**PenguinCounter User Guide for PointBlue**

Tim DeBenedictis ([tdebenedictis-RA@pointblue.org](mailto:tdebenedictis-RA@pointblue.org))

Version 1.0 • 21 Feb 2023

PenguinCounter is a program that refines YOLO’s raw penguin predictions and visualizes them on a scaled-down version of an orthomosaic. It generates statistics on the sizes of penguins predicted by YOLO, and compares the predictions against human-validated penguin labels. Finally, it generates a miniaturized version of an input orthomosaic.

Terminology:

* **Orthomosaic:** a single image of an entire penguin colony stitched together from many individual drone images, usually hundreds of thousands of pixels on a side.
* **Tile:** a small subsection of an orthomosaic, usually 512x256 pixels. Tiles may overlap by 20 pixels!
* **Prediction:** a penguin found by YOLO within a tile. Has a **class** (adult, adult-stand, or chick), a **probability** (likelihood of belonging to that class), and a **bounding box** with a center position (x,y) and size (w,h).
* **Validation Label:** a penguin labelled by a human within a tile. Has same properties as a Prediction, but probability is 1.0.
* **Validated Tile:** a tile that has been inspected by a human, and contains zero or more validation labels.

PenguinCounter is written in C++. It runs on Linux or MacOS, and can probably be compiled for Windows, too. These instructions describe how to compile and run PenguinCounter.

**Running**

These instructions assume you have an orthomosaic on your local machine, as well as directories containing YOLO predictions and validation labels. All inputs to PenguinCounter are supplied as command line arguments, in the following order:

ortho index adults stands chicks validations val\_map raw\_pred\_map ref\_pred\_map small\_ortho

Here is a description of each argument.

* ortho input path to full-size orthomosaic in GeoTIFF format
* index input path to tile index in CSV format
* adults input path to directory containing adult penguin predictions in TOLO .txt format
* stands input path to directory containing adult standing penguin predictions in TOLO .txt format
* chicks input path to directory containing penguin chick predictions in TOLO .txt format
* validations input path to directory containing human-validated labels in YOLO .txt or .csv format
* val\_map output path to validations map image, or “none”
* raw\_pred\_map output path to raw predictions map image, or “none”
* ref\_pred\_map output path to refined predictions map image, or “none”
* small\_ortho output path to small version of input orthomosaic, or “none”

The following example assumes you are running on a Mac or Linux host.

* You have an orthomosaic **croz\_20211127.tif** located in a directory **orthos/2021-11-27**
* The tile index file **croz\_20211127\_tilesGeorefTable.csv** is located in **tiles/2021-11-27**
* Your YOLO predictions are located in **counts/2021-11-27/adults\_s2\_best, adult\_stand\_s5\_best,** and **chick\_s\_best**
* Your validation data **croz\_20211127\_validation\_labels.csv** located in **counts/2021-11-27**
* You want the output images **validations.png, raw\_predictions.png**, **refined\_predictions.png**, and **croz\_20211127\_small.jpg** to be saved in a directory **output/2021-11-27**

Here is the command to run PenguinCounter to generate counts and output images from those inputs. Yes, this is all one line!

./counter orthos/2021-11-27/croz\_20211127.tif \

tiles/2021-11-27/croz\_20211127\_tilesGeorefTable.csv \

counts/2021-11-27/adult\_s2\_best \

counts/2021-11-27/adult\_stand\_s5\_best \  
counts/2021-11-27/chick\_s\_best \

counts/2021-11-27/croz\_20211127\_validation\_labels.csv \

output/2021-11-27/validations.png \

output/2021-11-27/raw-predictions.png \

output/2021-11-27/refinded-predictions.png \

output/2021-11-27/croz\_20211127\_small.jpg

An example of successful output looks like this:

Read metadata from /home/PointBlue/orthos/croz\_20211127.tif; width = 185998; height = 178549

Allocated storage for 286903 tiles in 379 rows x 757 cols.

Read tile index /home/PointBlue/tiles/croz\_2021-11-27/croz\_20211127\_tilesGeorefTable.csv with 65456 entries.

Found 318561 adult predictions in /home/PointBlue/counts/croz\_2021-11-27/adult\_s2\_best/labels

Found 235091 adult stand predictions in /home/PointBlue/counts/croz\_2021-11-27/adult\_stand\_s5\_best/labels

Found 40810 chick predictions in /home/PointBlue/counts/croz\_2021-11-27/chick\_s\_best/labels

Found 22280 tiles with no predictions.

Read 9145 validation labels from /home/PointBlue/counts/croz\_2021-11-27/validation\_data/croz\_20211127\_validation\_labels.csv

Counted 1000 tiles with validation labels.

Counted 6596 adult validation labels.

Counted 2105 adult stand validation labels.

Counted 0 chick validation labels.

Counted 444 validated tiles with no validation labels.

Counted 8006 adult predictions in validated tiles.

Counted 5418 adult stand predictions in validated tiles.

Counted 1044 chick predictions in validated tiles.

Counted 451 validated tiles with no predictions.

Wrote validations map /home/PointBlue/counts/croz\_2021-11-27-validations.png

Wrote raw predictions map /home/PointBlue/counts/croz\_2021-11-27-predictions-raw.png

Predicted Adult Min sizex=10.000 sizey=8.000

Predicted Adult Max sizex=512.000 sizey=256.000

Predicted Adult Mean sizex=31.949 sizey=32.332

Predicted Adult Stdv sizex=10.125 sizey=9.374

Predicted Adult Stand Min sizex=7.000 sizey=9.000

Predicted Adult Stand Max sizex=512.000 sizey=256.000

Predicted Adult Stand Mean sizex=29.564 sizey=27.641

Predicted Adult Stand Stdv sizex=34.606 sizey=18.741

Predicted Chick Min sizex=10.000 sizey=10.000

Predicted Chick Max sizex=512.000 sizey=256.000

Predicted Chick Mean sizex=30.363 sizey=30.364

Predicted Chick Stdv sizex=9.846 sizey=8.933

Validated Adult Min sizex=5.000 sizey=3.000

Validated Adult Max sizex=57.000 sizey=56.000

Validated Adult Mean sizex=29.023 sizey=30.903

Validated Adult Stdv sizex=7.444 sizey=8.415

Deleted 1482 outsized Adult penguins.

Validated Adult Stand Min sizex=3.000 sizey=5.000

Validated Adult Stand Max sizex=50.000 sizey=49.000

Validated Adult Stand Mean sizex=25.044 sizey=25.357

Validated Adult Stand Stdv sizex=5.202 sizey=5.786

Deleted 3974 outsized Adult Stand penguins.

Validated Chick Min sizex=inf sizey=inf

Validated Chick Max sizex=-inf sizey=-inf

Validated Chick Mean sizex=0.000 sizey=0.000

Validated Chick Stdv sizex=0.000 sizey=0.000

Deleted 40810 outsized Chick penguins.

Deleted 0 improbable Penguin predictions.

Deleted 170259 duplicate Penguin predictions.

Deleted 53 duplicate Penguin validation labels.

Counted 6352 adult predictions in validated tiles.

Counted 2670 adult stand predictions in validated tiles.

Counted 0 chick predictions in validated tiles.

Counted 465 validated tiles with no predictions.

Counted 250518 adult predictions in all tiles.

Counted 127419 adult stand predictions in all tiles.

Counted 0 chick predictions in all tiles.

Counted 26090 tiles with no predictions.

Wrote refined prediction map /home/PointBlue/counts/croz\_2021-11-27-predictions-refined.png

Confusion Matrix:

TP 7447 FP 1575

TN 213 FN 473

Classification Matrix:

0 0 0 0

0 1385 274 0

0 310 5478 0

0 0 0 0

TIFF width: 185998

TIFF height: 178549

TIFF tile width: 0

TIFF tile height: 0

TIFF bits per sample: 8

TIFF samples per pixel: 4

TIFF planar config: 1

TIFF photometric: 2

TIFF compression: 5

TIFF scanline size: 743992

Processed row 0 of 5580.

Processed row 1 of 5580.

Processed row 2 of 5580.

…

Processed row 5578 of 5580.

Processed row 5579 of 5580.

Wrote small version of ortho /home/PointBlue/orthos/croz\_20211127\_small.jpg

The output maps are scaled-down versions of the original orthomosaic. Colored pixels in the output maps represent individual penguins:

* **Red** pixels are adult penguins
* **Green** pixels are adult\_stand penguins
* **Blue** pixels are chick penguins

The final output image is a small (1/32 scale) version of the original orthomosaic. Each pixel in the scaled-down version is the average color of the corresponding 32x32 pixel area in the full-size orthomosaic.

**Confusion and Classification Matrices**

The confusion matrix shows the number of true and false positives and negatives for penguins in tiles that a human has validated. These are defined as follows:

* **TP = True Positive**: a penguin predicted by YOLO matched a penguin labelled by a human.
* **FP = False Positive**: a penguin predicted by YOLO did not match any penguin labelled by a human.
* **FN = False Negative**: no penguin predicted by YOLO matched any penguin labelled by a human.
* **TN = True Negative**: a tile where YOLO predicted zero penguins and humans also labelled zero penguins.

Here, *match* means: the center of one penguin’s bounding box fell within another’s bounding box. For the confusion matrix, the program does not care what class of penguin (adult, adult\_stand, chick) the penguin predicted by YOLO (or labelled by a human) belongs to; they simply have to match (or not).

This is different in the classification matrix, which shows the number of penguins classified correctly (or incorrectly) in each category. The classification matrix is a 4x4 table showing the number of YOLO-predicted adult, adult\_stand, and chick penguins matched with human-labelled penguins in each category, in all of the human-validated tiles.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **none (predicted)** | **adult (predicted)** | **adult\_stand (predicted)** | **chick (predicted)** |
| **none (labelled)** |  |  |  |  |
| **adult (labelled)** |  |  |  |  |
| **adult\_stand (labelled)** |  |  |  |  |
| **chick (labelled)** |  |  |  |  |

Ideally, the diagonal elements of this matrix should have the highest counts, and the off-diagonal elements would be zero. This would mean YOLO is classifying penguins in exactly the same categories that humans labelled them.

**Compiling**

[TBD]

PenguinCounter has dependencies on the following common C and C++ libraries:

* **OpenCV** for image processing
* **LibTIFF** for raw TIFF image file access (OpenCV cannot handle images with more than 231 pixels!)
* **LibJPEG** for raw JPEG image file access (same reason).
* **LibGDAL** for extracting geotransform data from GeoTIFF images.

The build script contains lines to install these on Ubuntu linux if they are not already present.

A few constants are hard-coded in main.cpp. Their names, descriptions, and current values are as follows.

|  |  |  |
| --- | --- | --- |
| **Constant** | **Description** | **Value** |
| kTileWidth | tile width in pixels | 512 |
| kTileHeight | tile height in pixels | 256 |
| kTileOverlap | tile overlap in pixels | 20 |
| kOutputScale | scale of output images relative to input orthomosaic | 1/32.0 |
| kMinProbability | minimum probability of YOLO predictions to accept | 0.01 |

You may wish to change them and recompile the program if (for example) a different tile size is used in future YOLO analysis.

**Under the Hood**

Here is what PenguinCounter does.

1. Reads metadata (width, height, geotransform) from the input orthomosaic, but does not read raster pixel data.
2. Reads tile index and allocates storage for all tiles generated from the orthomosaic.
3. Reads YOLO predictions of adult, adult-standing, and chick penguins.
4. Reads validation data.
5. Counts validation labels of adult, adult\_stand, and chick penguins.
6. Counta predictions of adult, adult stand, and chick penguins in validated tiles
7. Converts all penguin positions (both predictions and validations) from tile (local) to ortho (global) coordinates.
8. Writes output maps of validation labels and raw predictions.
9. Get statistics on the positions and sizes of adult, adult stand, and chick penguins predicted by YOLO, and validated by humans.
10. Deletes YOLO penguin predictions with very low probabilities below kMinProbability as defined above.
11. Deletes YOLO penguin predictions larger or smaller than the largest/smallest validation labels for each class.
12. Deletes YOLO penguin predictions and human-validated labels duplicated across overlapping tile edges.
13. Counts refined (de-duplicated) predictions both in validated tiles, and in all tiles.
14. Writes refined (de-duplicated) penguin prediction map.
15. Generates confusion matrix and classification matrix.
16. Writes scaled-down (small) version of original input orthomosaic.

For de-duplication, the program looks for penguins which overlap each other. Any penguin predictions whose bounding-box center falls inside the bounding box of another penguin is considered to be a duplicate. In this case, the lower-probability penguin is deleted, and the higher-probability penguin survives.