

Existing AI tools and services: A Galaxy Platform focus

27 Janvier 2026

Using Galaxy-Ecology as a sustainable Biodiversity common virtual research environment: Focus on integration, access and automation of AI algorithms

Yvan Le Bras

MNHN / PNDB-Data Terra / UAR DoHNÉE
French BON
France



Coline Royaux
MNHN / PNDB – DATA TERRA / UAR DoHNÉE / French BON
France

Marie Jossé
CNRS / – DATA TERRA / French BON
France

Pauline Segueineau
MNHN / PNDB – DATA TERRA / UAR DoHNÉE / French BON
France

Arthur Barreau
SU / PNDB – DATA TERRA / UMR ISYEB / French BON
France

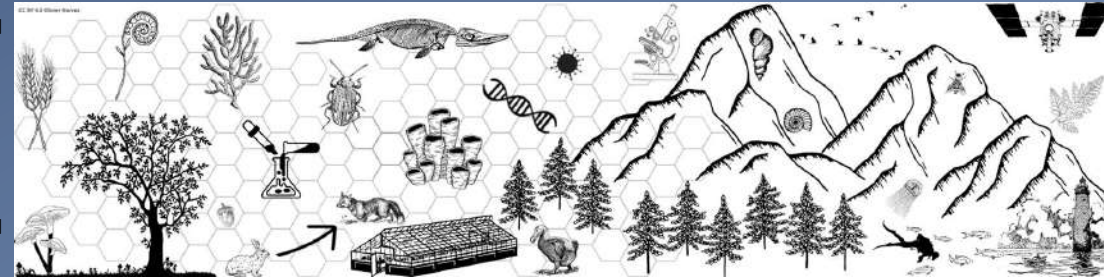
Nadine Le Bris
SU / PNDB – DATA TERRA / UMR ISYEB / French BON
France

Olivier NORVEZ
CNRS / PNDB & Theia – DATA TERRA / French BON
France

Francis Clément
MNHN / PNDB – DATA TERRA / French BON
France

Nicolas Buisine
MNHN / PNDB – DATA TERRA / UAR DoHNÉE / French BON
France

Jian-Sheng Sun
MNHN / PNDB – DATA TERRA / UAR DoHNÉE / French BON
France



Galaxy Ecology: From a platform to best practices in Biodiversity data analysis

Applying “Galaxy-E” best practices to reuse!

Atomization & Generalization to reuse & transpose



GigaScience, 2025, 14, 1–12
DOI: 10.1093/gigascience/giae122
Review

Guidance framework to apply best practices in ecological data analysis: lessons learned from building Galaxy-Ecology

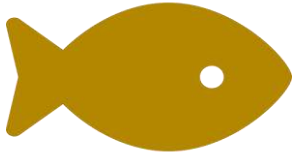
Coline Royaux^{1,2,*}, Jean-Baptiste Mihoub³, Marie Jossé⁴, Dominique Pelletier⁵, Olivier Norvez⁶, Yves Anne Fouilloux⁸, Helena Rasche¹⁰, Saskia Hiltmann¹¹, Bérénice Batut^{12,13}, Eléaume Marc^{14,15}, Pauli Guillaume Massé¹⁶, Alan Amossé¹⁷, Claire Bissery¹⁸, Romain Lorrilliere¹⁹, Alexis Martin²⁰, Yves Bas^{2,20}, Thimothée Virgoulay^{21,22}, Valentin Chambon¹⁷, Elie Arnaud²², Elisa Michon²³, Clara Urfer^{2,24}, Eloïse Trigodet²⁴, Gregoire Lois², Romain Julliard²⁵, Björn Grünig²⁵, Yvan Le Bras², and The Galaxy-E community

Guidance framework to apply best practices in ecological data analysis: Lessons learned from building Galaxy-Ecology



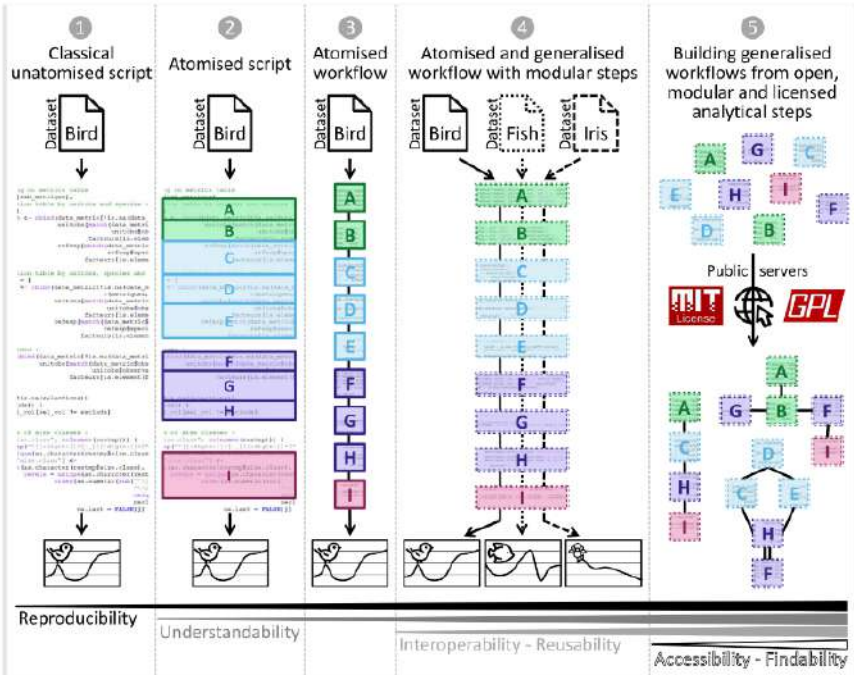
PCI recommendation: <https://doi.org/10.24072/pci.ecology.100694>

Royaux Coline^{1,2,*}, Mihoub Jean-Baptiste³, Jossé Marie⁴, Pelletier Dominique⁵, Norvez Olivier⁶, Reecht Yves^{7,8}, Fouilloux Anne⁹, Rasche Helena¹⁰, Hiltmann Saskia¹¹, Batut Bérénice^{12,13}, Eléaume Marc^{14,15}, Segueineau Pauline^{14,15}, Massé Guillaume¹⁶, Amossé Alan¹⁷, Bissery Claire^{18,19}, Lorrilliere Romain³, Martin Alexis¹⁹, Bas Yves^{3,20}, Virgoulay Thimothée^{21,22}, Chambon Valentin¹⁷, Arnaud Elie², Michon Elisa²³, Urfer Clara^{2,24}, Trigodet Eloïse^{21,24}, Delannoy Marie³, Lois Gregoire³, Julliard Romain³, Grünig Björn²⁵, The Galaxy-E community, Le Bras Yvan²



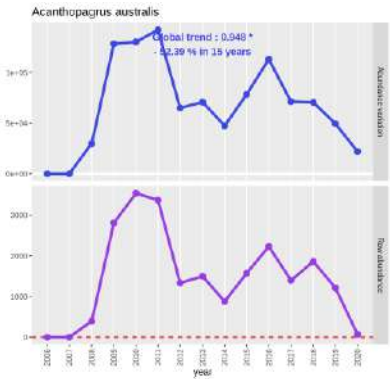
PAMPA

Graphical abstract – Levels of attainable best practices through the atomisation – generalisation framework

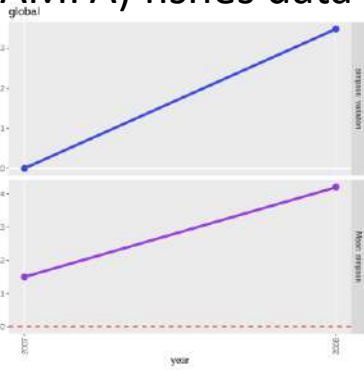


Keywords: Biodiversity; Reproducible analyses; Galaxy; Best practices; Atomisation; Generalisation; Workflows; Ecoinformatics; Conda; Container; Common Workflow Language; RO-CRATE

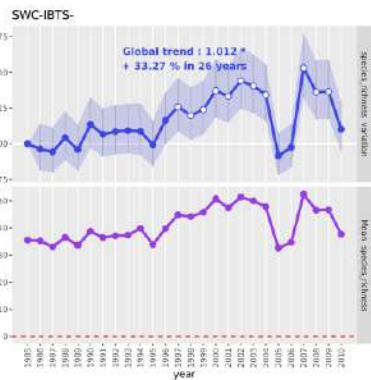
REEF LIFE SURVEY fishes data



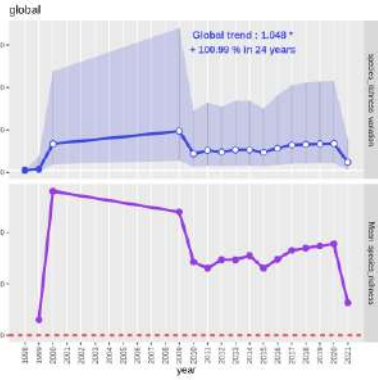
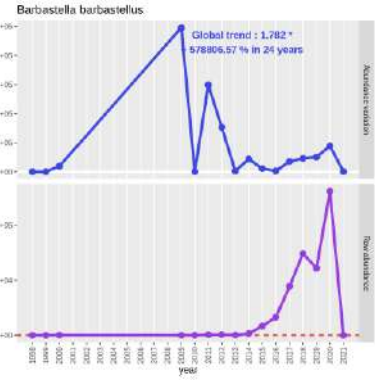
Marine Protected Areas (PAMPA) fishes data



DATRAS fishing data



VIGIE CHIRO bats data

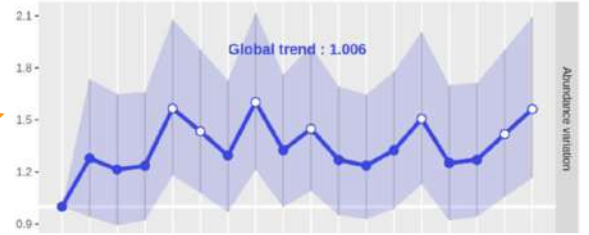
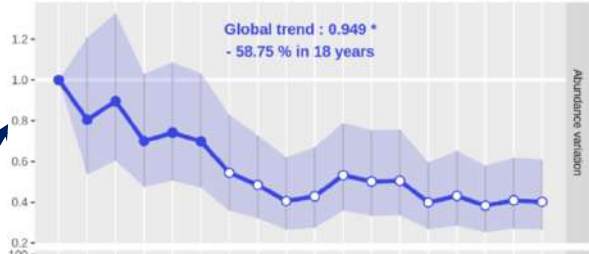
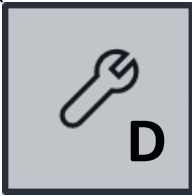
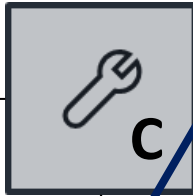
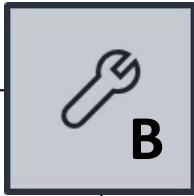
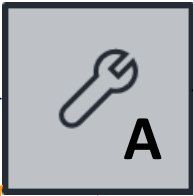


Applying “Galaxy-E” best practices to reuse!

Different data, same primary variables, same workflow

Location Year Species Occurrence

	carre	annee	espece	abond
1	2	2016	ACCGEN	0
2	2	2017	ACCGEN	0
3	2	2018	ACCGEN	0
4	2	2019	ACCGEN	0
5	183	2016	ACCGEN	0
6	183	2017	ACCGEN	0



	Survey	Year	Quarter	Area	AphiaID	Species	LngtClass	CPUE_number_per_hour
1	BITS	1991	1	22	126281	Anguilla anguilla	0	0.000000
2	BITS	1991	1	22	126281	Anguilla anguilla	720	0.009160
3	BITS	1991	1	22	126417	Clupea harengus	0	0.000000
4	BITS	1991	1	22	126417	Clupea harengus	80	0.075785
5	BITS	1991	1	22	126417	Clupea harengus	85	0.000000
6	BITS	1991	1	22	126417	Clupea harengus	95	0.000000
7	BITS	1991	1	22	126417	Clupea harengus	100	0.000000
8	BITS	1991	1	22	126417	Clupea harengus	105	0.012492
9	BITS	1991	1	22	126417	Clupea harenqus	110	0.618357



	Unit	bs	rotation	codeSp	sexe	taille	classe_taille	poids	nb_ind
1	AS140155	3		Hemifasc	-999	-999	P	-999	1
2	AS140159	1		Nasosp.	-999	-999	P	-999	3
3	AS140159	3		Gompvari	-999	-999	P	-999	1
4	AS140160	3		Gompvari	-999	-999	P	-999	1



Applying “Galaxy-E” best practices to reuse!

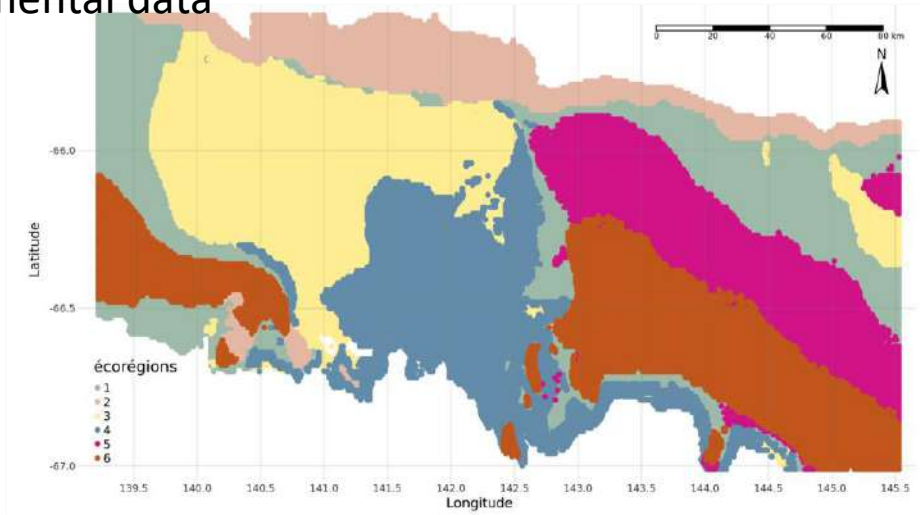
Different data, same primary variables, same workflow

Transpose to other taxonomic and/or geographical and/or temporal context

From Antarctica to North East Atlantic

From Invertebrates to Fishes

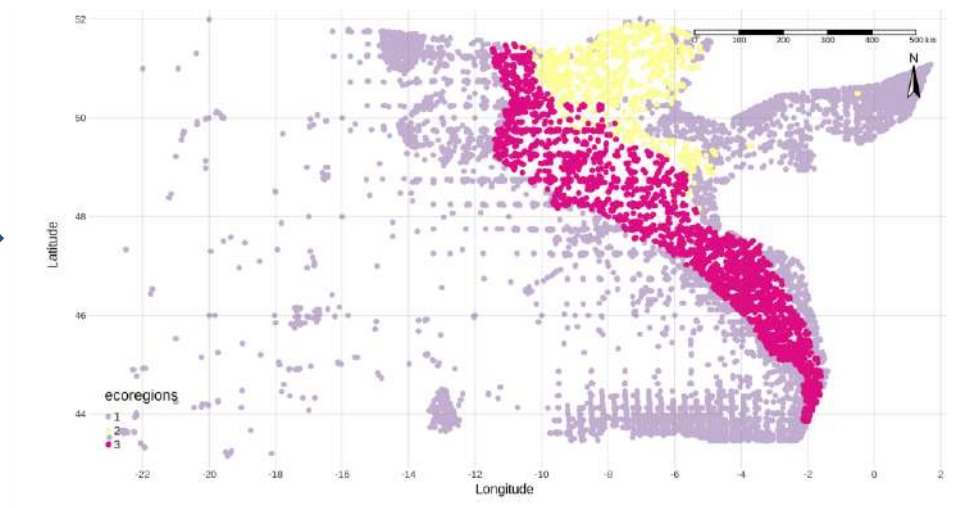
Project data (invertebrates)
Project environmental data



North Atlantic

OBIS marine occurrence data (fishes)
OBIS marine environmental data (all)

c1: decimallongitude x c2: decimallatitude x c5: shoredistance x c6: bathymetry x c7: sst x c8: sss x



Applying “Galaxy-E” best practices to reuse!

Different data, same primary variables, same workflow

Test several scenarios

From Antarctica to North East Atlantic

From Invertebrates to Fishes

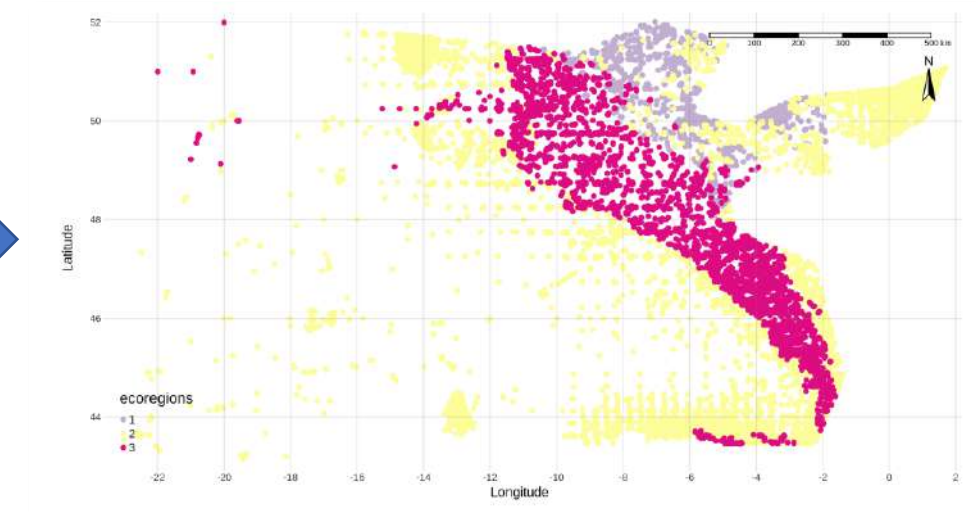
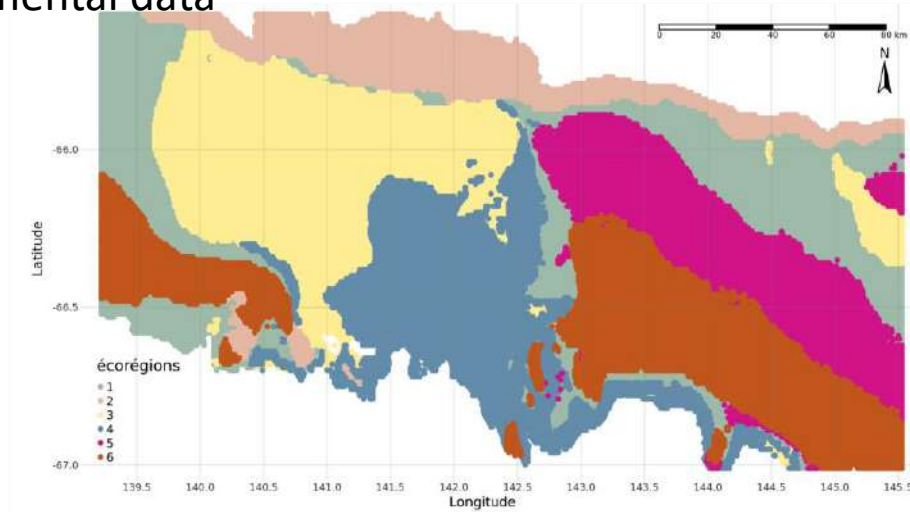
Modifying modelling parameters

North Atlantic

OBIS marine occurrence data (fishes)

OBIS marine environmental data (sst / bath

Project data (invertebrates)
Project environmental data



Training tutorials for Ecology



Data and Metadata Management

These tutorials are focusing on data and metadata management in Ecology.

Lesson	Slides	Hands-on	Recordings	Input dataset	Workflows
Cleaning GBIF data using OpenRefine biodiversity					
Creating FAIR Quality assessment reports and draft of Data Papers from EML metadata with MetaShRIMPS Metadata EML FAIR Data Paper biodiversity					
Creating metadata using Ecological Metadata Language (EML) standard with EML Assembly Line functionalities Metadata EML Ecology Biodiversity FAIR Data Paper					
Data submission using ENA upload Tool					

Data access

These lessons focus on ways to access data classically used in Ecology

Lesson	Slides	Hands-on	Recordings	Input dataset	Workflows
QGIS Web Feature Services earth-system GIS Geographical Information System WFS Spatial data Maps OGC biodiversity					

Data preprocessing

These lessons focus on manners to preprocess data used in Ecology

Lesson	Slides	Hands-on	Recordings	Input dataset	Workflows
Biodiversity data exploration Taxonomic data data quality biodiversity					
Checking expected species and contamination in bacterial isolate illumina bacteria microgalaxy					
Cleaning GBIF data for the use in Ecology gbif data management data cleaning biodiversity					

Data visualization

These tutorials showcase data visualization in Ecology

Lesson	Slides	Hands-on	Recordings	Input dataset	Workflows
Visualization of Climate Data using NetCDF xarray Map Plotting paraglider					
Visualize EBV cube data with Panoply netCDF viewer Interactive-tools EBV cube Data visualization					

Data analysis

These lessons focus on ways to analyse data in Ecology

Lesson	Slides	Hands-on	Recordings	Input dataset	Workflows
Champs blocs indicators Ecosystem EBV class EBV dataset EBV workflow Marine ecosystems biodiversity					
Compute and analyze biodiversity metrics with PAMPA toolsuite Species population EBV class Community composition EBV class EBV dataset EBV workflow modeling biodiversity					
Ecoregionalization workflow tutorial ecology taxonomic data EBV workflow modeling gbif ocean earth-system Interactive-tools biodiversity					
From NDVI data with OpenEO to time series visualisation with Holoviews earth-system land degradation NDVI copernicus holoviews					
Life Traits Ecoregionalization workflow biodiversity					
Marine Omics Identifying biosynthetic gene clusters earth-system ocean marine omics biodiversity					
Metabarcoding/eDNA through Obtools Genetic composition EBV class Community composition EBV class EBV dataset EBV workflow eDNA Metabarcoding biodiversity					
Obis marine indicators earth-system ocean marine omics biodiversity					
Phylodiversity analysis quick tutorial ecology taxonomic data					
Preparing genomic data for phylogeny reconstruction phylogeny data handling functional annotation biodiversity					
RAD-Seq Reference-based data analysis RAD-seq Genetic composition EBV class Species population EBV class EBV dataset EBV workflow biodiversity					
RAD-Seq de-novo data analysis RAD-seq Genetic composition EBV class Species populations EBV class EBV dataset EBV workflow biodiversity					
RAD-Seq to construct genetic maps RAD-seq Genetic composition EBV class EBV dataset EBV workflow biodiversity					
Regional GAM Species populations EBV class Species traits EBV class EBV dataset EBV workflow biodiversity					
Sentinel 2 biodiversity Remote sensing biodiversity					
Species distribution modeling Interactive-tools modeling gbif species populations EBV class biodiversity					
Taxonomic Analysis of eDNA biodiversity					

Training tutorials for Ecology

Data and Metadata Management

These tutorials are focusing on data and metadata management in Ecology.

github.com/mjoudy/training-material/blob/yolo_predict_tutorial_deepsea/topics/imaging/tutorials/yolo_prediction/tutorial.md

training-material / topics / imaging / tutorials / yolo_prediction / tutorial.md



Preview Code Blame 286 lines (287 loc) · 12.9 KB

Part 1: Detection of marine species

Dataset

👉 We will use selected images from the SEANOE dataset (% cite lebeaud2024deepsea %).

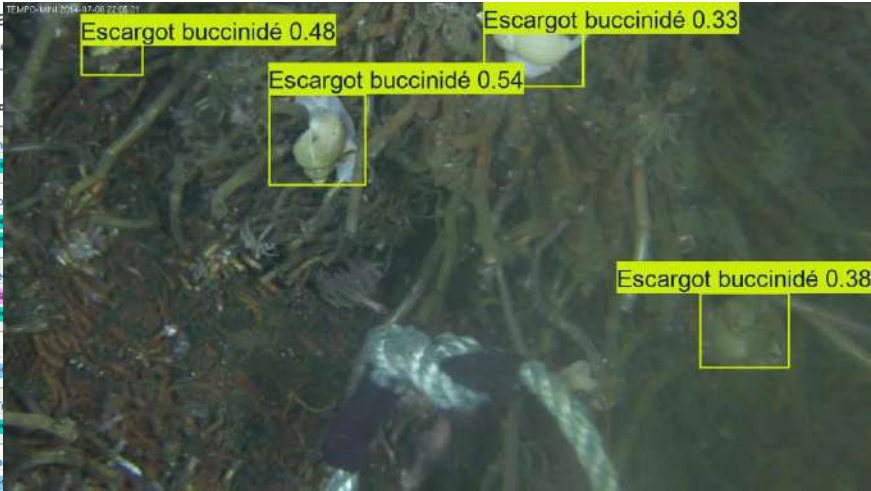
The [SEANOE](#) collection features real underwater images captured by deep-sea observatories as part of a citizen science initiative called DeepSea Spy. These non-destructive imaging stations continuously monitor marine ecosystems and provide snapshots of various fauna. In this dataset, multiple annotators—including trained scientists and enthusiastic citizen scientists—have manually labeled images with polygons, lines, or points highlighting marine organisms. These annotations were then cleaned and converted into bounding boxes to create a training-ready dataset for object detection with YOLOv8. Though the exact species vary, images often include deep-sea fish, species, making this dataset well-suited for practicing detection tasks.



Data

These k

LESSON



Escargot buccinidé 0.48

Escargot buccinidé 0.33

Escargot buccinidé 0.54

Escargot buccinidé 0.38

barcoding/eDNA through Obtools

Genetic composition EBV class Community composition EBV class EBV dataset EBV workflow eDNA

barcoding biodiversity

marine indicators

System down marine omics biodiversity

biodiversity analysis quick tutorial

phylogeny taxonomic data

Comparing genomic data for phylogeny reconstruction

Genomics data handling functional annotation biodiversity

RNA-Seq Reference-based data analysis

RNA-seq Genetic composition EBV class Species population EBV class EBV dataset EBV workflow

RNA-Seq de-novo data analysis

RNA-seq Genetic composition EBV class Species populations EBV class EBV dataset EBV workflow

RNA-Seq to construct genetic maps

RNA-seq Genetic composition EBV class EBV dataset EBV workflowz biodiversity

Phylogenetic GAM

Species populations EBV class Species traits EBV class EBV dataset EBV workflow biodiversity

Phylogenetic 2 biodiversity

Phylogenetic sensing biodiversity

Species distribution modeling

Active-tools modeling golf species populations EBV class biodiversity

Phylogenetic Analysis of eDNA

Genomics biodiversity



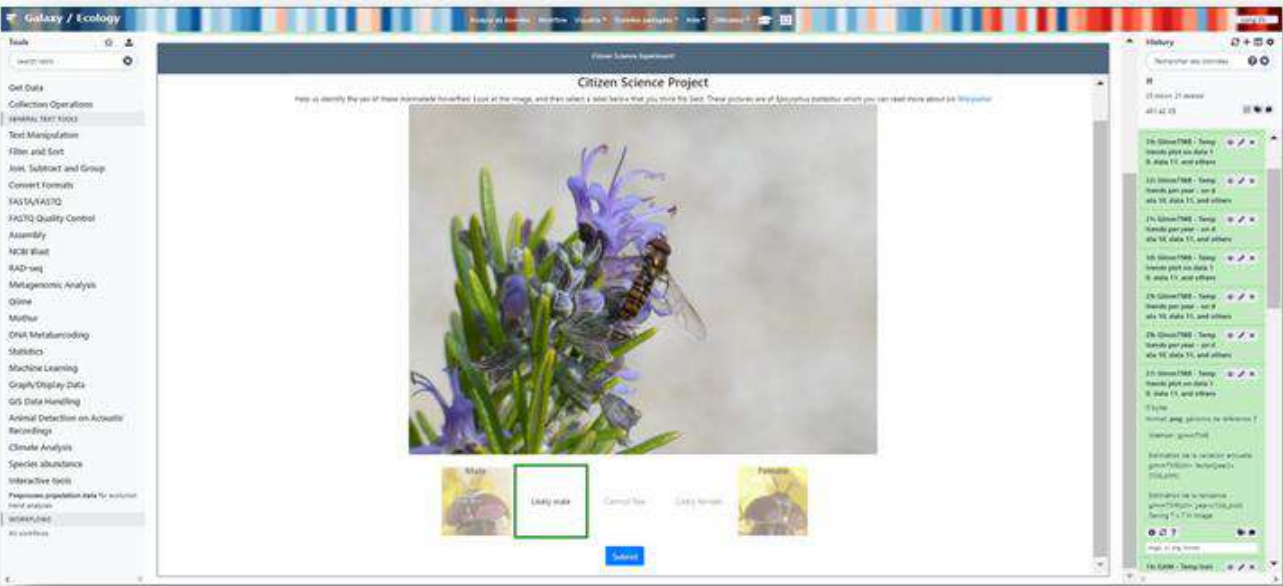
New tutorial using
AI model (*in
review*)

[Link to the PR](#)

Galaxy Ecology: From citizen science to AI

Citizen science & crowdsourcing

Crowdsourcing image



VIGIENATURE
Un réseau de citoyens qui fait avancer la science

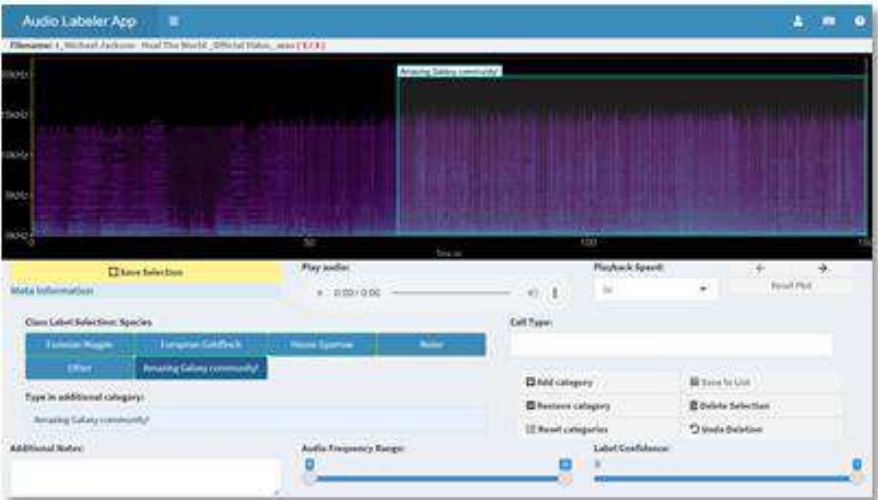


Gamification of participatory science for training and education purposes



>3400 classifications per month / >110 per day

Annotation of sounds



Annotation of images

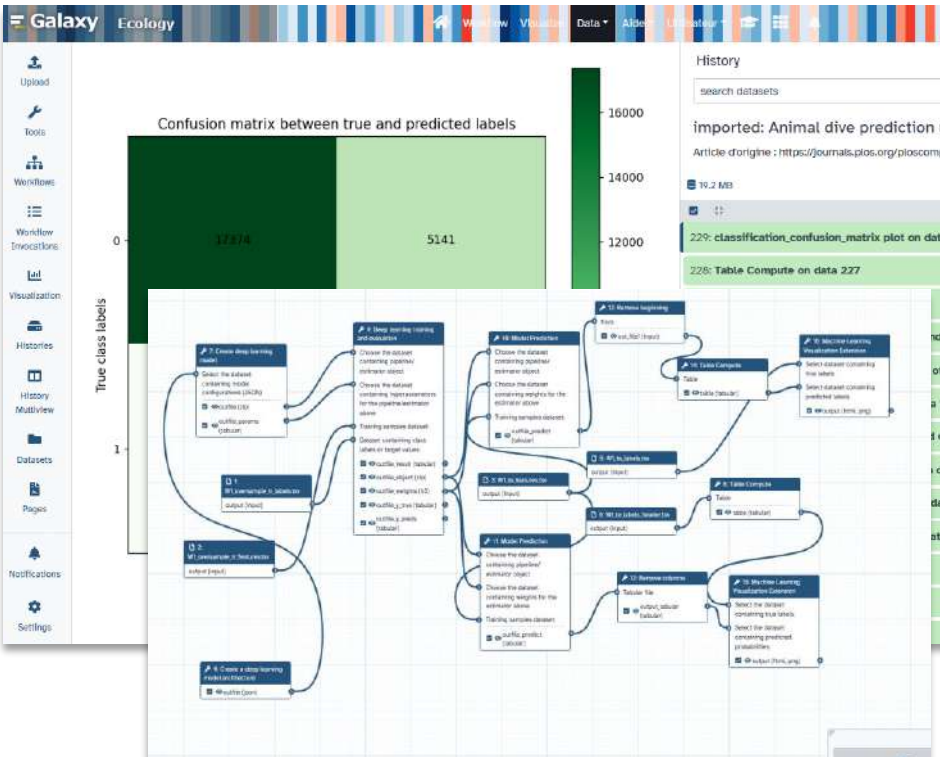


Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*

Adapt existing tools to use cases
From Interactive annotation to training
to labeling

Using Jupyter notebook
From camera trap video to species
detection with megadetector &
pytorch-wildlife

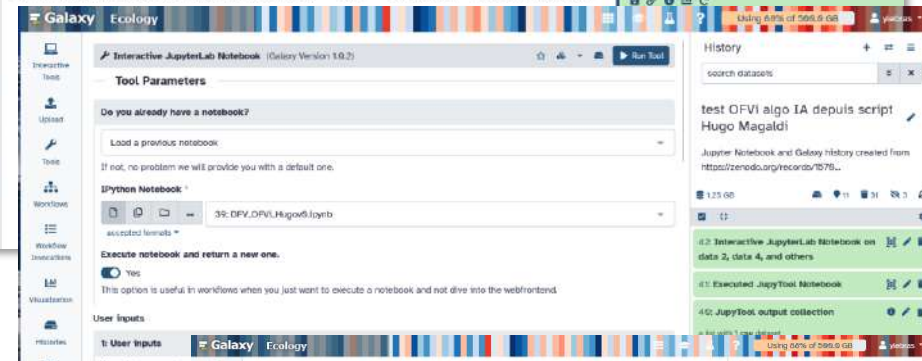


AnyLabeling Interactive
Effortless data labeling with AI support from
YOLO and Segment Anything!

Convert AnyLabeling JSON to YOLO text
with ultralytics

Perform YOLO training
with ultralytics

Perform YOLO image labeling
with ultralytics



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

Amédée Roy, Sophie Lanco Bertrand, Ronan Fablet

Version 2 Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1008990>

MOOREV



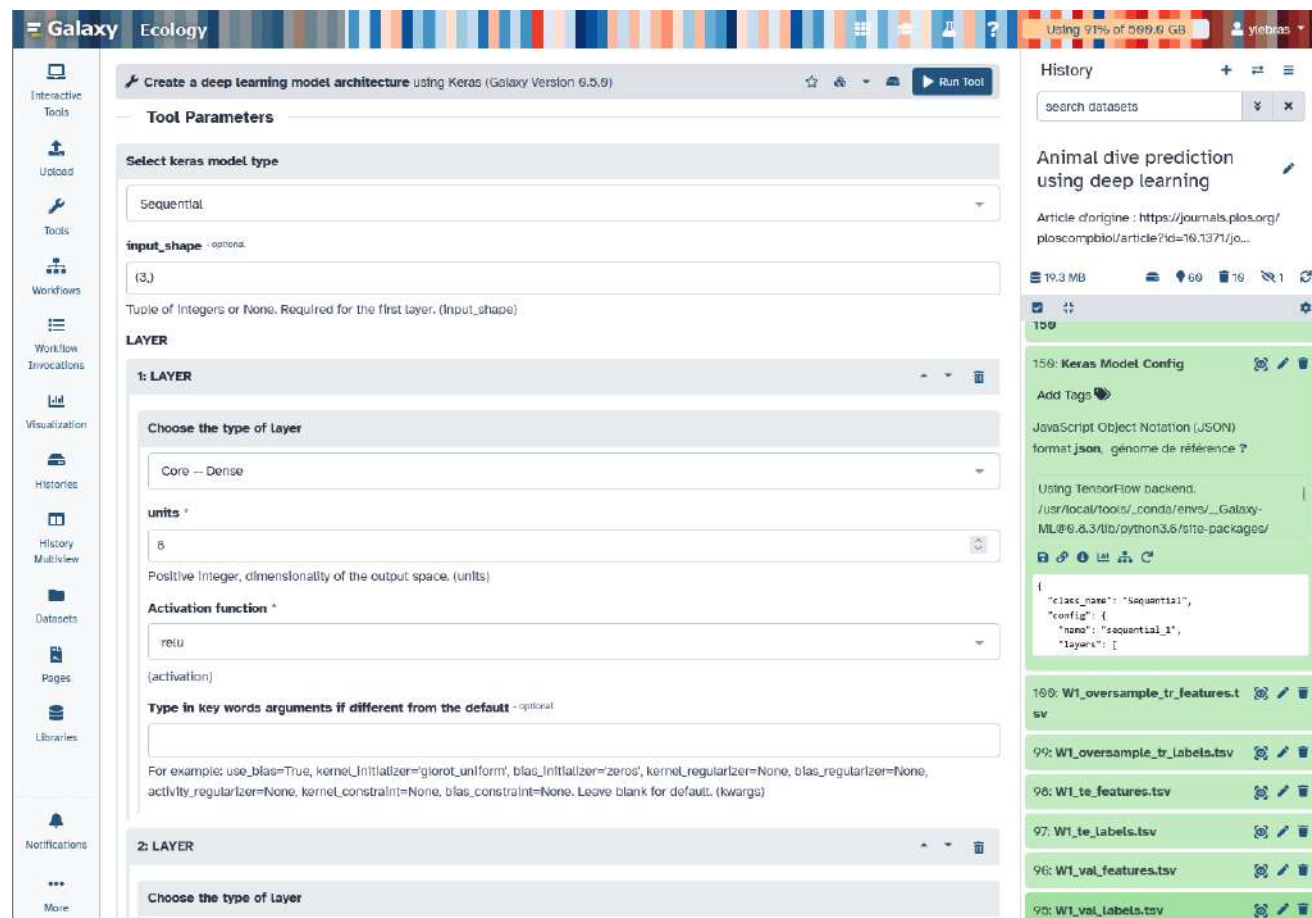
Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



RESEARCH ARTICLE

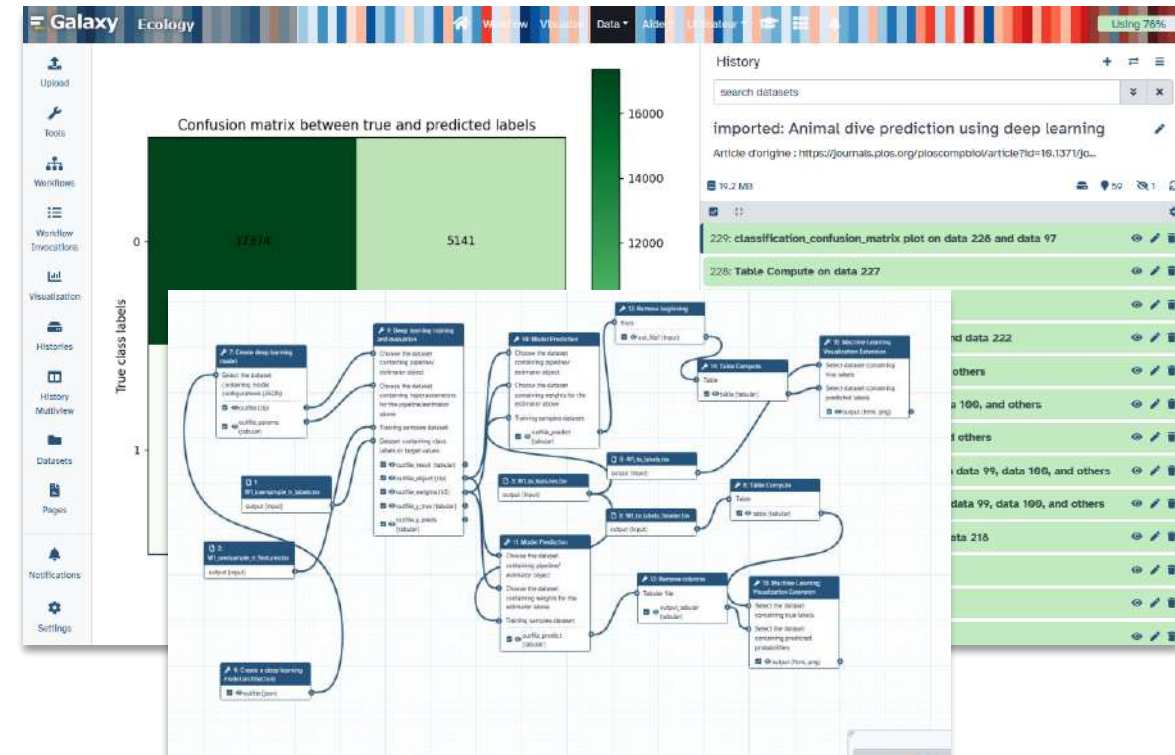
Amédée Roy , Sophie Lanco Bertrand, Ronan Fablet

Version 2 Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>



Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



Galaxy Ecology

Using 91% of 500.9 GB

Ylebras

Create deep learning model with an optimizer, loss function and fit parameters (Galaxy Version 0.5.0)

Run Tool

Tool Parameters

Choose a building mode

Build a training model

Select the dataset containing model configurations (JSON) *

159: Keras Model Config

accepted formats

Do classification or regression? *

KerasGClassifier

Compile Parameters

Select a loss function *

binary_crossentropy

Select an optimizer

Adam - Adam optimizer

Optimizer Advanced Options

Select metrics - optional

acc / accuracy * binary_accuracy *

switch to column select

Fit Parameters

epochs *

20

batch_size -optional

History

search datasets

Animal dive prediction using deep learning

Article d'origine : <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1009890>

152: get_params for Keras Model Builder on data 159

Add Tags

121 lines

format tabular, génome de référence ?

KerasGClassifier(amsgrad=False, batch_size=32, beta_1=0.9, beta_2=0.999, callbacks=[callback_selection])

151: Keras Model Builder on data 159

Add Tags

1.4 KB

format zip, génome de référence ?

KerasGClassifier(amsgrad=False, batch_size=32, beta_1=0.9, beta_2=0.999, callbacks=[callback_selection])

PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

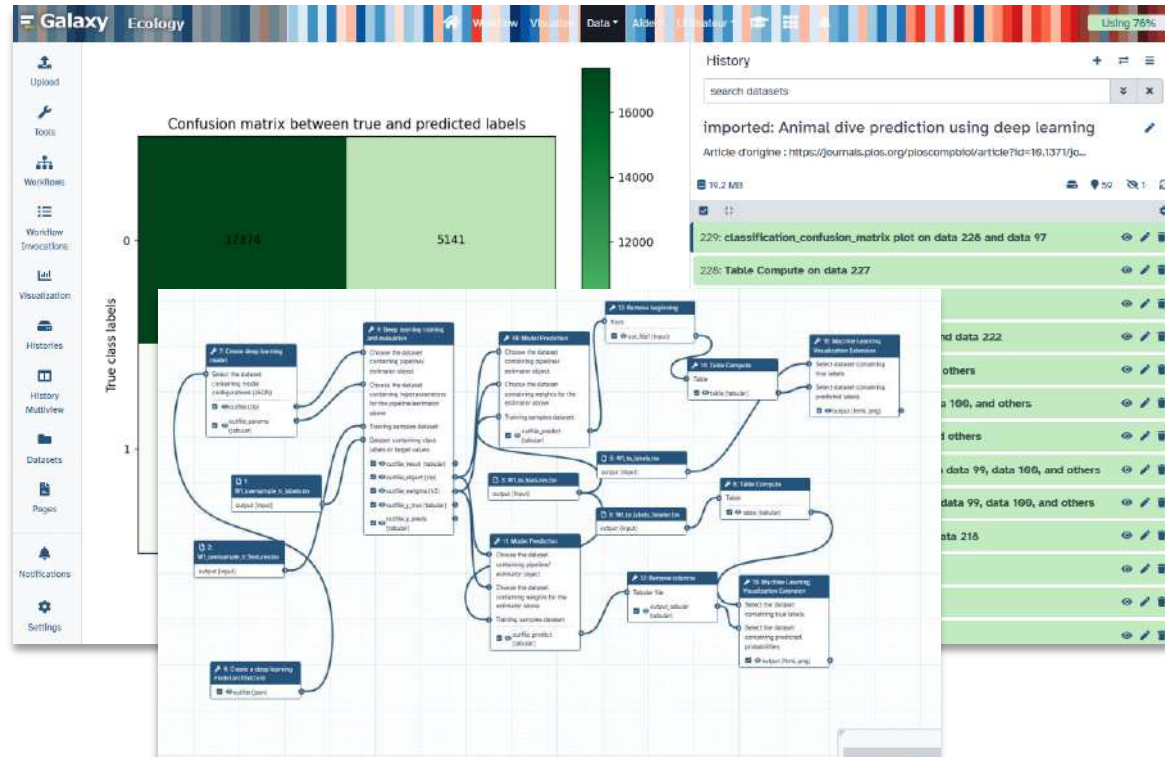
Deep inference of seabird dives from GPS-only records: Performance and generalization properties

Amédée Roy, Sophie Lanco Bertrand, Ronan Fablet

Version 2 Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>

Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

Amédée Roy, Sophie Lanco Bertrand, Ronan Fablet

Version 2 Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>

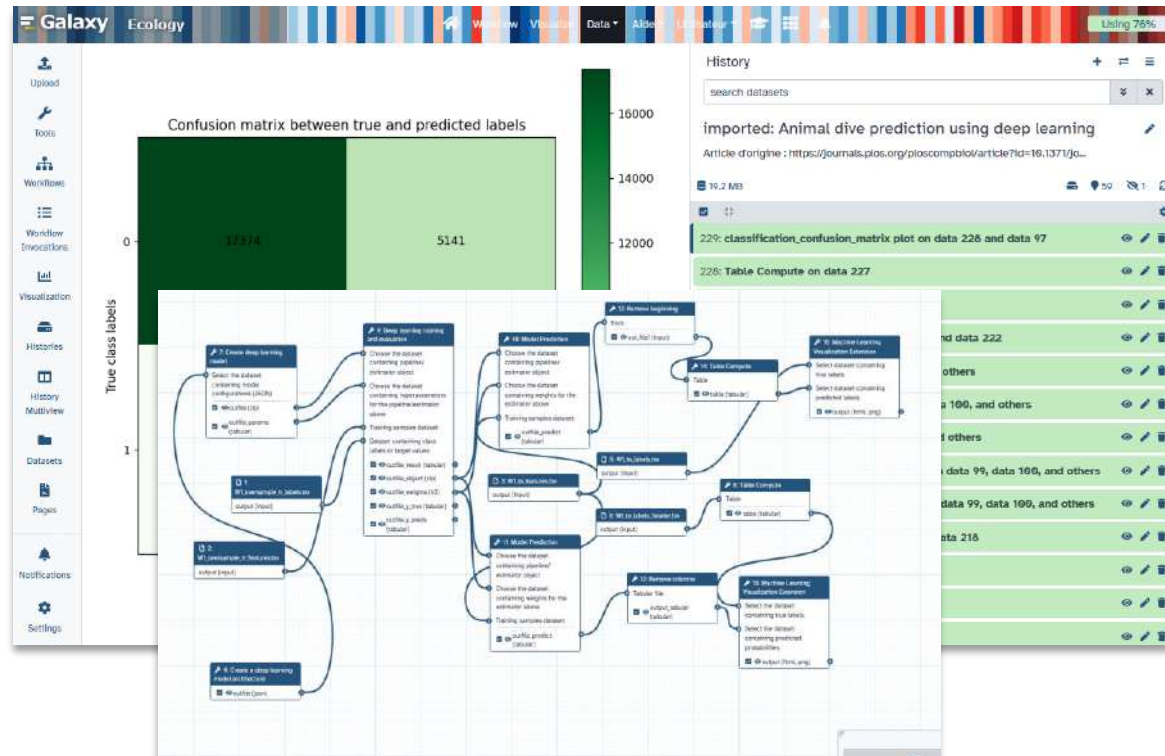
The screenshot displays the Galaxy web interface for the 'Deep learning training and evaluation' tool. The tool parameters are configured as follows:

- Select a scheme:** Train and Validate
- Choose the dataset containing pipeline/estimator object:** 151: Keras Model Builder on data 150
- Hyperparameter Swapping:** (Expanded)
- Validation holdout:** (Expanded)
- Metrics for evaluation:** (Expanded)
- Select input type:** tabular data
- Training samples datasets:** 109: W1_oversample_tr_features.tsv
- Does the dataset contain header:** No
- Choose how to select data by column:** All columns
- Dataset containing class labels or target values:** 99: W1_oversample_tr_labels.tsv

The 'History' panel on the right lists recent jobs, including 'Animal dive prediction using deep learning' and '163: All predictions on data 99, data 100, and others'.

Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

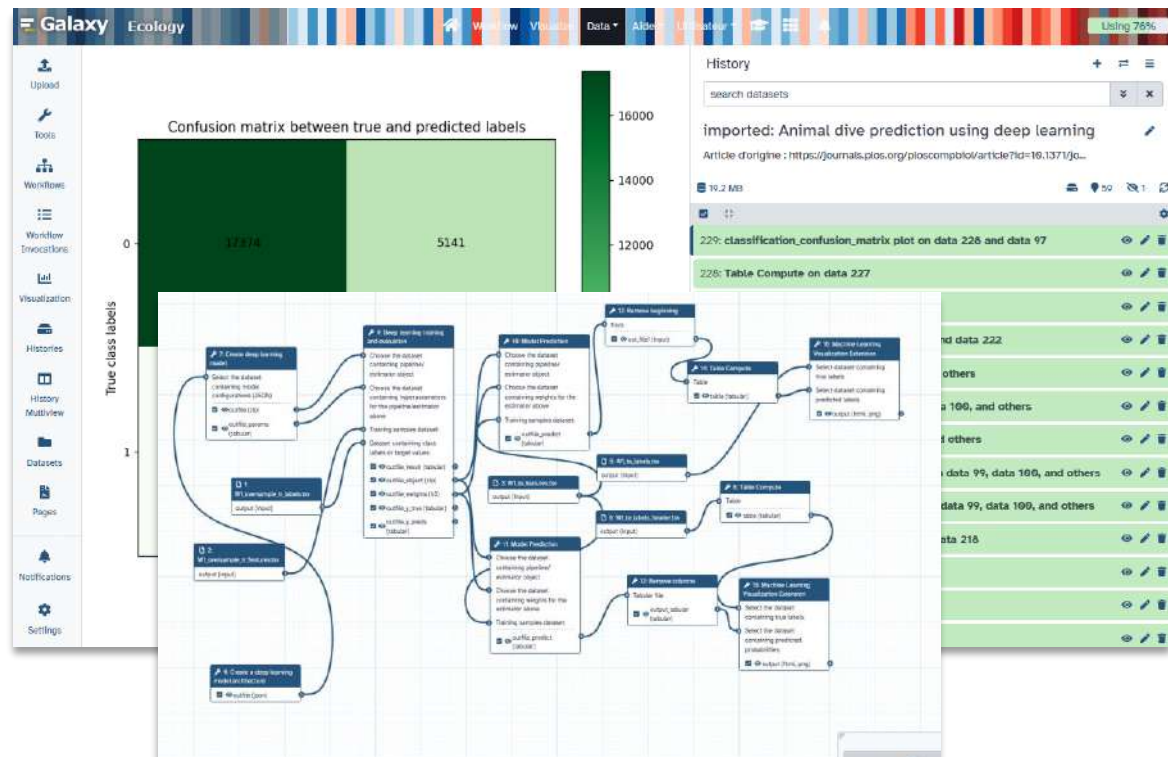
Amédée Roy, Sophie Lanco Bertrand, Ronan Fablet

Version 2 Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>

The screenshot shows the Galaxy web interface with the "Evaluate a Fitted Model" tool selected. The tool parameters are configured as follows: "Choose the dataset containing pipeline/estimator object" is set to "160: Fitted estimator or estimator skeleton on data 99, data 100, and others"; "Choose the dataset containing weights for the estimator above" is set to "161: Weights trained on data 99, data 100, and others"; "Select the primary metric (scoring)" is set to "Classification -- 'roc_auc'"; "Additional scoring used in multi-metric mode" is set to "optional"; "Select input types" is set to "tabular data"; "Dataset containing features" is set to "98: W10_features.tsv"; "Does the dataset contain header" is set to "No"; and "Choose how to select data by columns" is set to "switch to column select". The history panel on the right shows the workflow steps, including "160: Evaluate a Fitted Model on data 99, data 100, and others" and "162: True Labels/target values on data 99, data 100, and others".

Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



This screenshot shows the 'Model Prediction' tool interface in Galaxy. The tool parameters section includes:

- Choose the dataset containing pipeline/estimator object:** 160: Fitted estimator or estimator skeleton on data 99, data 166, and others.
- Choose the dataset containing weights for the estimator above:** 161: Weights trained on data 99, data 166, and others.
- Select invocation method:** predict.
- Select input data type for prediction:** tabular data.
- Training samples dataset:** 96: W1_t0_features.tsv.
- Does the dataset contain header:** No.
- Choose how to select data by column:** All columns.

The right panel shows the history of the tool, including '160: Model Prediction on data 96, data 161, and data 160' and '165: Evaluate a Fitted Model on data 97, data 98, and others'.

PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

Amédée Roy, Sophie Lanco Bertrand, Ronan Fablet

Version 2 Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>

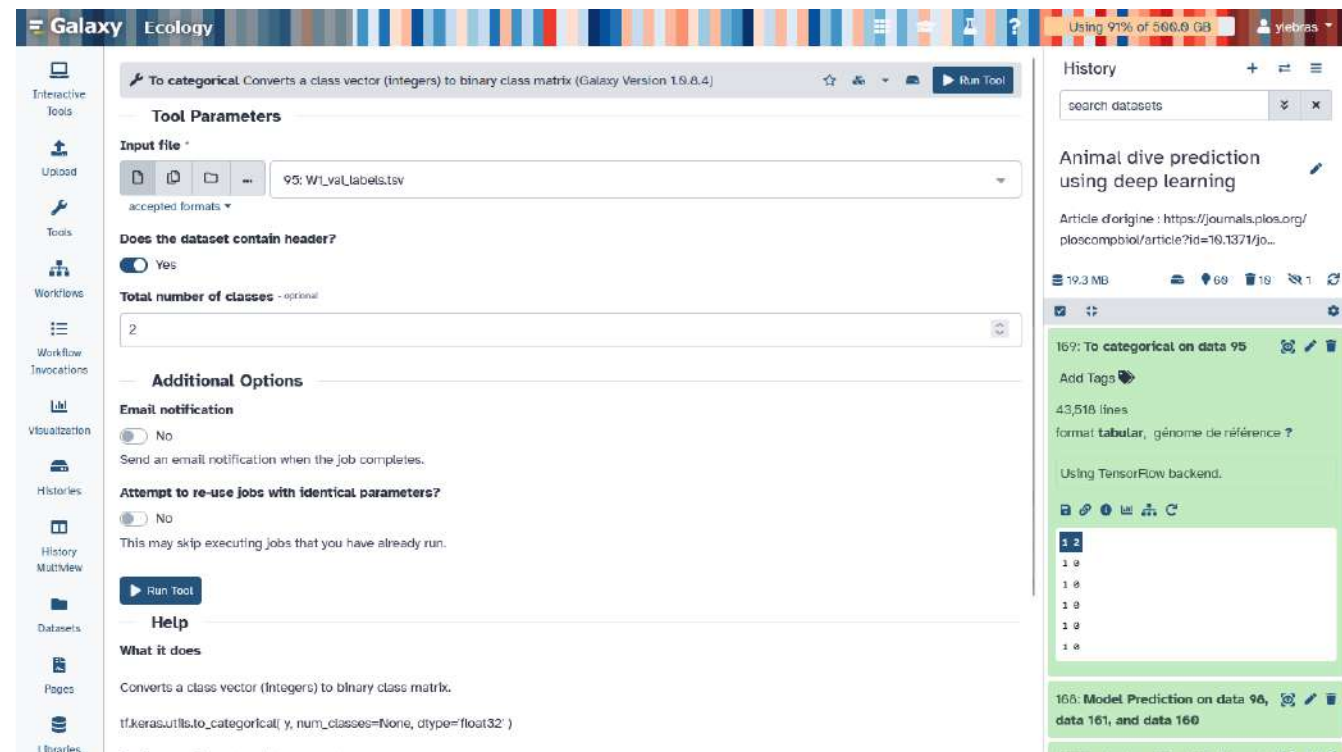
Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*



RESEARCH ARTICLE

Amédée Roy , Sophie Lanco Bertrand, Ronan Fablet

Version 2 Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>



Annotations & AI

Reproduce analyses from research article *Deep learning from creating the model to training and evaluation*

Machine Learning Visualization Extension includes several types of plotting for machine learning (Galaxy Version 1.9.8.4)

Tool Parameters

Select a plotting type

Confusion matrix for classes

Select dataset containing true labels *

97: W1_te_labels.tsv

accepted formats

Does the dataset contain header:

No

Choose how to select data by columns

All columns

Select dataset containing predicted labels *

228: Table Compute on data 227

accepted formats

Does the dataset contain header:

No

Plot title - optional

Confusion matrix between true and predicted labels

Choose plot color *

Greens

History

search datasets

Animal dive prediction using deep learning

Article d'origine : <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1009890>

10.3 MB

229: classification_confusion_matrix plot on data 228 and data 97

Confusion_matrix x Results x

Add Tags

34.6 KB

format png, génome de référence ?

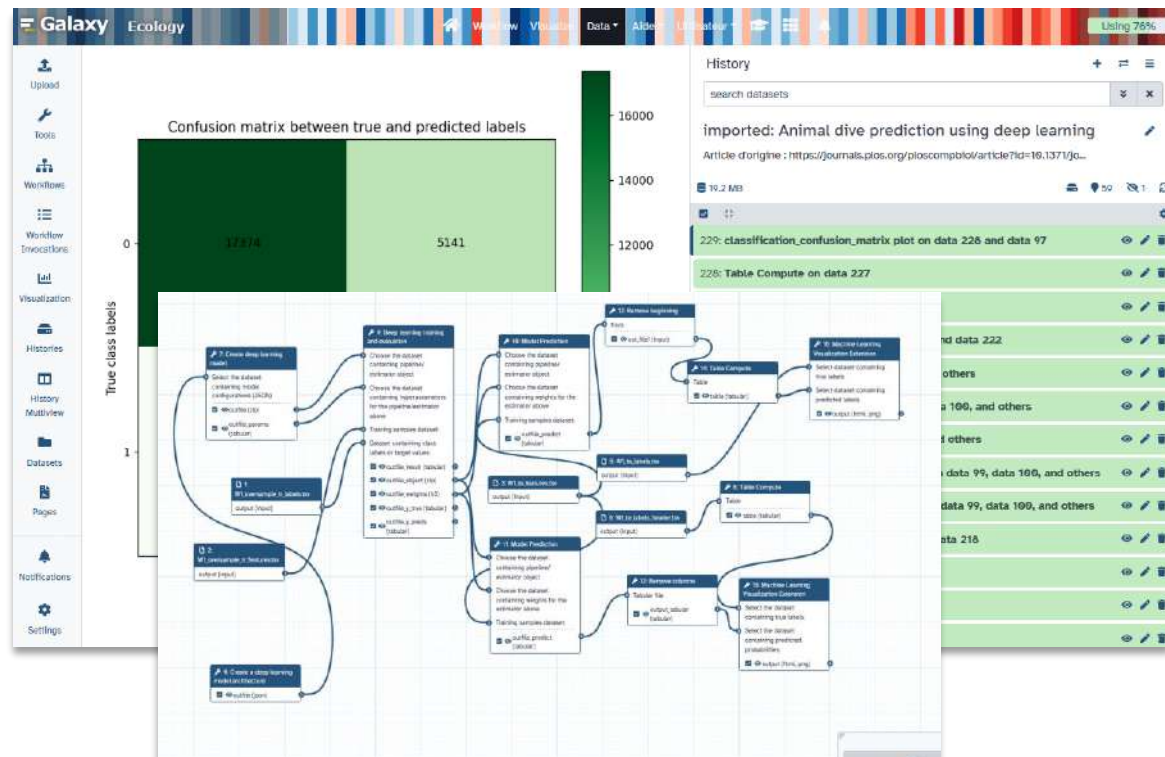
Using TensorFlow backend.

/usr/local/tools/conda/envs/_Galaxy-ML@9.8.3/lib/python3.6/site-packages/

Image in png format

228: Table Compute on data 227

227: Remove beginning on data 2



PLOS COMPUTATIONAL BIOLOGY

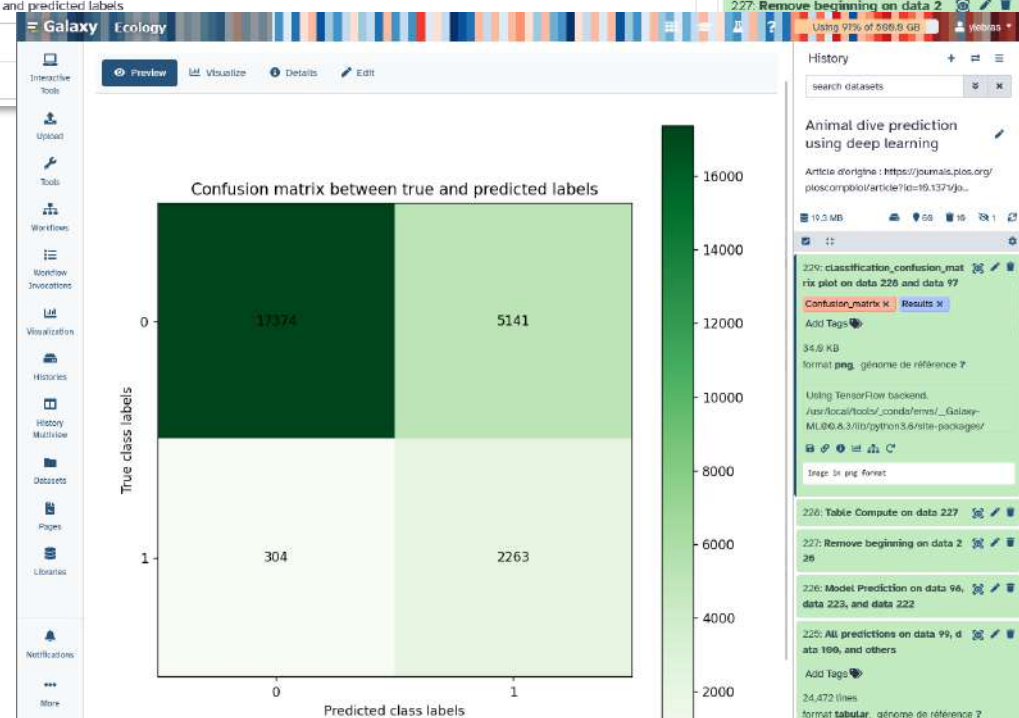
OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Deep inference of seabird dives from GPS-only records: Performance and generalization properties

Amédée Roy, Sophie Lanco Bertrand, Ronan Fablet

Version 2 Published: March 11, 2022 • <https://doi.org/10.1371/journal.pcbi.1009890>



Annotations & AI

Adapt existing tools to use cases
*From Interactive annotation to training
to labeling*



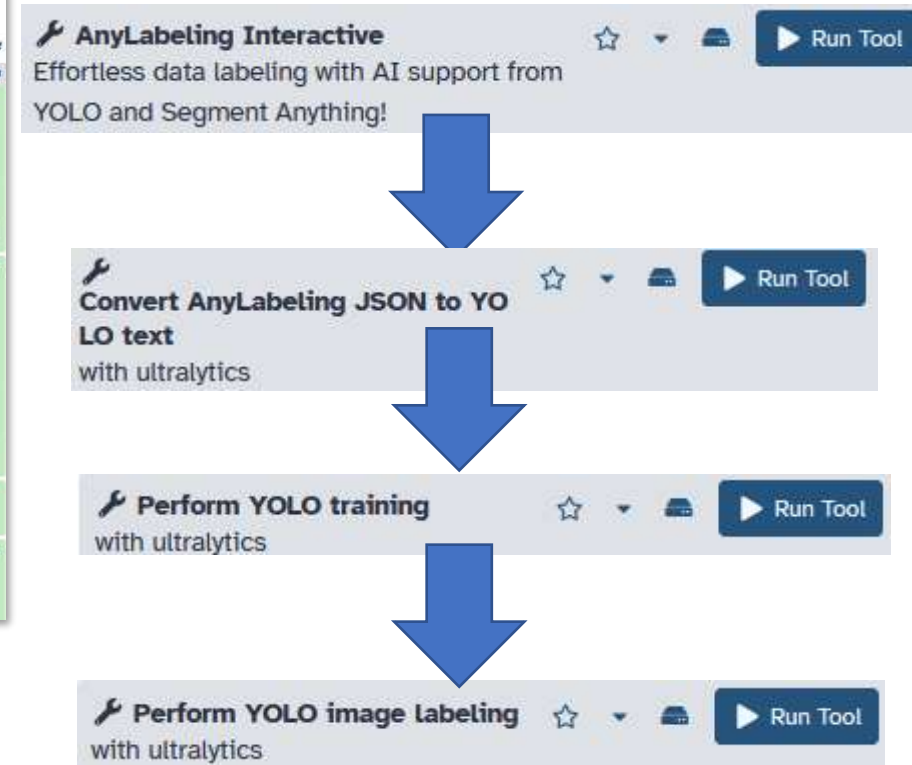
MOOREV



Annotations & AI



Adapt existing tools to use cases
*From Interactive annotation to training
to labeling*



Annotations & AI

Adapt existing tools to use cases
*From Interactive annotation to training
to labeling*

Galaxy

Ecology

Interactive Tools

Upload

Tools

Workflows

Workflow Invocations

Visualization

Histories

History Mainview

Datasets

Pages

Libraries

Notifications

Perform YOLO training with ultralytics (Galaxy Version 0.3.0+galaxy2)

Tool Parameters

Input Images

accepted formats

Search for options

Unselected (14)

131: MED-Annotation-Crispi2822_DSC_6494-copie.jpg

130: MED-Annotation_DSC63716-laser-Zoanthids-lasers.jpg

129: DSC69761.JPG

128: DSC69117-messif-madrepore-lophelia-en-cave.jpg

127: DSC66736-copie.JPG

126: DSC69291-copie.JPG

125: DSC69114-plastique-Lophelia-madrepore.jpg

124: Dpertusum-rose_DSC69226-laser.jpg

Selected (26)

64: C2-MAX_28236223_DE_W6662_P6661-6656.jpg

63: C2-MAX_28236223_DE_W6662_P6661-6654.jpg

62: C2-MAX_28236223_DE_W6662_P6661-6656.jpg

61: C2-MAX_28236223_DE_W6662_P6661-6647.jpg

60: C2-MAX_28236223_DE_W6662_P6661-6639.jpg

59: C2-MAX_28236223_DE_W6662_P6661-6637.jpg

58: C2-MAX_28236223_DE_W6662_P6661-6633.jpg

57: C2-MAX_28236223_DE_W6662_P6661-6636.jpg

Input YOLO txt files

accepted formats

Search for options

Unselected (21)

226: AnyLabeling Interactive on data 17: version.txt

213: Training Metrics

209: AnyLabeling Interactive on data 17: version.txt

206: AnyLabeling Interactive on data 17: version.txt

166: AnyLabeling Interactive on data 17: version.txt

146: MED-Annotation-Crispi2822_DSC_6494-copie.jpg

139: Dpertusum-rose_DSC69226-laser.jpg

138: MED-Annotation_DSC63716-laser-Zoanthids-lasers.jpg

Selected (26)

92: C2-MAX_28236223_DE_W6662_P6661-6656.txt

91: C2-MAX_28236223_DE_W6662_P6661-6654.txt

90: C2-MAX_28236223_DE_W6662_P6661-6656.txt

89: C2-MAX_28236223_DE_W6662_P6661-6647.txt

88: C2-MAX_28236223_DE_W6662_P6661-6639.txt

87: C2-MAX_28236223_DE_W6662_P6661-6637.txt

86: C2-MAX_28236223_DE_W6662_P6661-6633.txt

85: C2-MAX_28236223_DE_W6662_P6661-6636.txt

The YOLO text files, each text file must correspond to one input image (same name different extension).

model_url

YOLO11n-seg

Training Parameters

History

search datasets

Test MOOREV et yolo anylabeling tools

511 MB

216

6

4

214: Training Plot

213: Training Metrics

Add Tags

169 lines 21 columns, 1 comments

format csv, génome de référence ?

Creating new Ultralytics Settings v0.6 file

View Ultralytics Settings with 'yolo settings' or

1.epoch 2.time 3.train/box_loss 4.train/seg_loss

epoch time train/box_loss train/seg_loss

1 1.81749 1.27588 3.58629

2 2.13389 1.53666 2.76807

3 2.76657 1.39071 2.99332

4 1.1382 1.04353 5.75326

212: Last Model

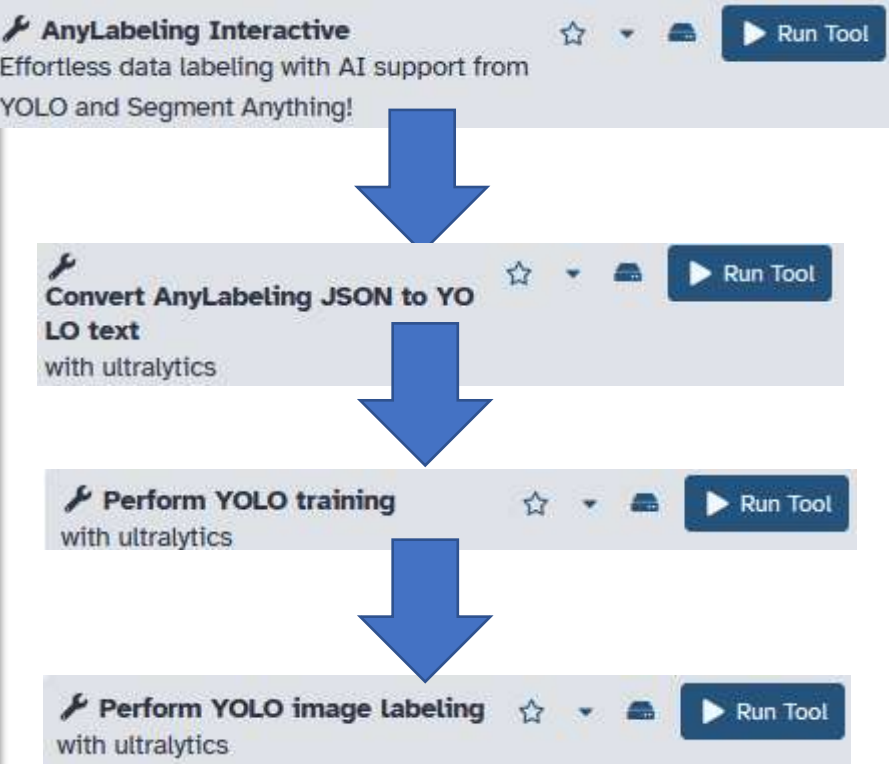
211: Best Model

Add Tags

5.7 MB

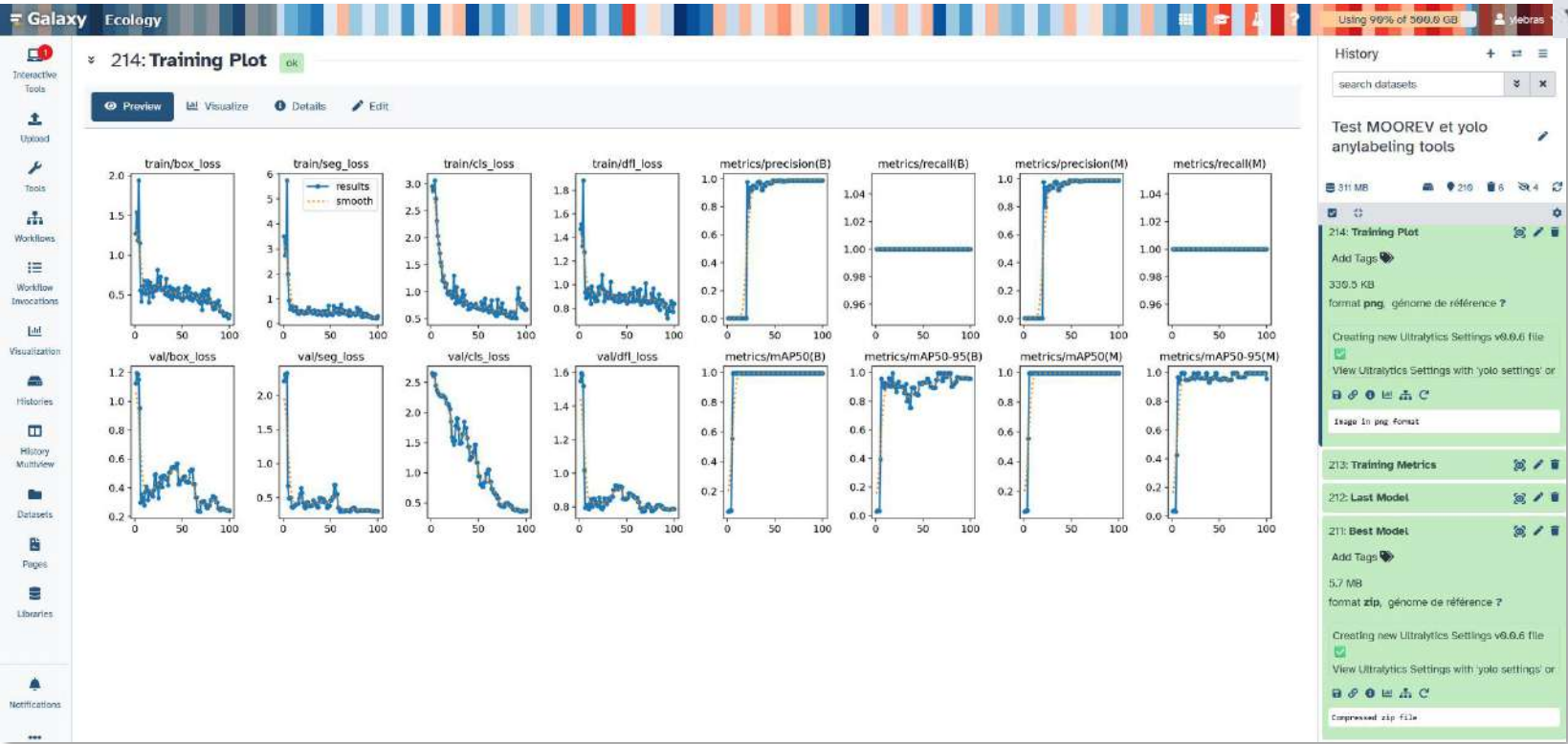
format zip, génome de référence ?

Creating new Ultralytics Settings v0.6 file



Annotations & AI

Adapt existing tools to use cases
*From Interactive annotation to training
to labeling*



AnyLabeling Interactive
Effortless data labeling with AI support from
YOLO and Segment Anything!

Convert AnyLabeling JSON to YOLO text
with ultralytics

Perform YOLO training
with ultralytics

Perform YOLO image Labeling
with ultralytics

MOOREV



Annotations & AI

Using Jupyter notebook
From camera trap video to species
detection with megadetector &
pytorch-wildlife

The image displays three screenshots of the Galaxy Ecology web interface, illustrating the workflow for species detection using JupyterLab notebooks.

Top Screenshot: Shows the '43: predictions_stride_1_thresh_0.2_2025-09-26 09-16-22' notebook. The 'Preview' tab displays a table of predictions:

	squirrel	water chevrotain	human	vehicle	Prediction	Confidence score
/ena	squirrel	water chevrotain	human	vehicle	Prediction	Confidence score
> 97	1.7414376462273278e-06	4.065219844225918e-06	0.0	0.0	elephant	0.9638436758335876

Middle Screenshot: Shows the '41: Executed JupyterLab Notebook' interface. The 'Tool Parameters' section asks 'Do you already have a notebook?'. The 'History' panel on the right shows a list of notebooks, including 'test OFVi algo IA depuis script Hugo Magaldi'.

Bottom Screenshot: Shows the '41: Executed JupyterLab Notebook' interface with the 'Code' tab selected. The code cell displays the following output:

```
Classifying... 80% | 24/27 [23:15<03:00, 62.77s/step]
Classifying... 93% | 25/27 [24:02<01:55, 57.95s/step]
Classifying... 96% | 26/27 [24:52<00:55, 55.53s/step]
Classifying... 100% | 27/27 [25:57<00:00, 52.37s/step]
Classifying... 100% | 27/27 [25:37<00:00, 56.95s/step]
```

The code cell also shows the following output:

```
CONSOLIDATING PREDICTIONS

In [18]: # Get a single prediction for each file (photo or video) by a majority vote over
the detections
gby_dict = {}
gby_dict['filepath'] = 'first'
gby_dict['Filename'] = 'first'
for taxon in taxons['human', 'vehicle']:
    gby_dict[taxon] = 'mean'
predictions_grouped = predictions.groupby('filepath').agg(gby_dict)
predictions_grouped['Prediction'] = predictions_grouped[taxons_all].apply(lambda
x: 'blank' if sum(x)==0 else taxons_all[np.argmax(x)], axis=1)
predictions_grouped['Confidence score'] = predictions_grouped[taxons_all].apply
(lambda x: 'blank' if sum(x)==0 else np.max(x), axis=1)
predictions_grouped.columns = [predictions_grouped.columns[1][0] for i in range
(len(predictions_grouped.columns))]
# Predict as 'blank' photos/videos for which there were no detections
for filepath in filepath_all:
    if filepath not in predictions_grouped['filepath']:
        prediction_blank = pd.DataFrame([filepath, filepath.split(os.sep,1)
[1]]+ [0]*(taxons.count2) + ['blank', 1-detection_threshold]), columns=list(pre
dictions_grouped.columns))
        predictions_grouped = pd.concat([predictions_grouped, prediction_blank])
predictions_grouped = predictions_grouped.reset_index(drop=True)
predictions_grouped = predictions_grouped.sort_values(by='filepath', axis=0)
predictions_grouped.to_csv(os.path.join(predictions_dir, 'predictions_' + 'stpd
e_' + str(stride) + '_thresh_' + str(detection_threshold) + '.' + timestamp + '.c
sv'), index = False)
print('PREDICTIONS SUCCESSFULLY SAVED TO FOLDER: ' + str(predictions_dir))

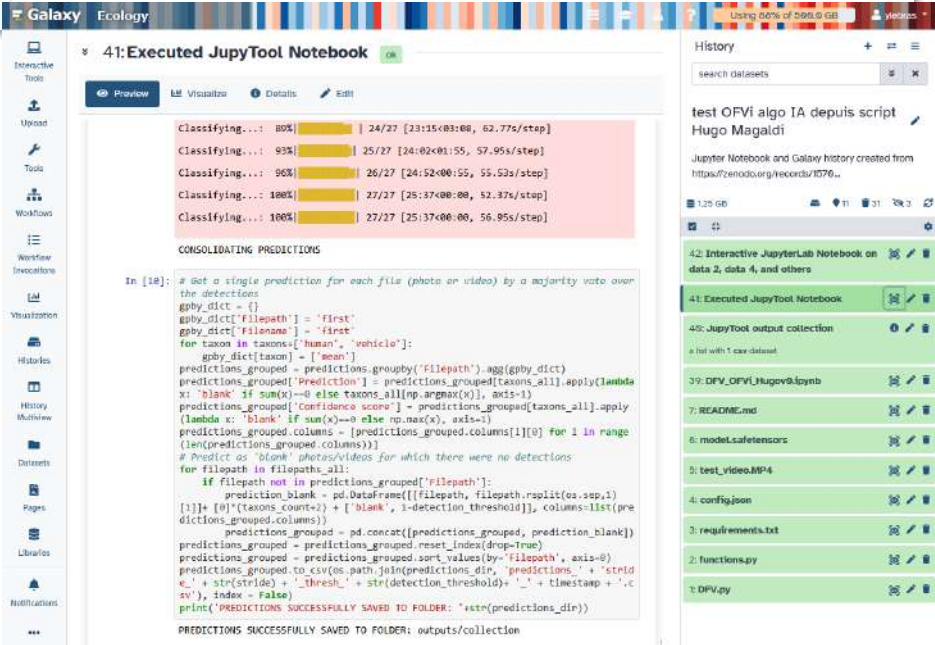

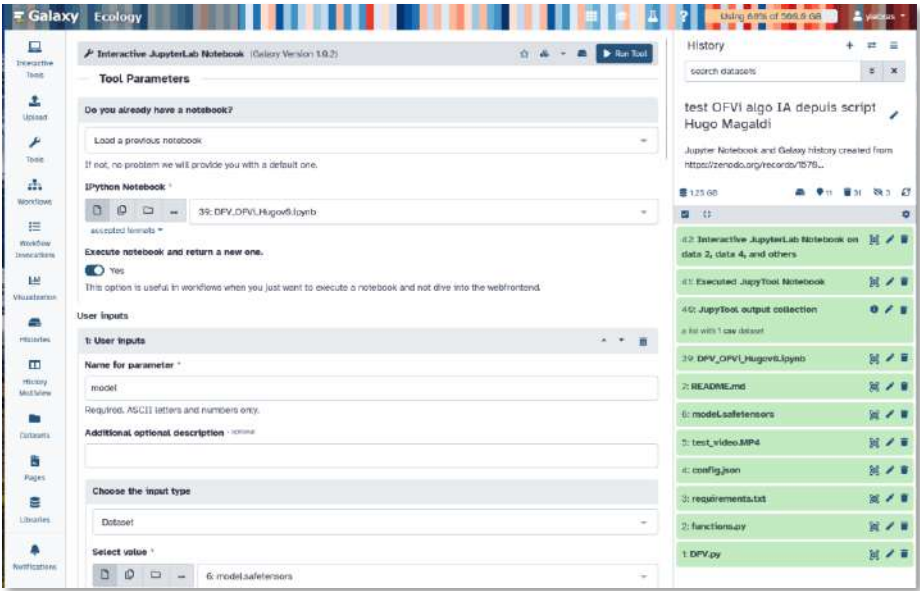
PREDICTIONS SUCCESSFULLY SAVED TO FOLDER: outputs/collection
```

The 'History' panel on the right shows a list of notebooks, including '42: Interactive JupyterLab Notebook on data 2, data 4, and others' and '41: Executed JupyterLab Notebook'.



Annotations & AI

Using Jupyter notebook
From camera trap video to species
detection with megadetector &
pytorch-wildlife



Annotations & AI

New: Using Hugging Face models in Galaxy Ecology

— User Preferences / My Repositories / Create New / Hugging Face Hub 😊

Create a Hugging Face Hub 🤖 File Source

Name *

My Hugging Face Account

Label this new file source with a name.

Description - optional

Provide some notes to yourself about this file source - perhaps to remind you how it is configured, where it stores the data, etc..

Hugging Face Hub Endpoint

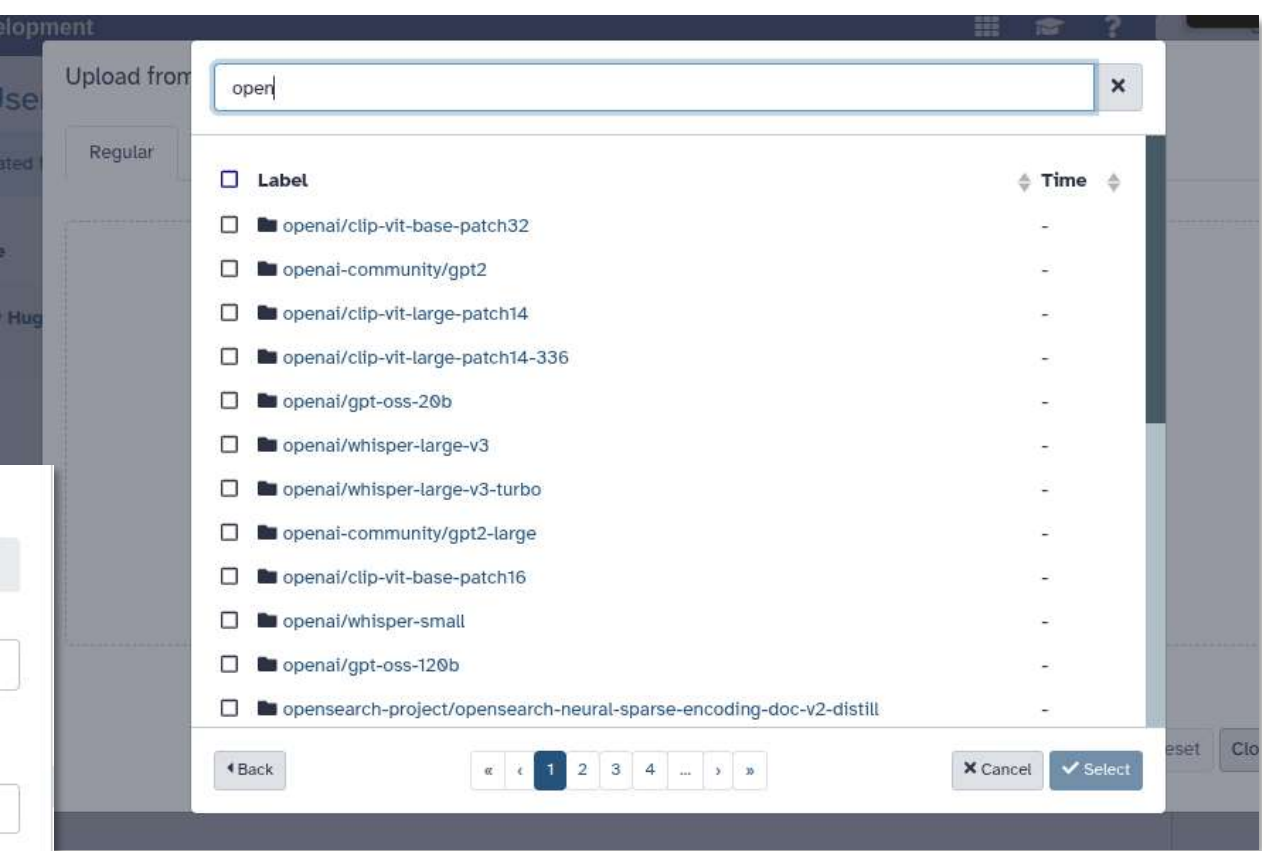
https://huggingface.co

Custom endpoint of the Hugging Face Hub you are connecting to. This should be the full URL including the protocol (http or https) and the domain name. You can leave this blank to use the default Hugging Face Hub endpoint (https://huggingface.co).

Hugging Face Access Token

The personal access token to use to connect to the Hugging Face Hub. You can generate a new token in your Hugging Face account settings. This will allow Galaxy to access private models if you have the necessary permissions.

Create



Galaxy Ecology initiative: A lot of young contributors!



Clara Urfer

Elisa Michon

Valentin Chambon

Timothée Virgoulay

Eloïse Trigodet

Alan Amossé

Claire Dussin



Elouan Le Mestric



Jean Le Cras



Olivier Norvez

Elie Arnaud

Yassine Ankerl



Molène Mahé



Kévin Payet



Najat Amoukou



Arthur Barreau



Triskell Cumunel



Marie Jossé



Pauline Segueineau



Simon Bénateau



Benjamin Yguel



Coline Royaux



2018

2019

2020

2021

2022

2023

2024

2025