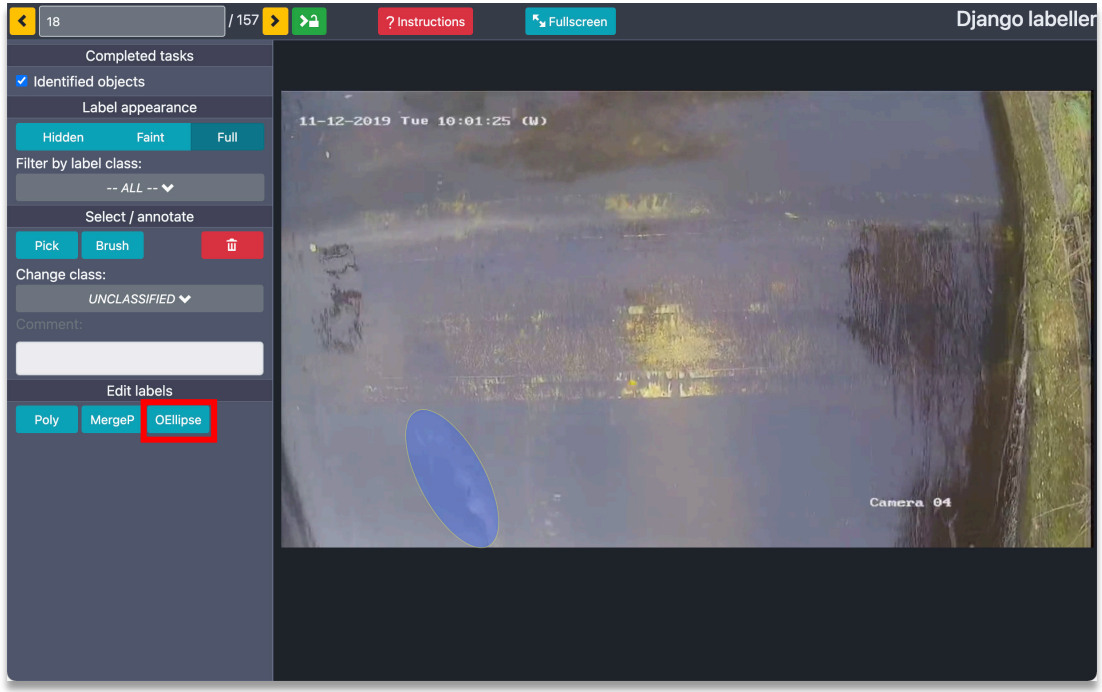
- The summary in (b) gives a brief breakdown of the images
- Pressing the *fetch from server* button (c) under the *label statistics* heading will show you how many of each class of creature/object have been annotated, giving you an idea of which classes are more abundant and which need more instances.



3. Annotate training images

The labelling tool

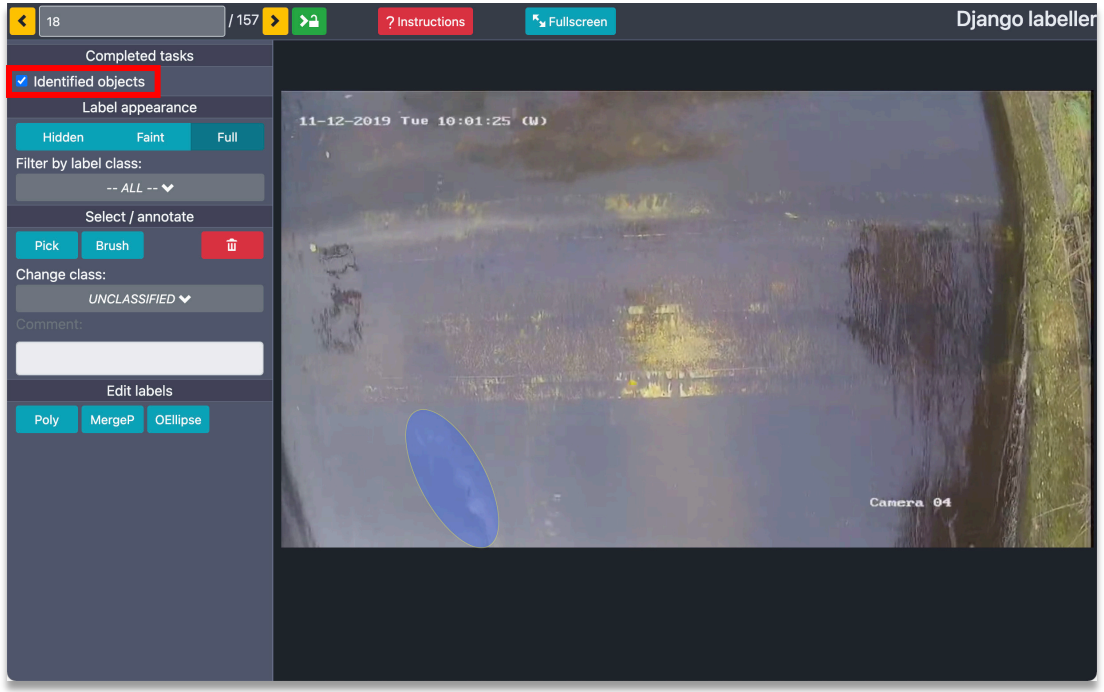
- The *instructions* button will give you detailed help on how to use the tool
- In general, draw oriented ellipses surrounding objects of interest (*OEllipse* button)
  - E.g. the ellipse covering the fish in the bottom left of the image
  - Try to get the ellipse to extend from tip to tail along the length of the creature



3. Annotate training images

The labelling tool

- Once all visible objects have ben annotated, click the *identified objects* checkbox to indicate this.
  - If you don't do this, the image will not be used for training; all images **not yet** marked as *complete* will be **ignored**

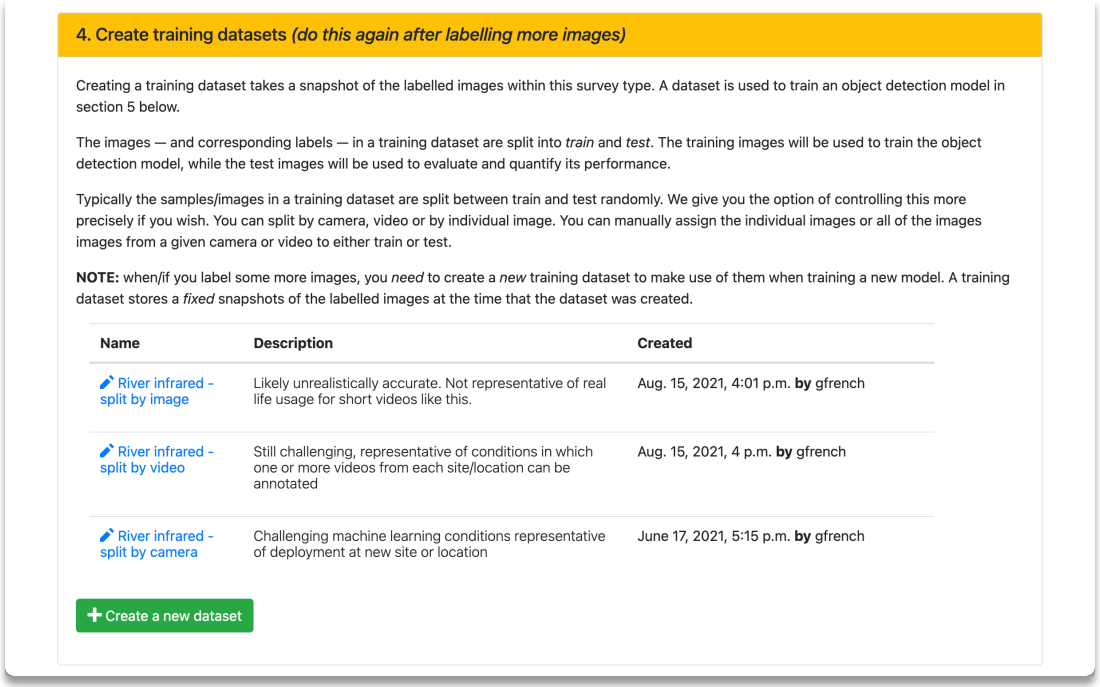


Using AVIMS

4. Build a training dataset

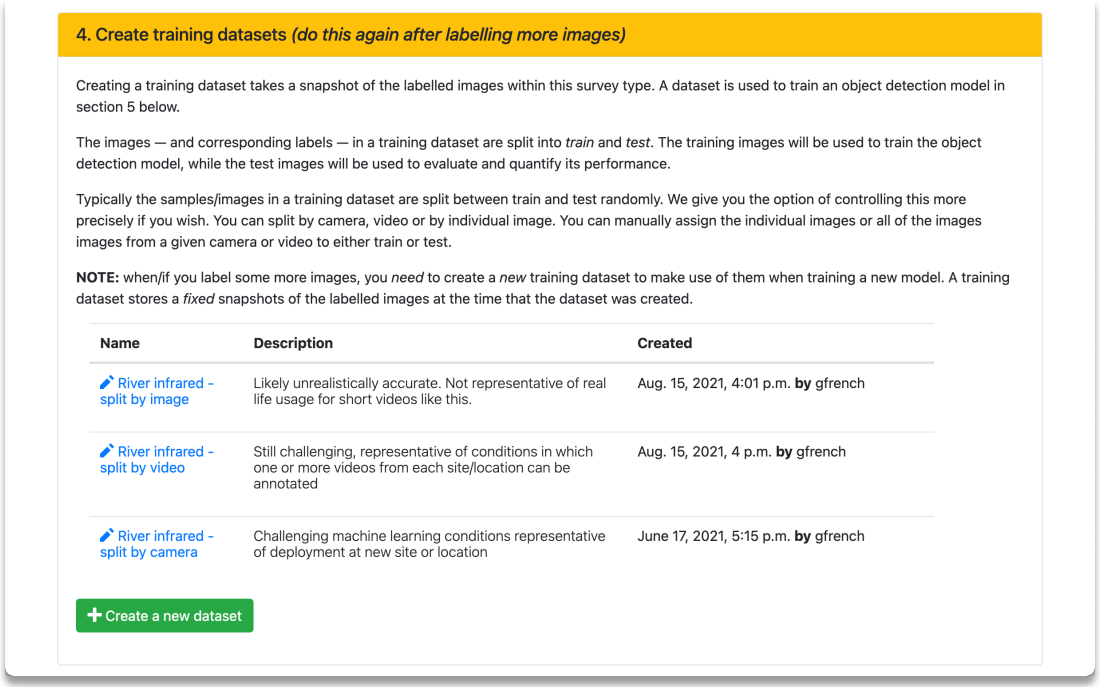
4. Build a training dataset

- Section 4 of the survey page lists the datasets that have been created and when this was done, along with their names and descriptions
- Section 4 will be grey and not usable if:
  - You have not yet labelled any images



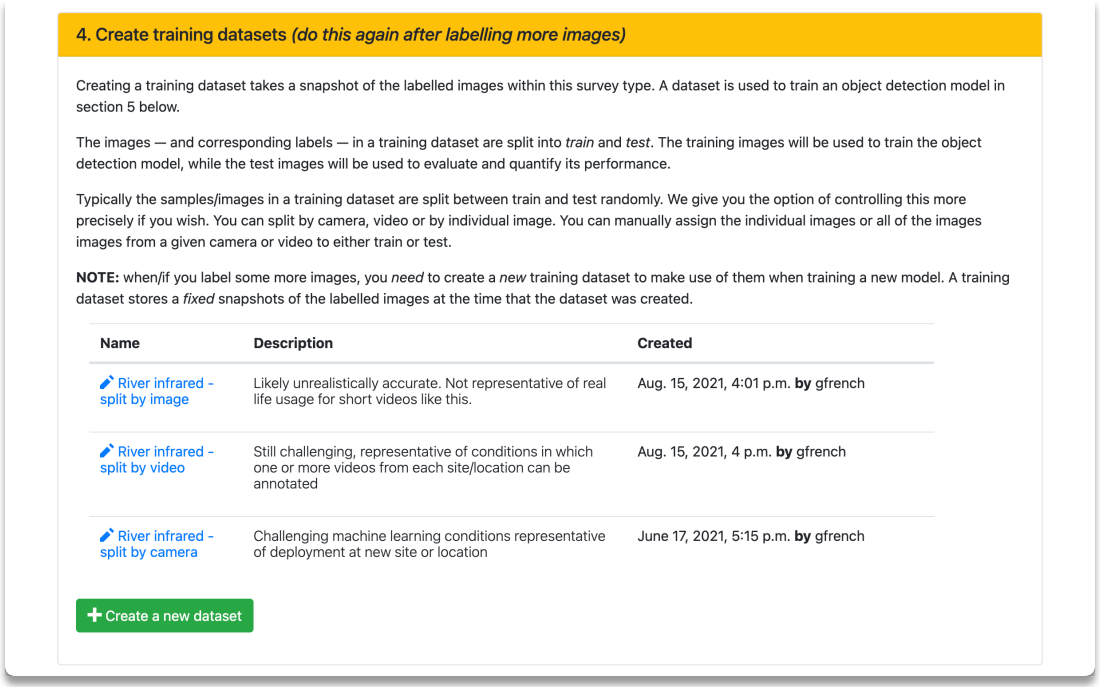
4. Build a training dataset

- Creating a training dataset takes a snapshot of:
  - All of the training images that have been marked as complete
  - Their corresponding labels in the state they are in at the time



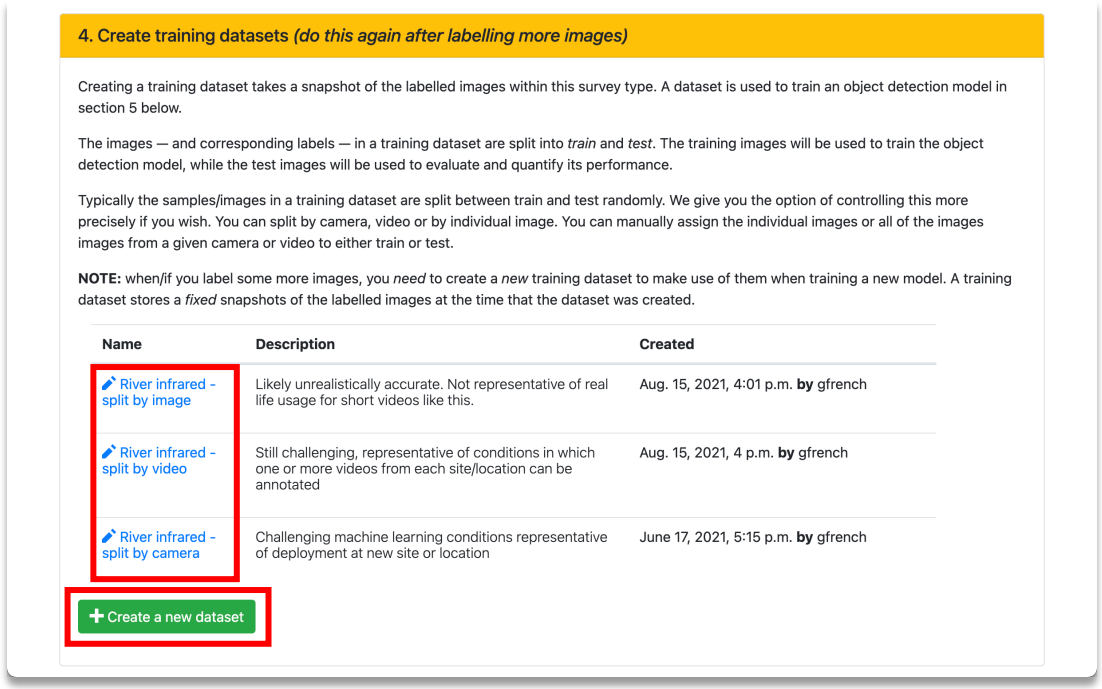
4. Build a training dataset

- If you:
  - Select more video frames for labelling
  - Label more images
  - Change the labels in the images currently marked as complete (e.g. to fix mistakes)
- You will need to create a *new* training dataset to take advantage of them when training a model



4. Build a training dataset

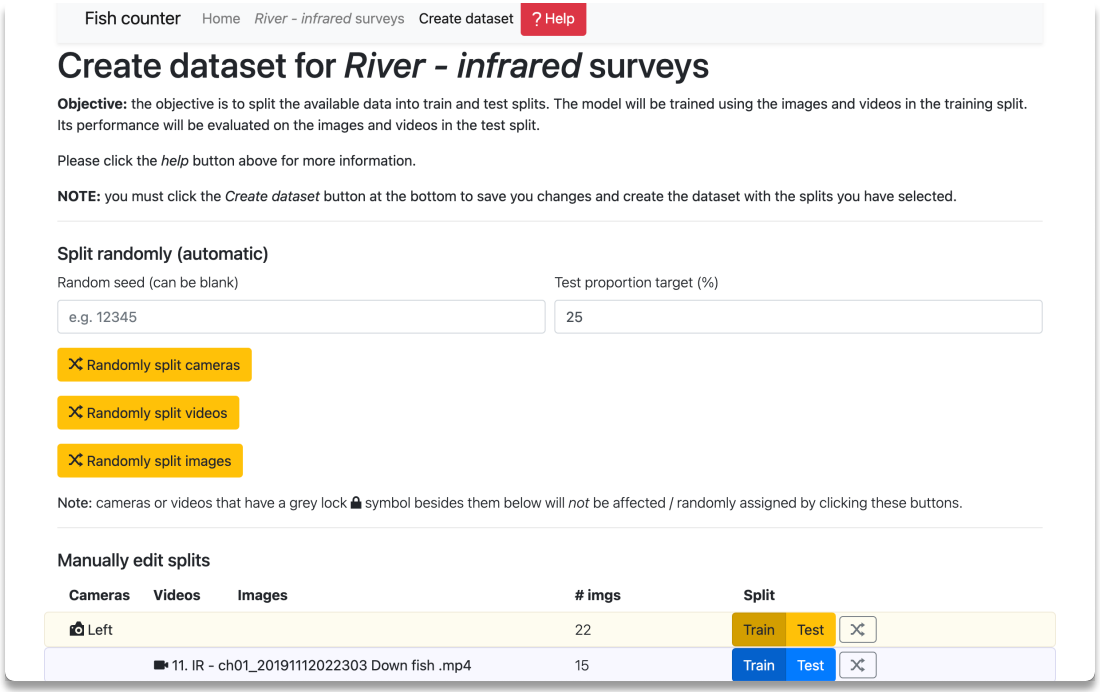
- Clicking one of the blue links will allow you to edit the name and description of the dataset
  - You will not be able to alter its contents as it is a fixed snapshot
- Create a new dataset by clicking the green *Create a new dataset* button



4. Build a training dataset

Dataset builder (random; top)

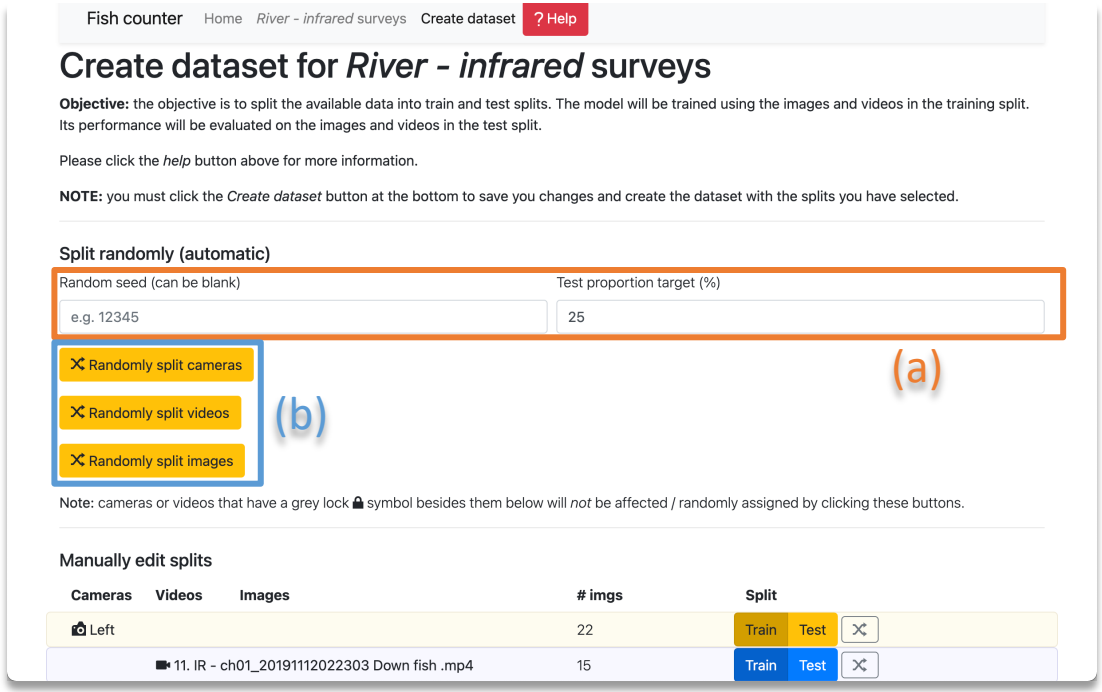
- The dataset builder lets you split the data into *train* and test (see discussion in the ML intro)
- The fastest way of setting the splits is to have the machine split the data randomly



4. Build a training dataset

Dataset builder (random; top)

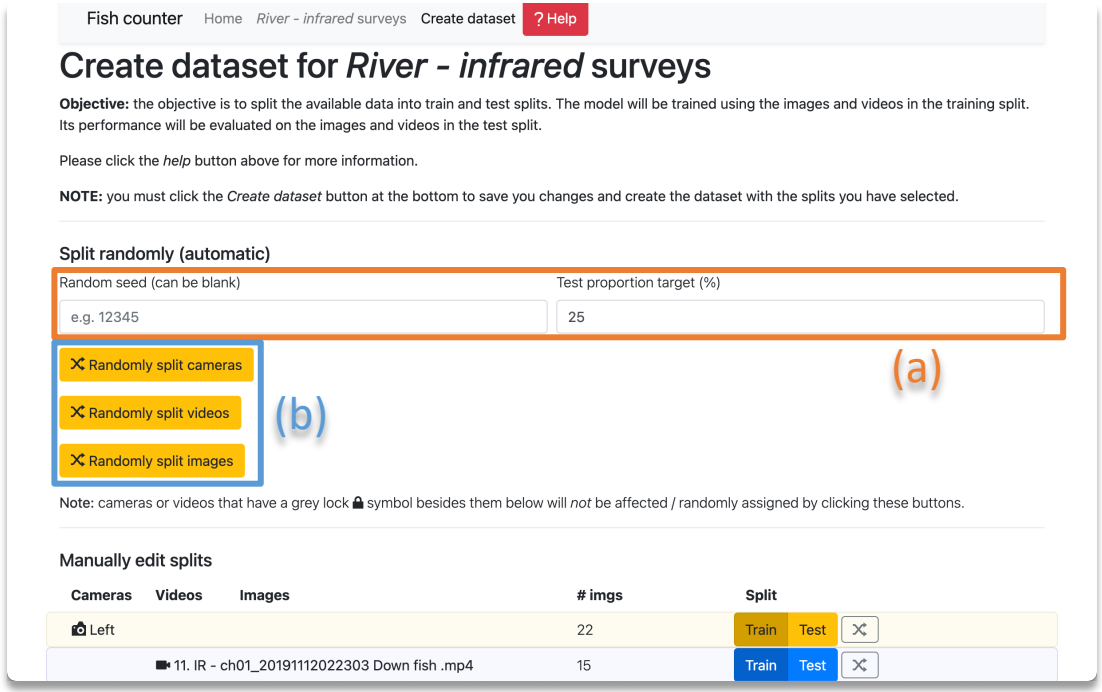
- The buttons in (b) will randomly assign the split by camera, video or individual image
  - E.g. clicking *Randomly split cameras* will assign the same split (train or test) to all of the images from videos from that camera
  - Similarly for *Randomly split videos*



4. Build a training dataset

Dataset builder (random; top)

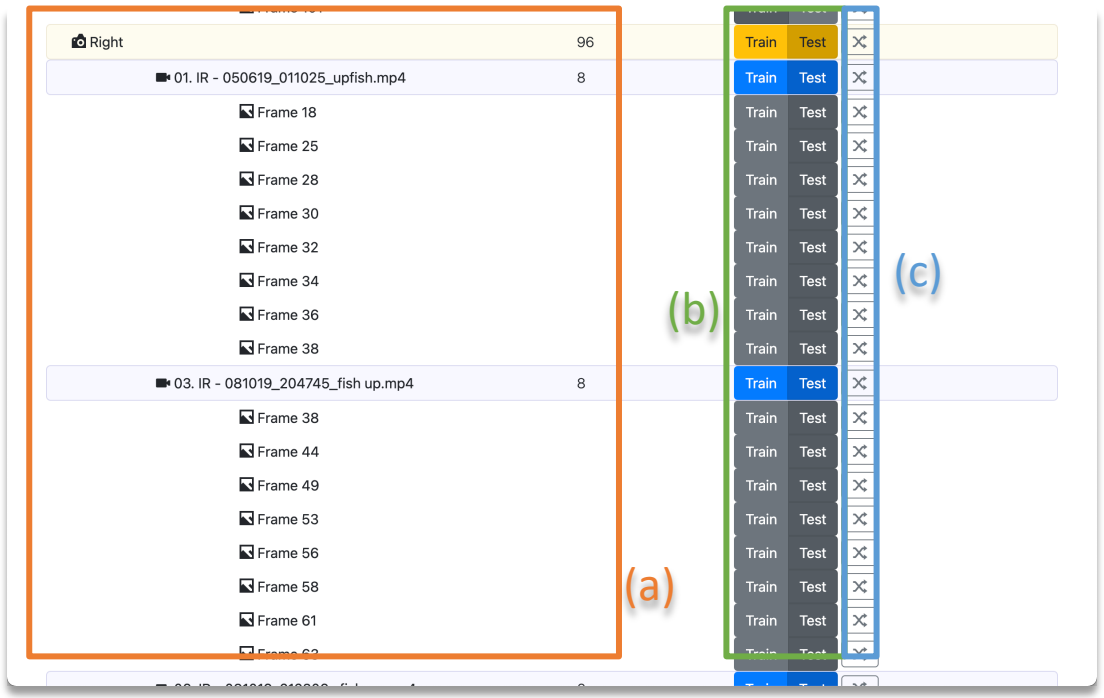
- When splitting randomly
  - You can determine the proportion of the data that will be assigned to the *test* split by changing the *Test proportion target* field in (a)
  - If the *Random seed* field (in (a)) is blank, a different split will be generated each time. Giving a value will control the random split that is used.



4. Build a training dataset

Dataset builder (manual)

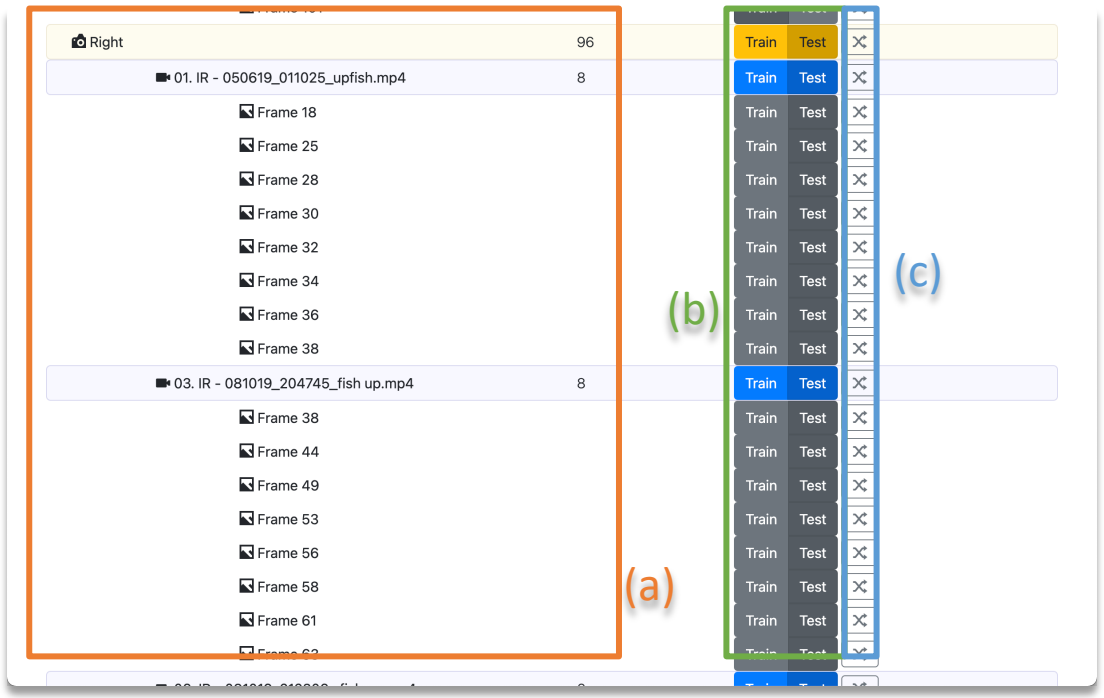
- Further down you will find manual controls for the splits for each image in the dataset
- The images in (a) are divided by camera and video



4. Build a training dataset

Dataset builder (manual)

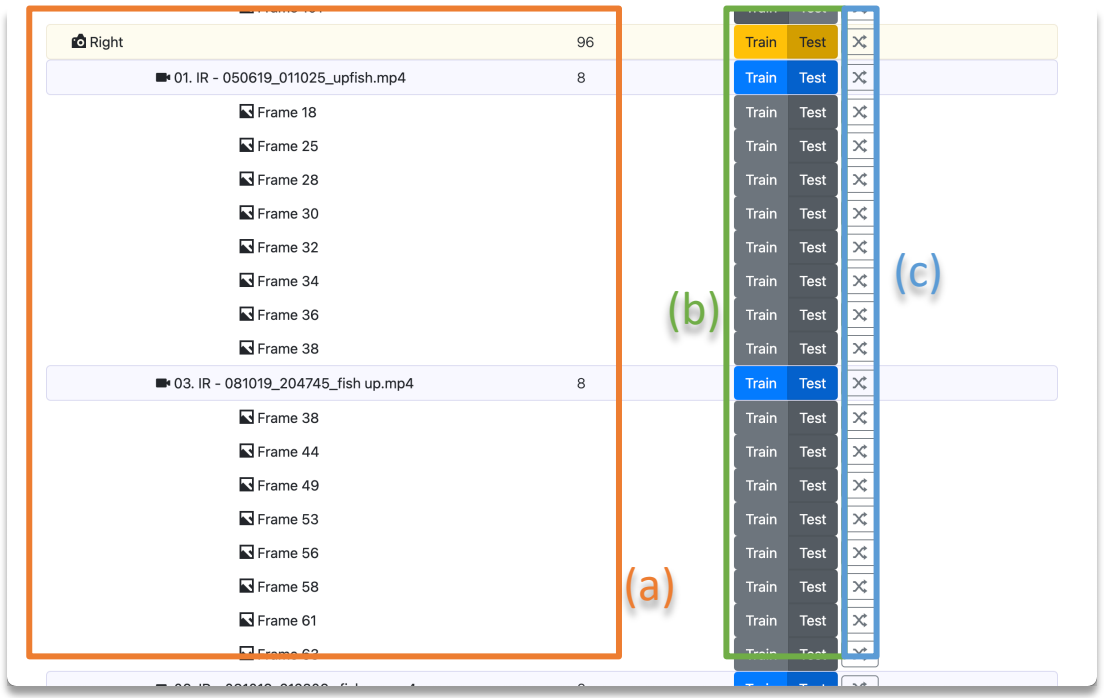
- You can assign images to train or test by clicking the *Train or Test* buttons in (b).
  - Clicking the grey buttons will assign the split of an individual image.
  - Clicking the blue or yellow buttons will assign the split to all images in the given video or camera respectively



4. Build a training dataset

Dataset builder (manual)

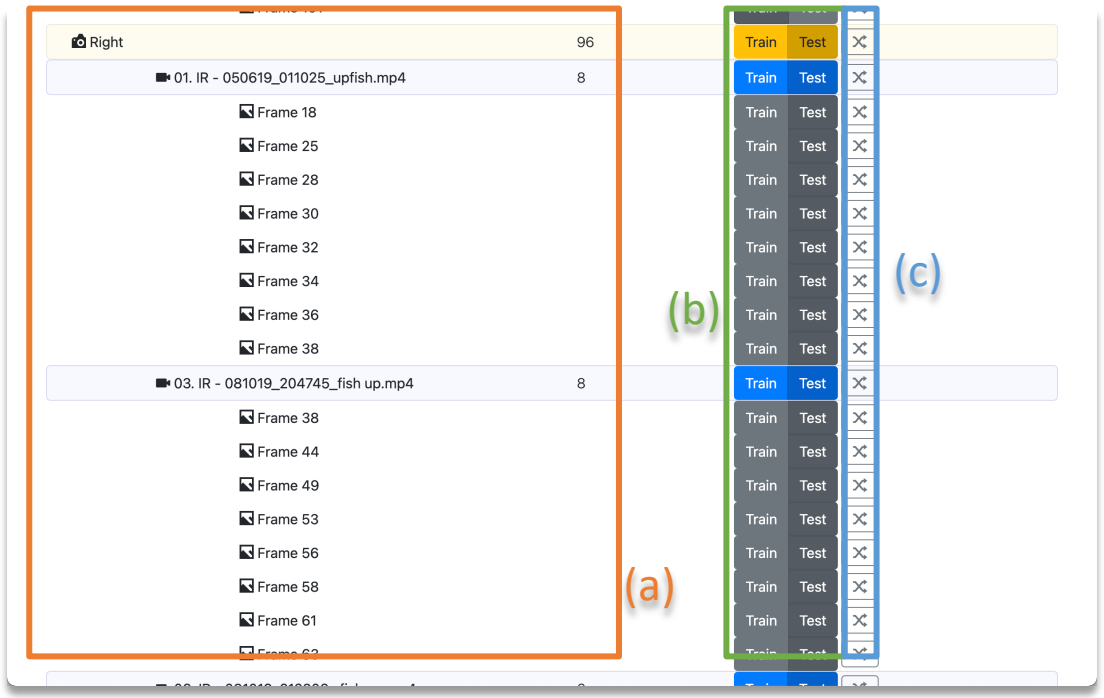
- Clicking the buttons in (c) will change the button to a lock icon. Locked images/videos/cameras will be unaffected if you click the *split randomly* buttons further up, as discussed previously
  - This allows you to specifically assign a split to certain images/videos/cameras, while allowing the rest to be determined randmoly



4. Build a training dataset

Dataset builder (manual)

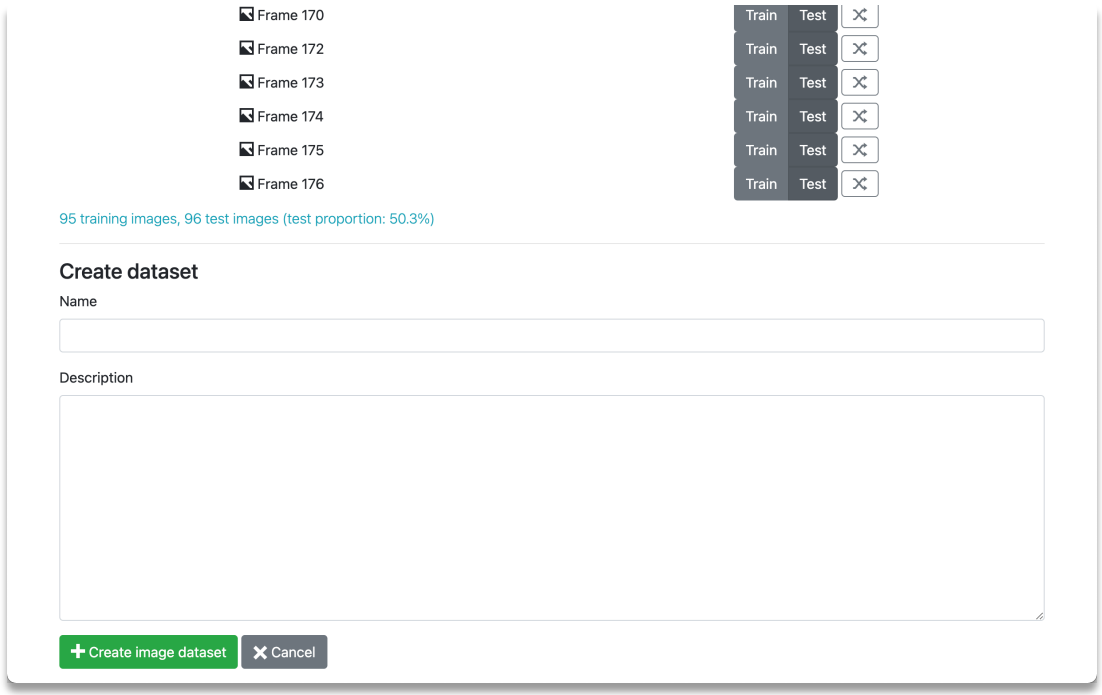
- Clicking the buttons in (c) will change the button to a lock icon. Locked images/videos/cameras will be unaffected if you click the *split randomly* buttons further up, as discussed previously
  - This allows you to specifically assign a split to certain images/videos/cameras, while allowing the rest to be determined randmoly



4. Build a training dataset

Dataset builder

- Finally, give the dataset a name and a description
- Click the green *Create image dataset* button to create it
  - As stated earlier, once created, only its name and description can be edited; the splits assigned to the images *cannot* be modified later

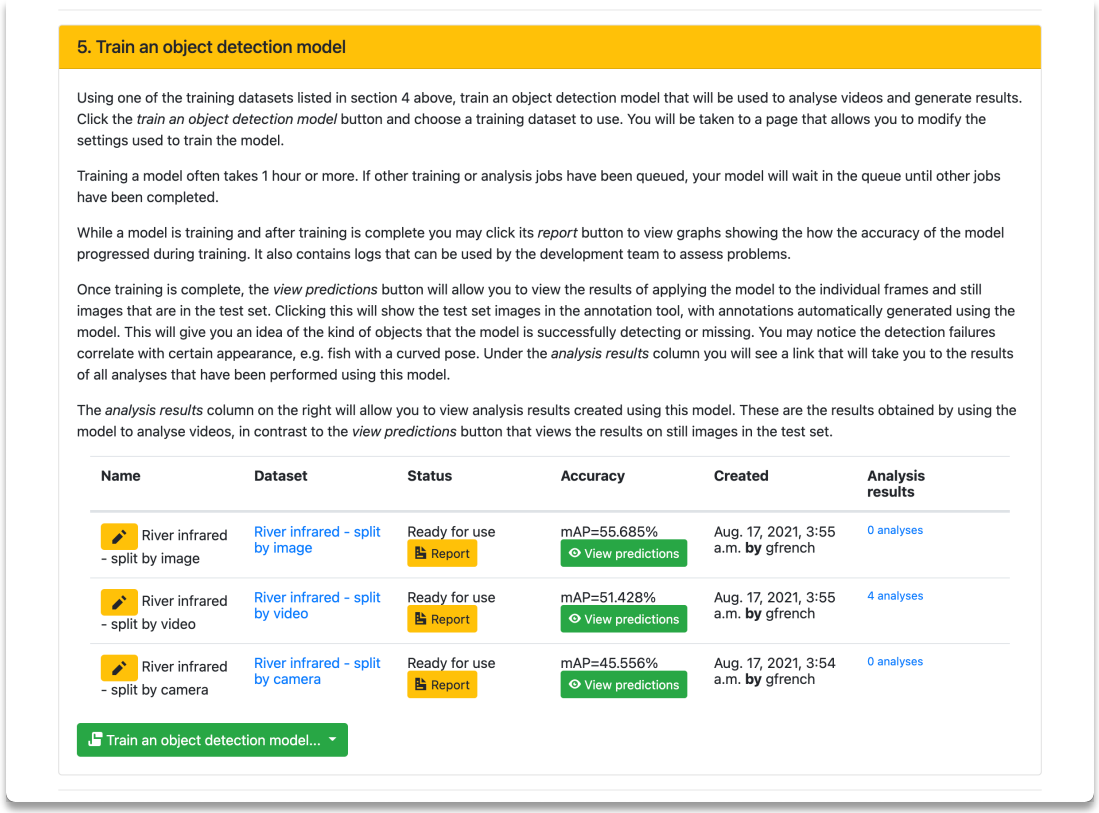


Using AVIMS

5. Train an object detection model

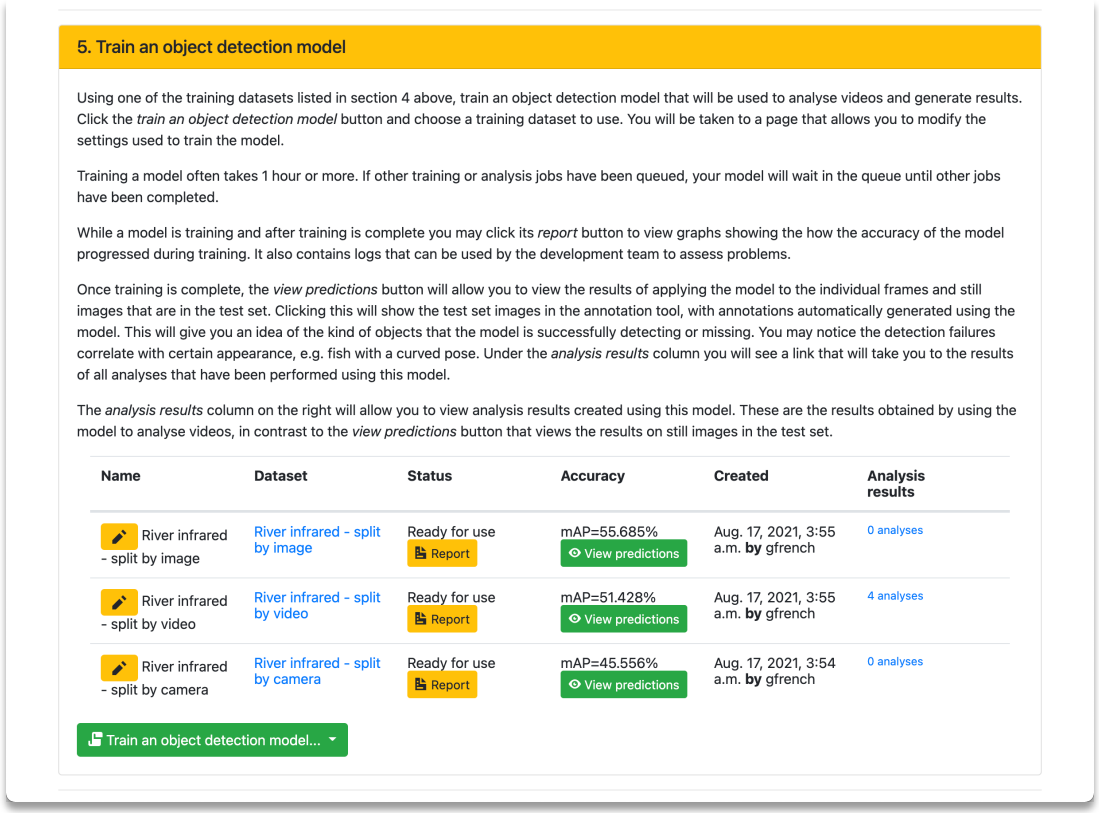
5. Train an object detection model

- Section 5 of the survey page lists the models that have been trained.
- Section 5 will be grey and not usable if:
  - You have not created any datasets



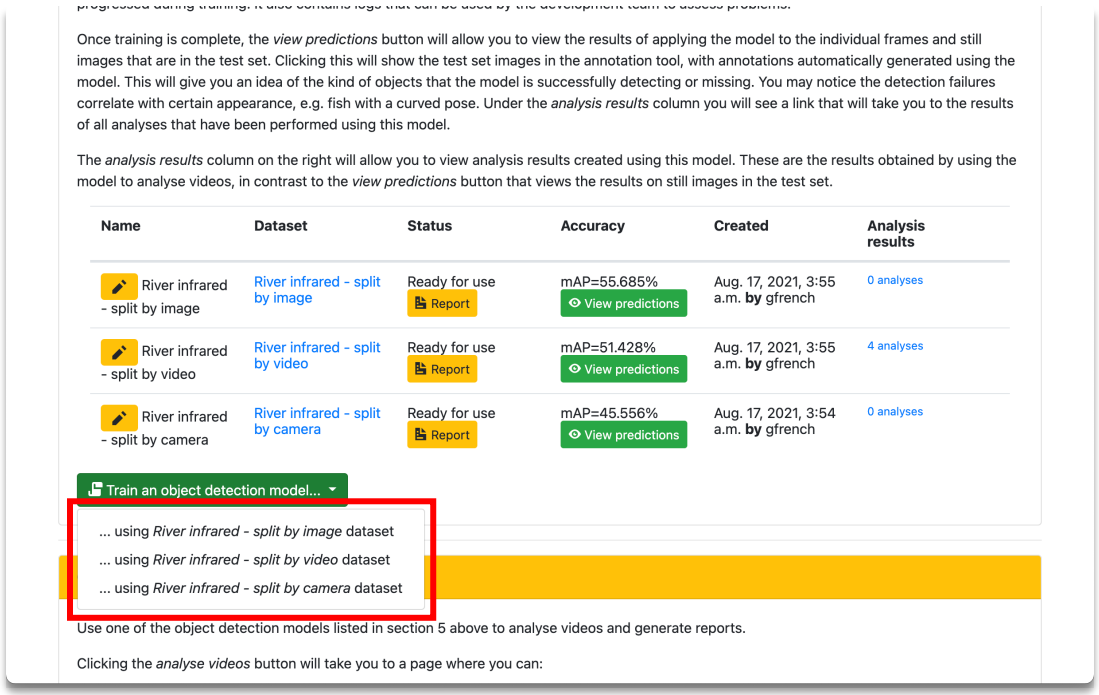
5. Train an object detection model

- Each model in the list shows:
  - The name of the model, with an edit button to change its' name
  - The dataset used to train it
  - A training report and log
  - Accuracy and predictions on the test set
  - Links to any analysis results generated using the model (see section 6, later)



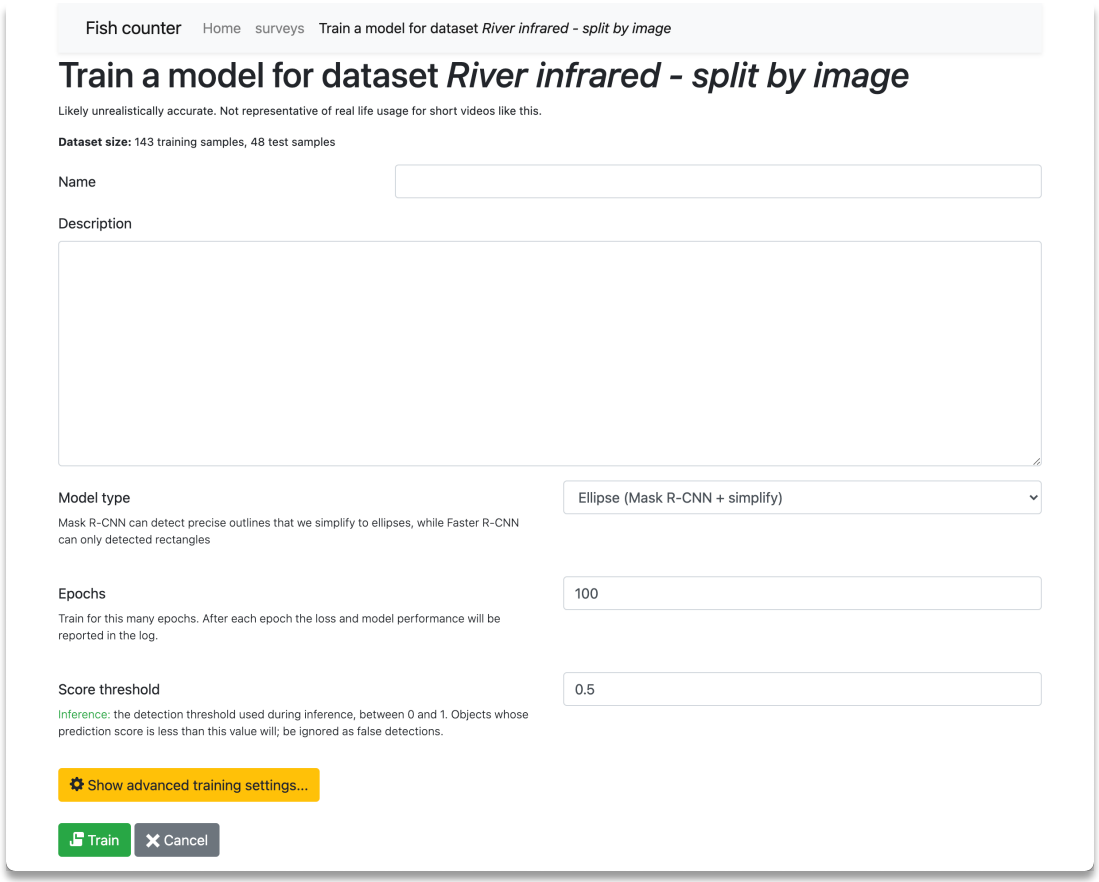
5. Train an object detection model

- To train an object detection model, click the green *Train an object detection model…* button
  - A menu will appear allowing you to select the dataset you wish to use
  - Choose a dataset



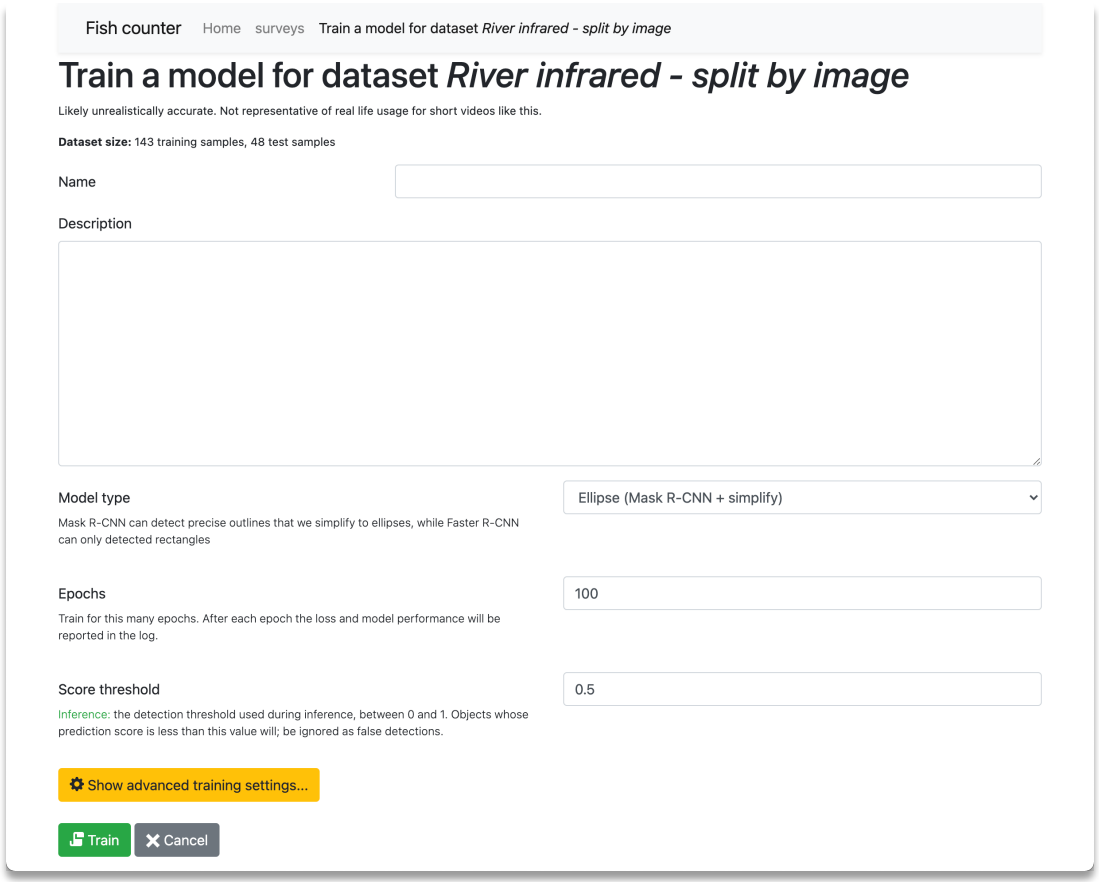
5. Train an object detection model

- The train model form:
  - Fill in the name and description
  - The model type allows you to choose between detecting objects as oriented ellipses, or non-oriented boxes
  - The default settings for *Epochs* and *Score threshold* should work in most cases. You can try other settings if you wish



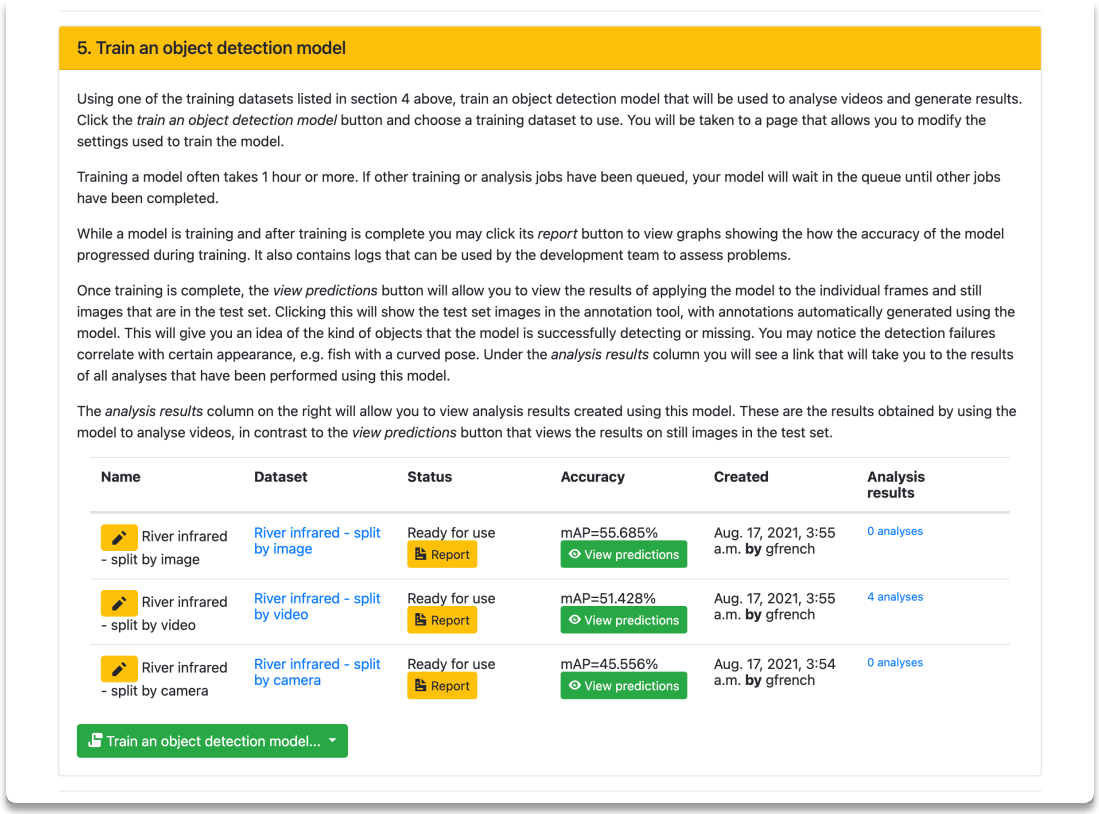
5. Train an object detection model

- The train model form:
  - The yellow *Show advanced training settings* button will show more fields. These can be left as is.
  - Speak to the developers about using different settings here
- Click the green *Train* button to train the model
  - Training takes time and will happen in the background



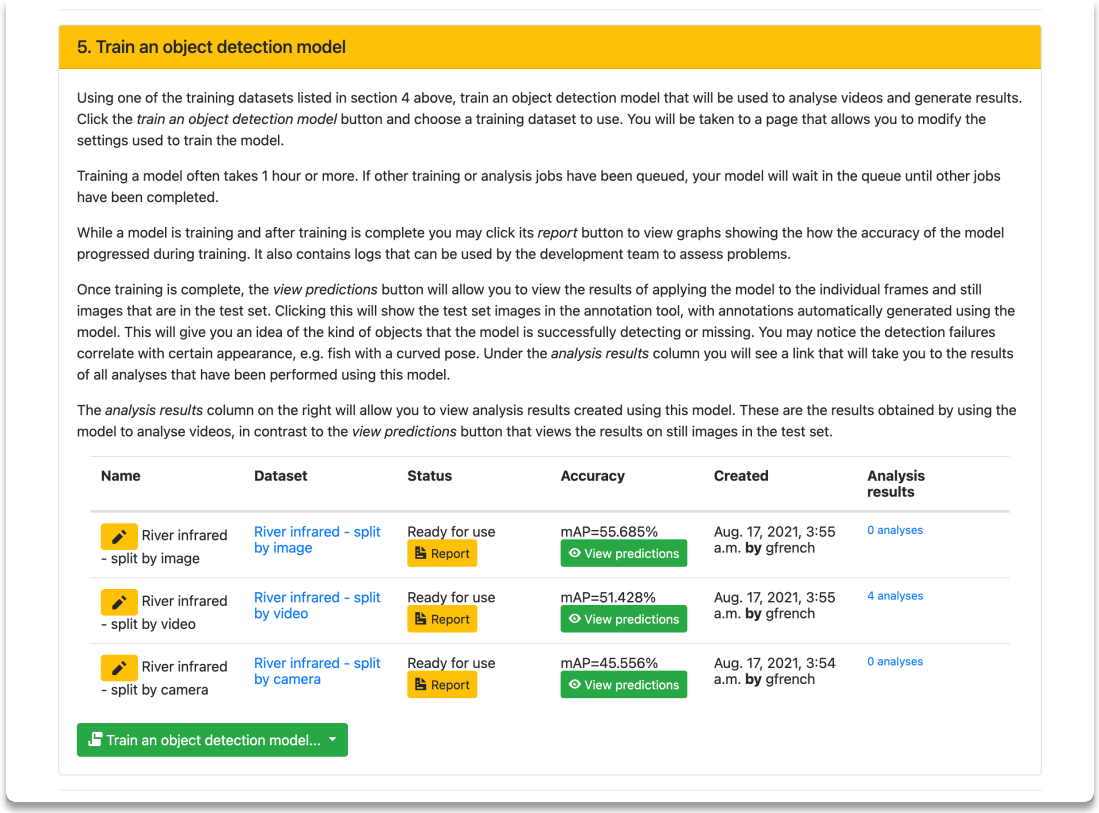
5. Train an object detection model

- The status will start out as *Queued*, indicating that the training job has not started. Once it starts, the yellow *Report* button will appear. You can click this during training to bring up a progress graph



5. Train an object detection model

- Once training has finished, the status will change to *Ready for Use*, as seen in the image to the right
  - Under the accuracy column you will see a measure of the accuracy as *mean average precision* (mAP)
  - The *View predictions* button will be available. Clicking it will show the images in the dataset test split in the labelling tool, with the labels that were automatically predicted by the model

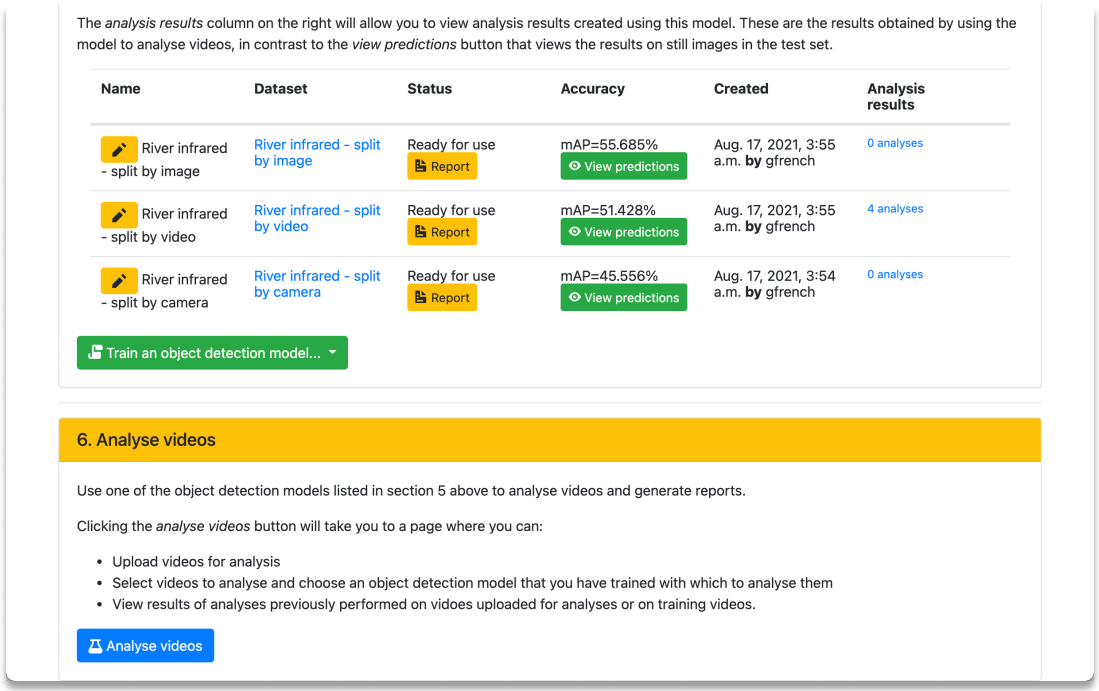


Using AVIMS

6. Analyse videos

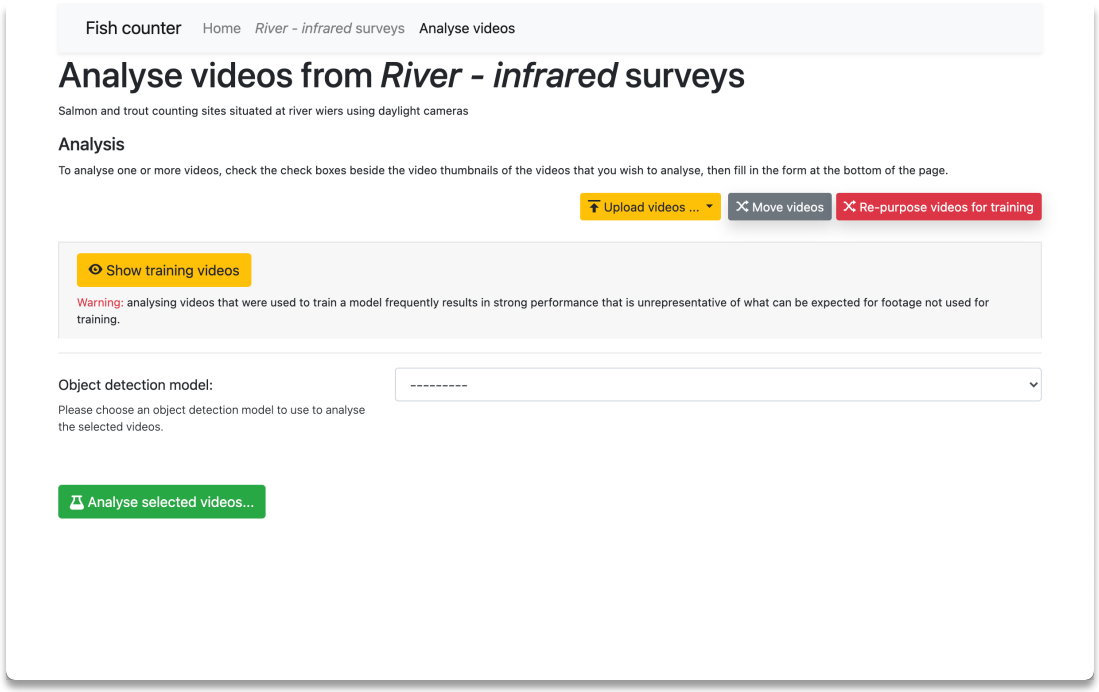
6. Analyse videos

The blue *Analyse videos* button In Section 6 will take you to the analysis page



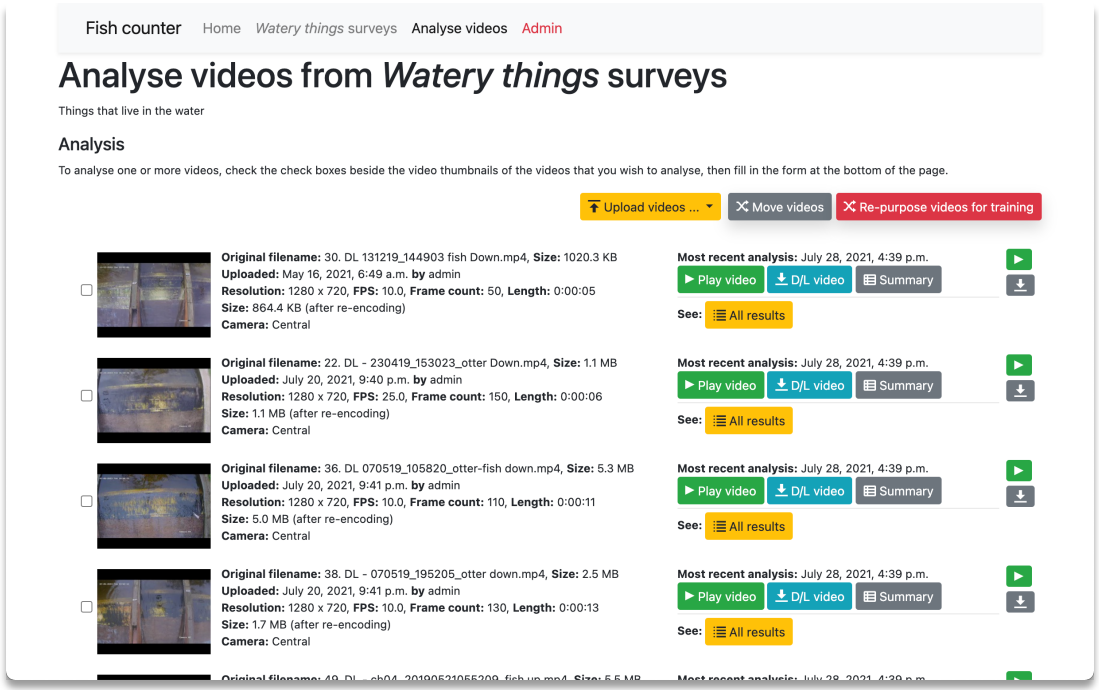
6. Analyse videos

If you haven't uploaded any videos for analysis yet, it will look like this:



6. Analyse videos

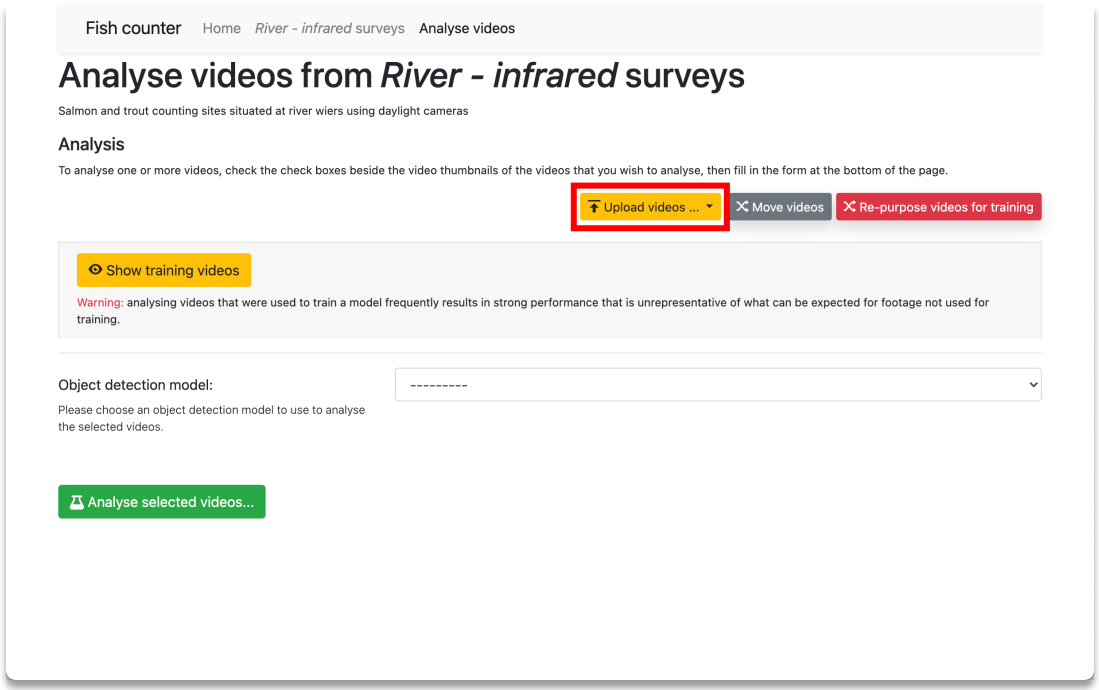
If you have, it will look like this:



6. Analyse videos

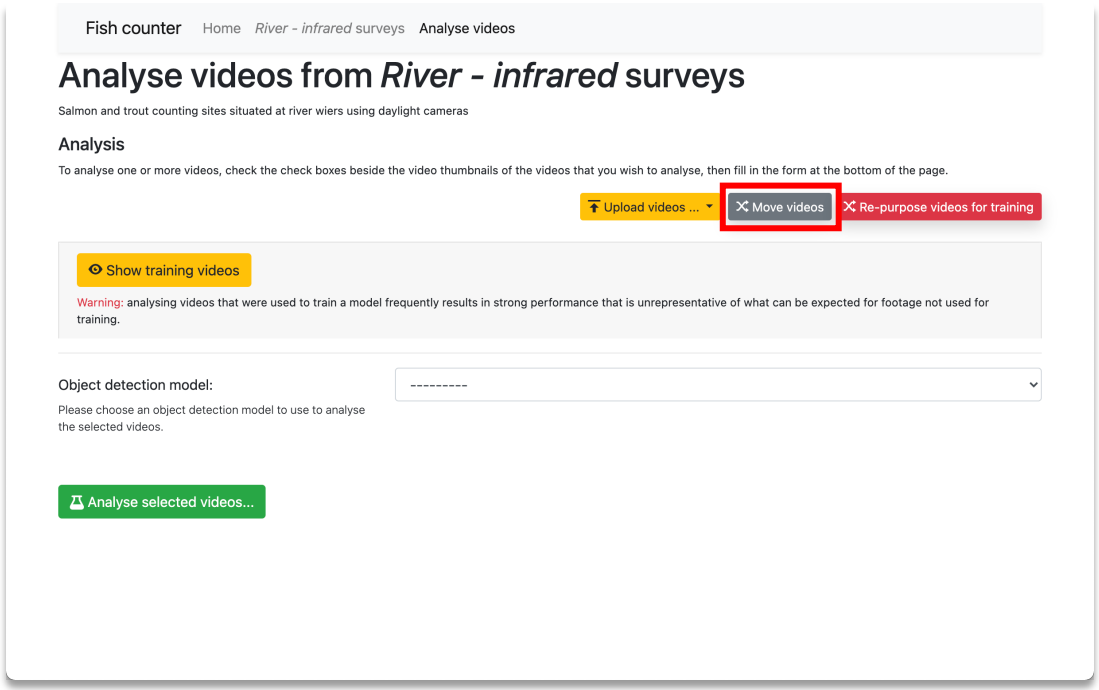
To analyse videos, you will need to upload them

- Click the yellow *Upload videos* button
- A menu will appear asking you to choose a camera to associate them with
  - If you need to create a new camera, you will need to return to the survey page, then to the upload page as in section 1
- You will be taken to an upload page where you can upload video files



6. Analyse videos

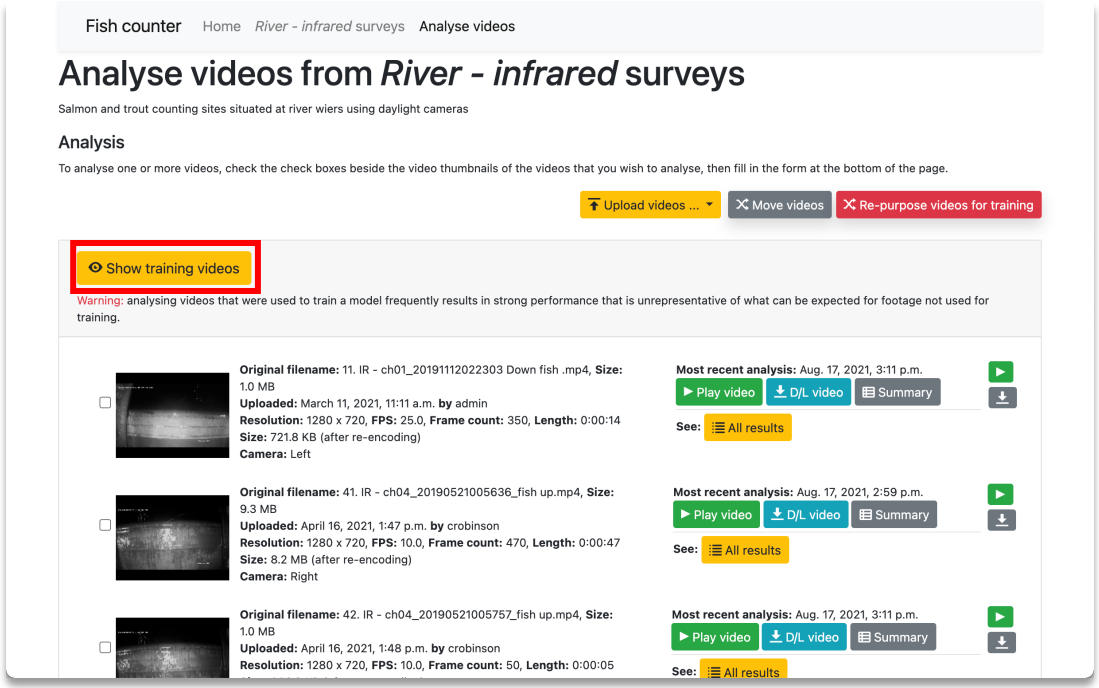
The *Move videos* button allows you to change the camera that videos are associated with, as with the upload page in section 1



6. Analyse videos

Clicking the yellow *Show training videos* will show the videos that were used for training.

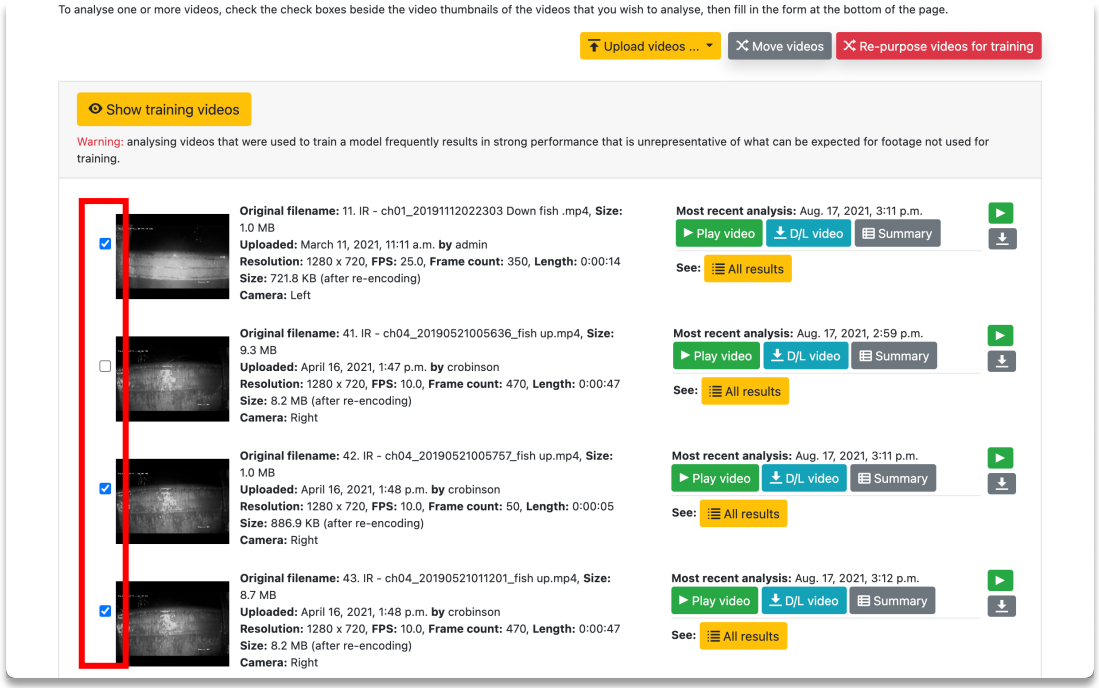
- You can analyse these too, with the caveat that analysing a video that was used for training will get strong performance unrepresentative of what can be expected on other footage



6. Analyse videos

You need to select the videos you want to analyse.

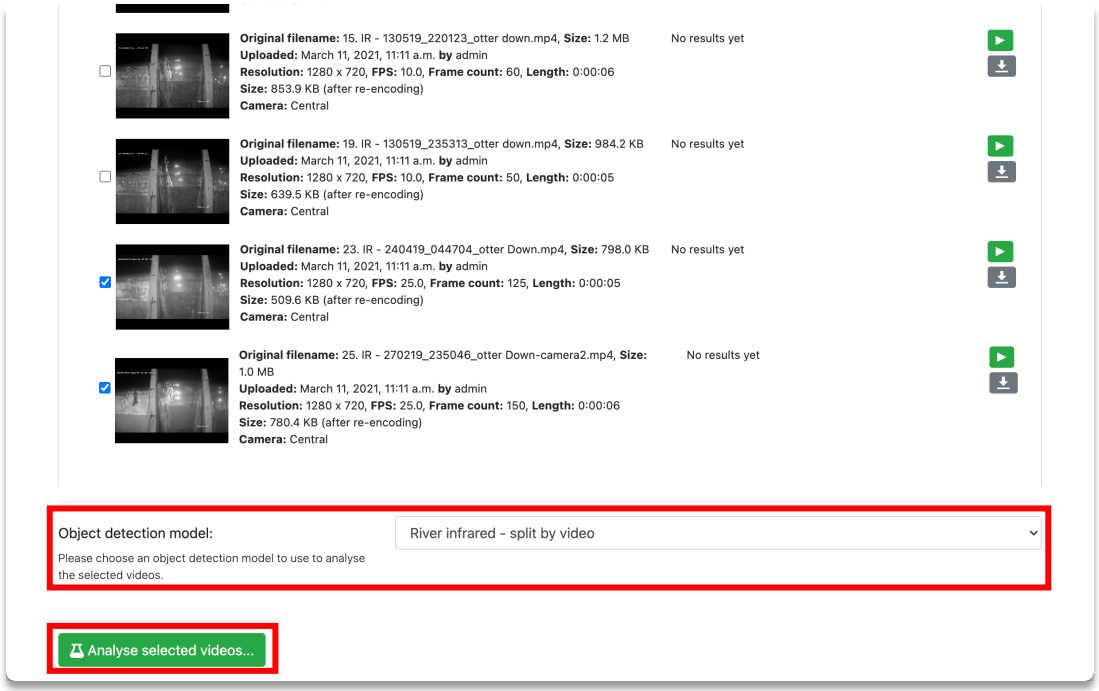
- For each video you want to analyse, tick the checkbox to the left of its' thumbnail image.



6. Analyse videos

Choose the object detection model using the drop-down near the bottom of the page. One is selected here.

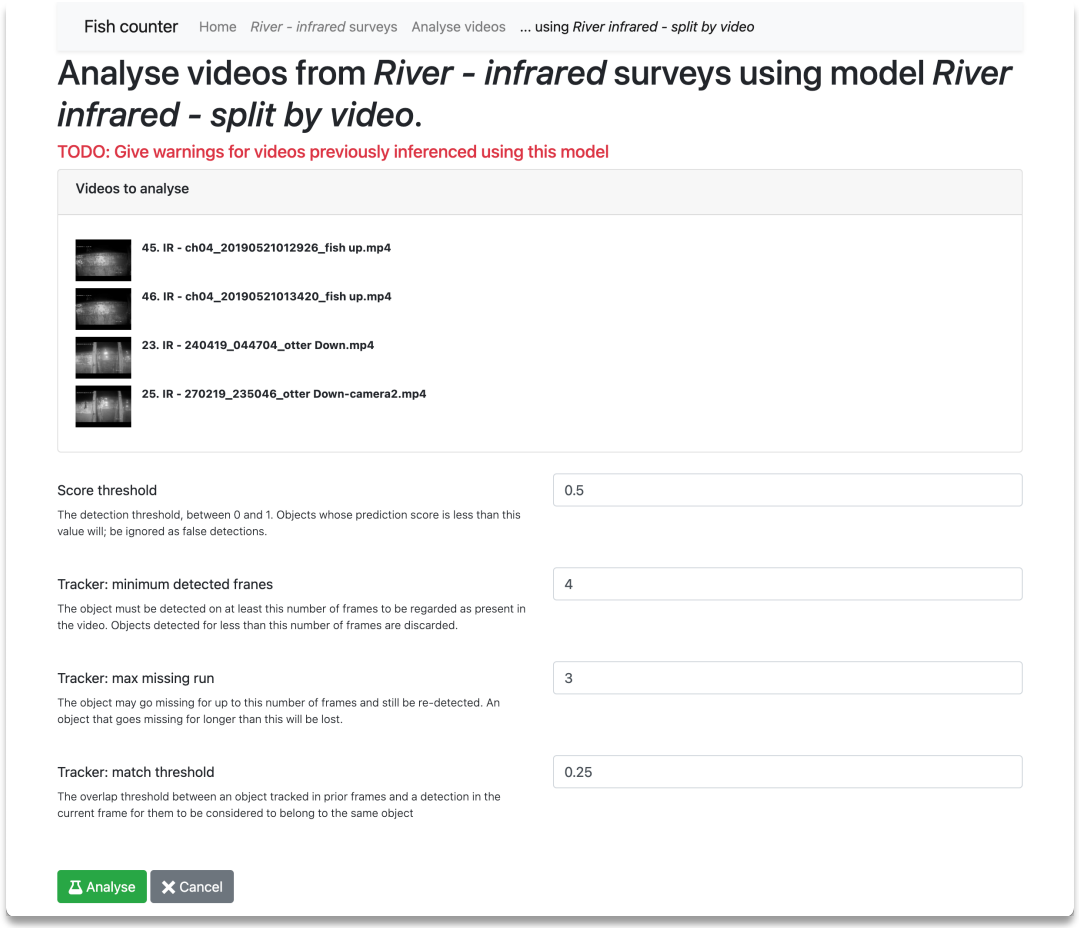
Then click the green *Analyse selected videos…* button



6. Analyse videos

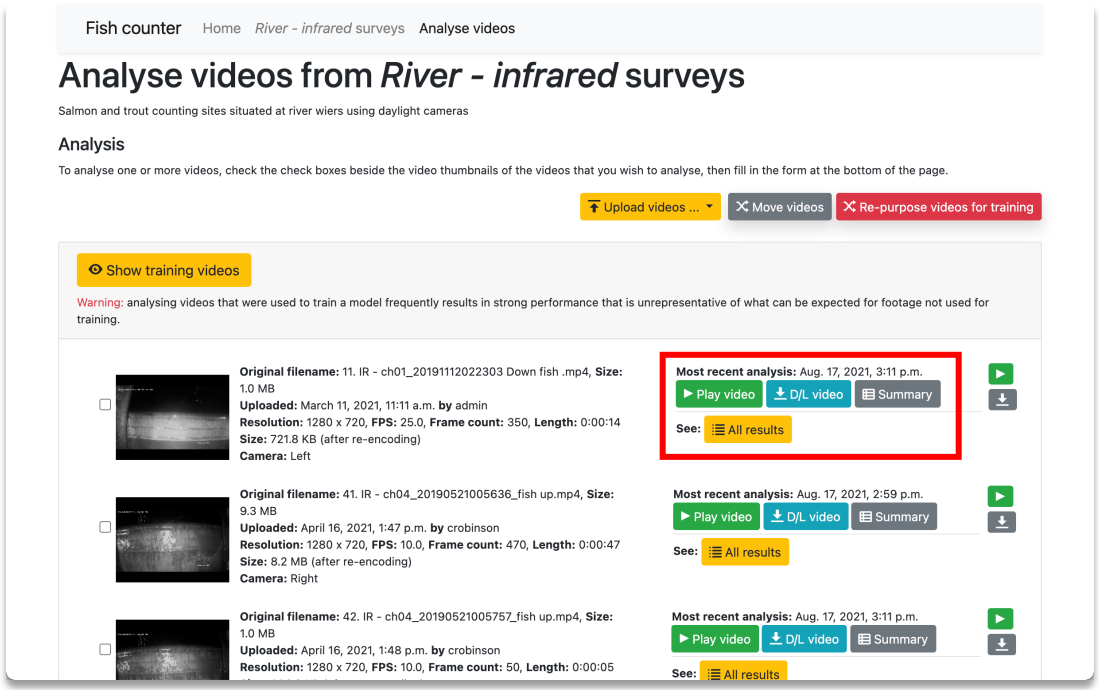
The settings in the *analyse videos* form can be left as is.

Click the green *Analyse* button to queue the analysis job.



6. Analyse videos

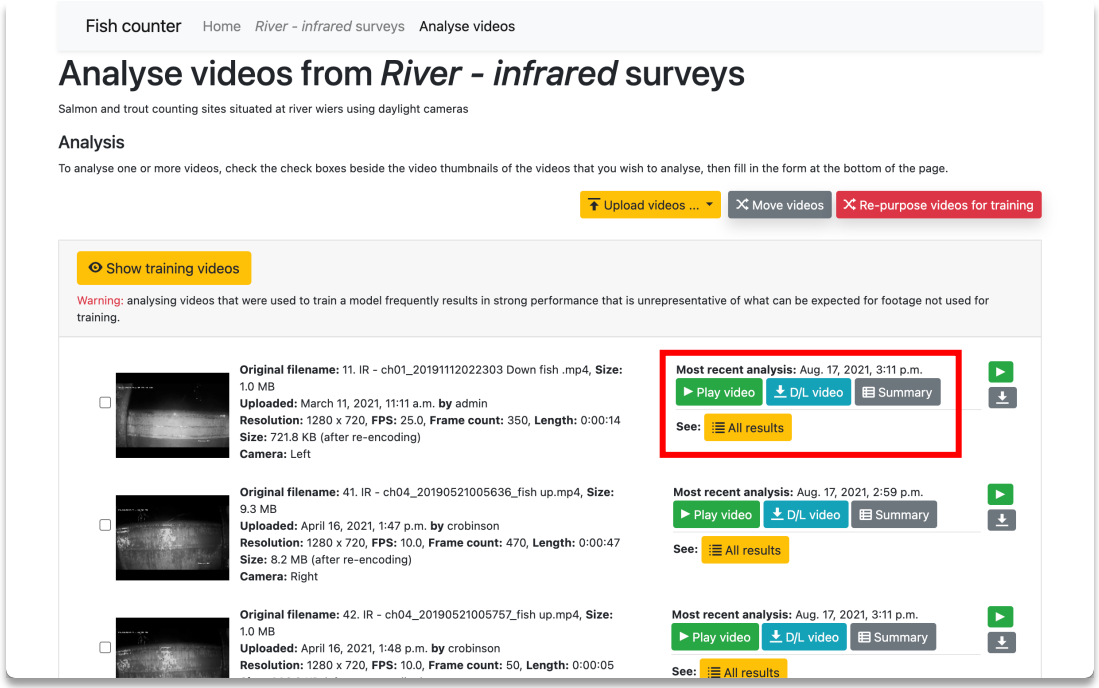
Once at least one analysis of a video is complete (a video can be analysed multiple times, e.g. by multiple models), the results can be quickly accessed from the analysis page.



6. Analyse videos

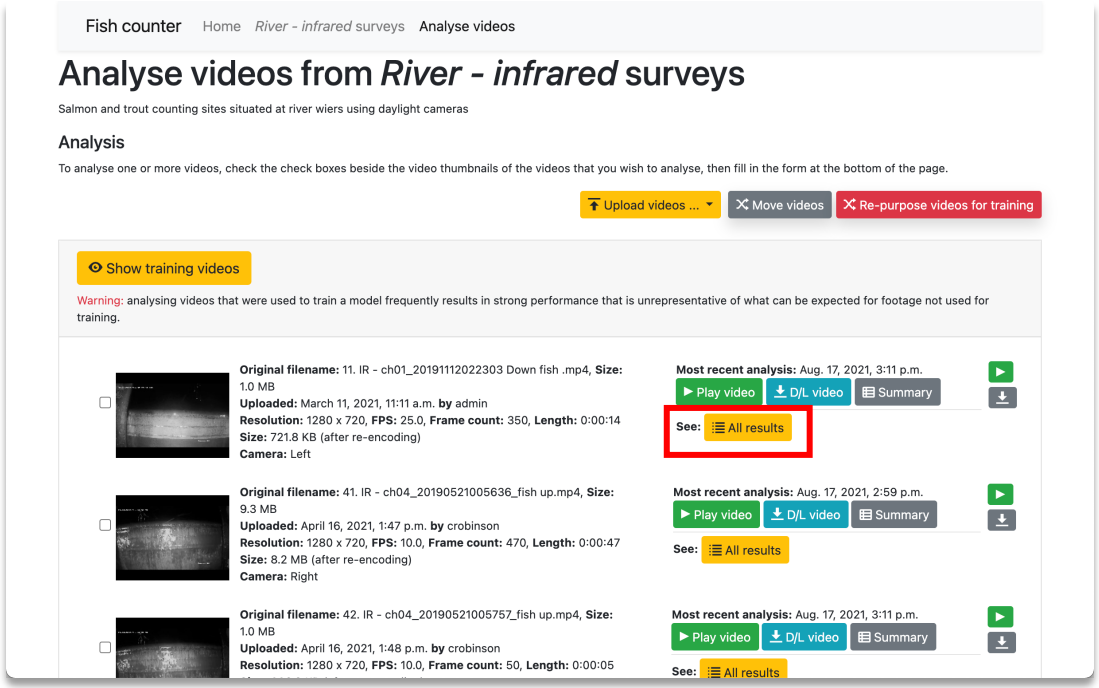
A video with detected and tracked objects can be played or downloaded.

A tabular summary of the detections can be downloaded as a .CSV file or viewed as a table.



6. Analyse videos

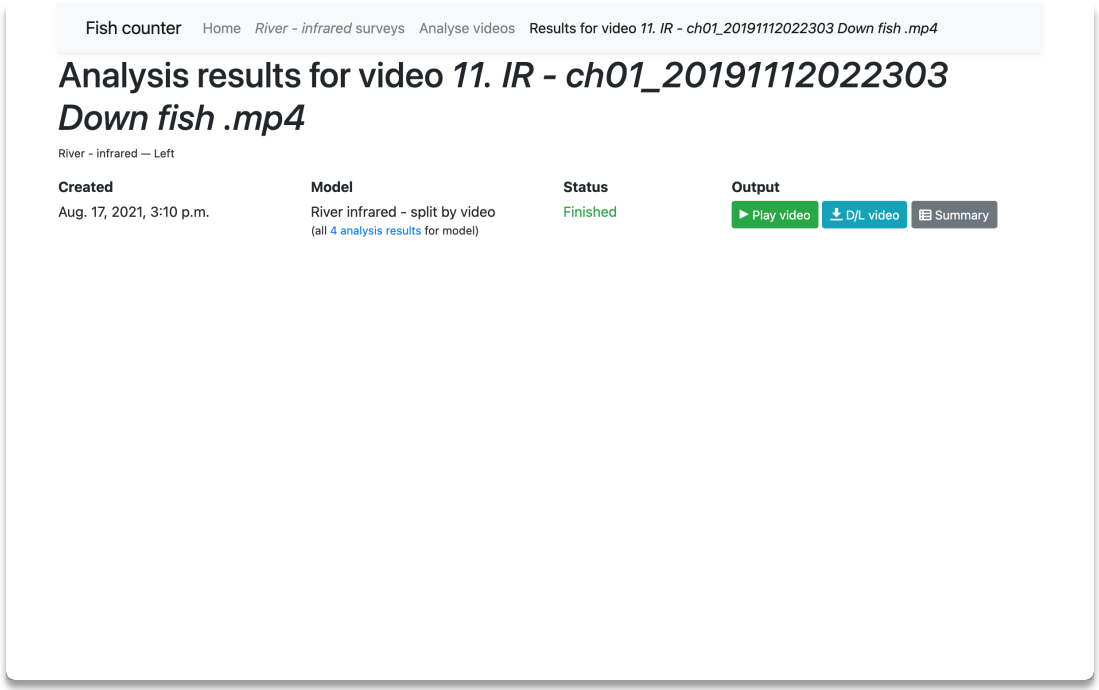
Clicking the yellow *All results* button will show you all the analysis results for that video…



6. Analyse videos

In this case, this video has only been analysed once, so only one entry is visible.

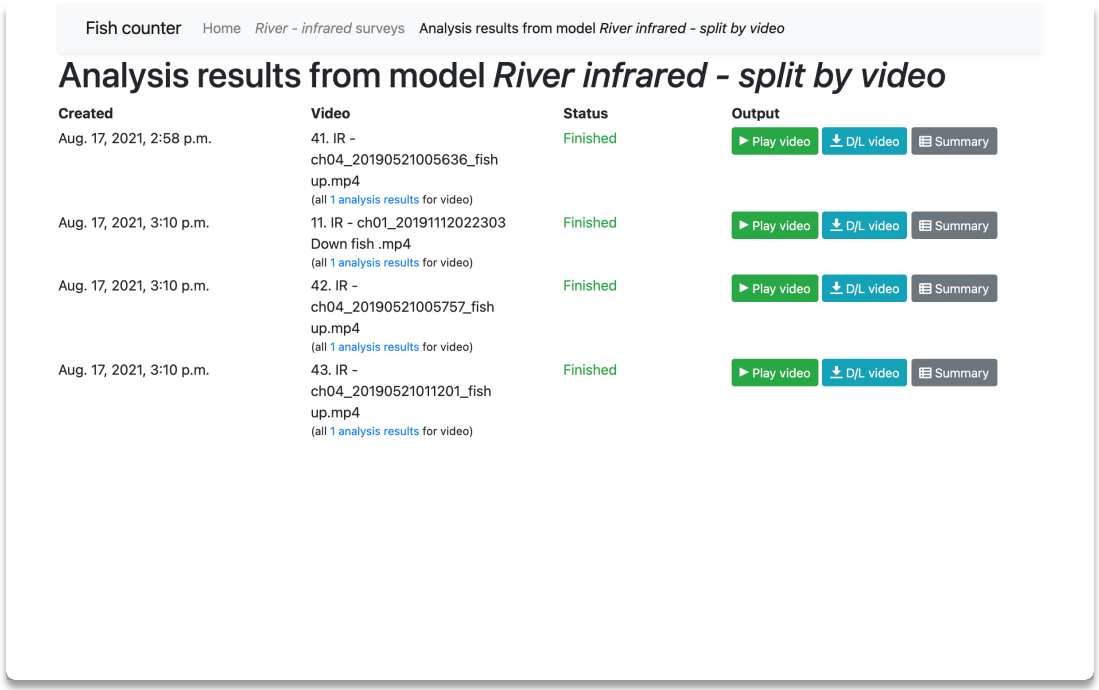
Clicking the small blue *‘4 analysis results’* in the *Model* column will show all the analysis results created using that model...



6. Analyse videos

Analysis results from a model

This can also be accessed from the survey page; the last column in the table in section 5 links here.



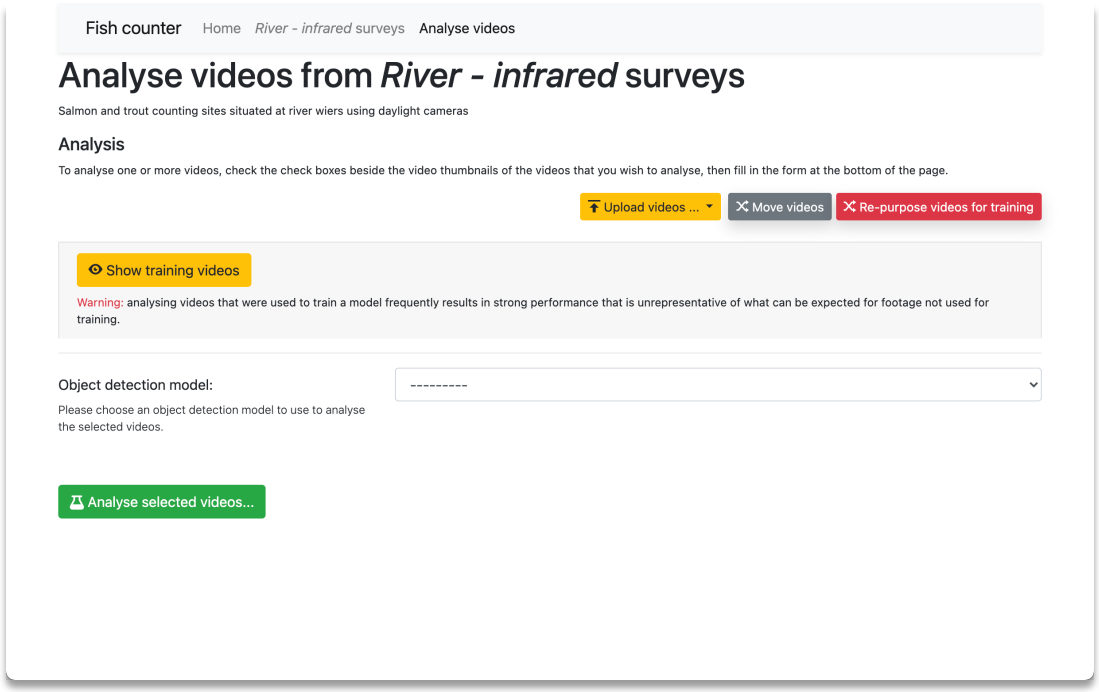
Using AVIMS

Circular work-flow

Circular work-flow

Let’s assume that after watching the video generated by analysis, you conclude that the performance needs improvement.

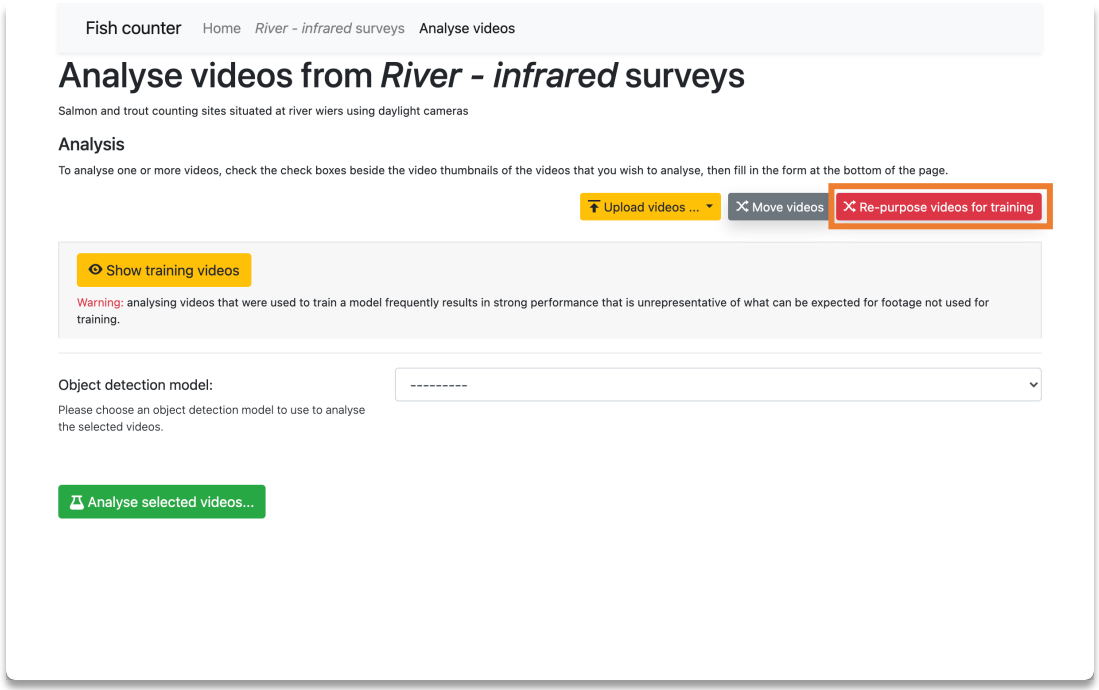
One option is to add this video to the training set.



Circular work-flow

Adding an analysed video to the training set

- Click the red *Re-purpose videos for training* button
- This will take you to a page where you can select videos to move to the training section



Circular work-flow

Adding an analysed video to the training set

- Return to the upload footage page (section 1) and choose frames for labelling
- Label these frames (section 3)
- Repeat the subsequent steps…

Thank you!

I hope you found it helpful

Contact

Email: craig.robinson@gov.scot

© Crown Copyright 2023

Marine Directorate of the Scottish Government
Marine Laboratory
375 Victoria Road
Aberdeen
AB11 9DB

Copies of this report are available from the Scottish Government Publications website at https://www.gov.scot/publications/