AI.FISH

DFO Phase IIB: Computer Vision Results

Review of Approach

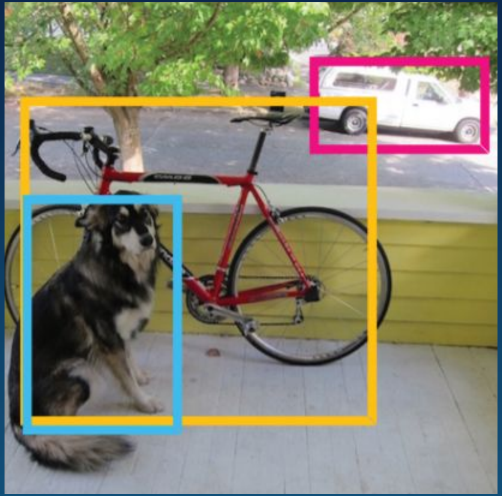
Goal: Predict species apportionment in an image

This involves:

- Localizing all fish in an image
  - 100s or 1000s or targets
- Classifying the species of all localized fish
- Computing proportions from these counts

Review of Approach

Existing computer vision algorithms inadequate.



Object detection

- Localization (boxes) + classification
- But **cannot handle dense scenes**

Crowd localization

- Localization in dense scenes
- But **cannot classify**

Review of Approach

AI.FISH

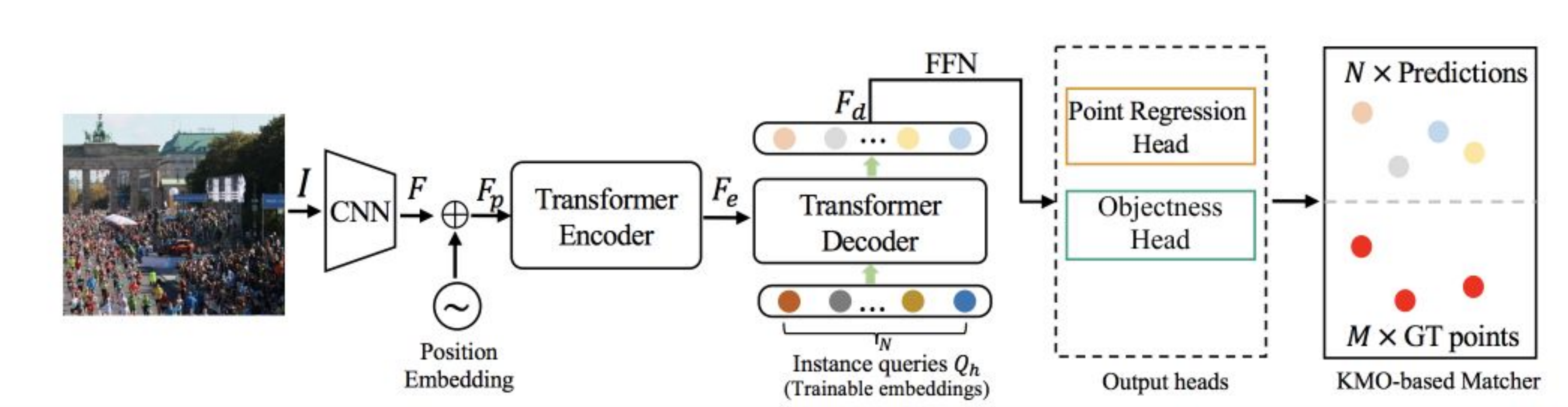
Goal: Dense localization and classification.

Dataset

- 9 vessels, 39 trips, 471 tows of raw data
- Three step annotation procedure:
    1. Localizations (“dots”)
    2. Expert species classifications
    3. Expert review
- Collected 477,889 annotations in 3,362 images

Algorithm Details

Starting point: Crowd Localization Transformer (CLTR)

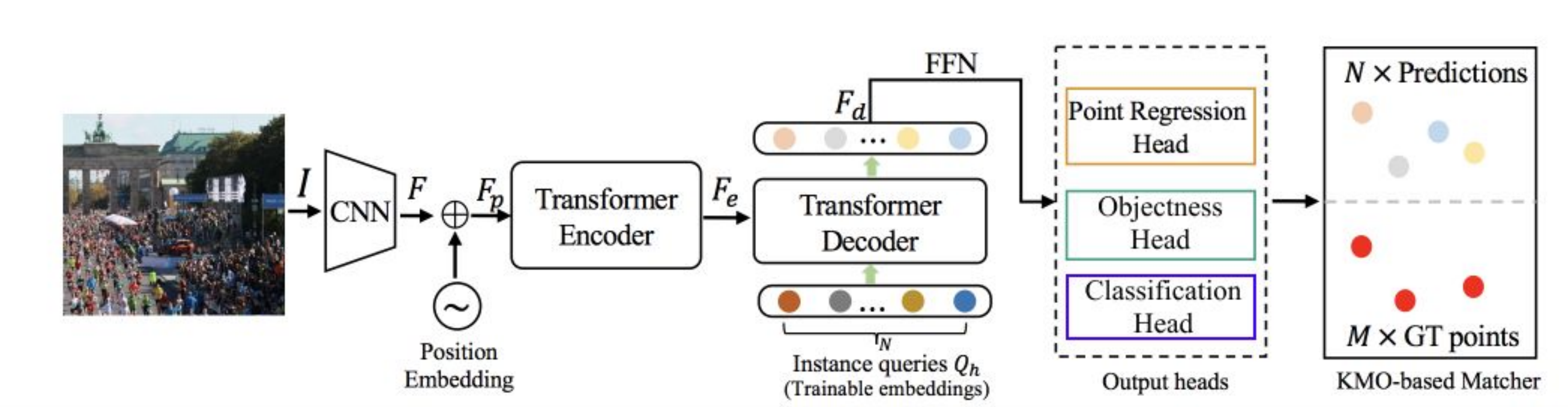


- State of the art for crowd localization
- But **cannot perform classification**

Liang, D., Xu, W. and Bai, X., 2022, October. An end-to-end transformer model for crowd localization. In European Conference on Computer Vision (pp. 38-54). Cham: Springer Nature Switzerland

Algorithm Details

Add a classification branch



- Additional neural network layers
- Additional classification loss during training (softmax + cross-entropy)
- Now **each point has a species classification**

Metrics

Two metrics were developed to evaluate our model in the context of the apportionment task:

1. **Dominant Species Accuracy**
2. **Weighted Classification Error**

We evaluated the model on its own (using our test dataset) as well as in comparison to trained reviewers.

Metrics

AI.FISH

Dominant Species Accuracy

Accuracy of predicting the most common class.

Typically an image is dominated by one species; predicting this species correctly will have the largest effect on overall accuracy.

This metric provides a simple "at a glance" measure of how well we do at identifying the majority class.

Metrics

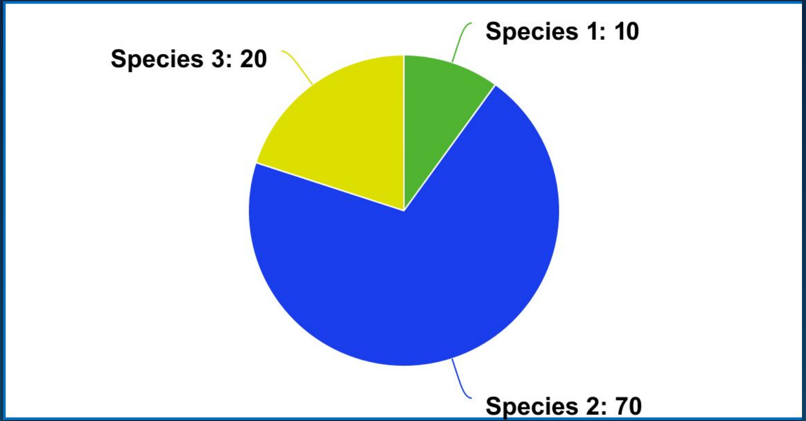
Weighted Classification Error

A more complex metric that takes into account all species present as well as the apportionment goal.

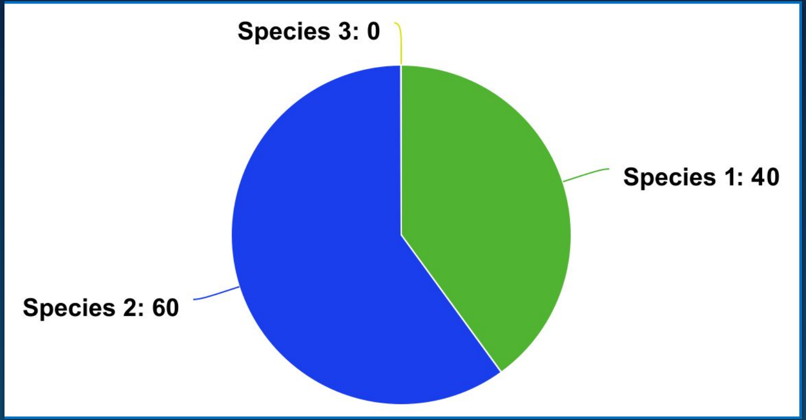
Metrics

Weighted Classification Error

Ground Truth



Predicted

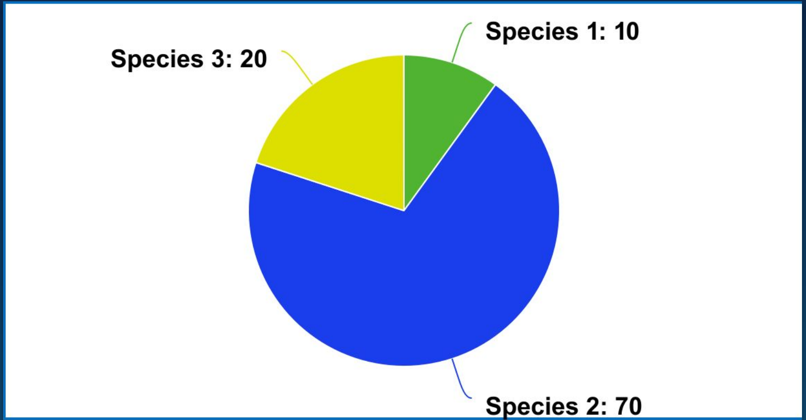


Metrics

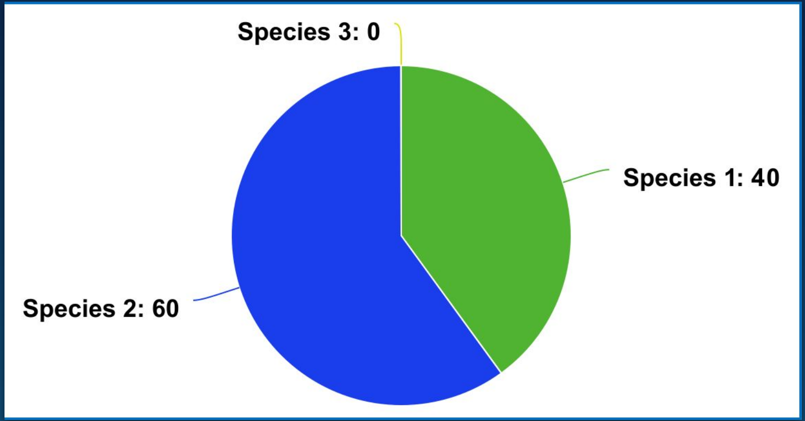
AI.FISH

Weighted Classification Error

Ground Truth



Predicted

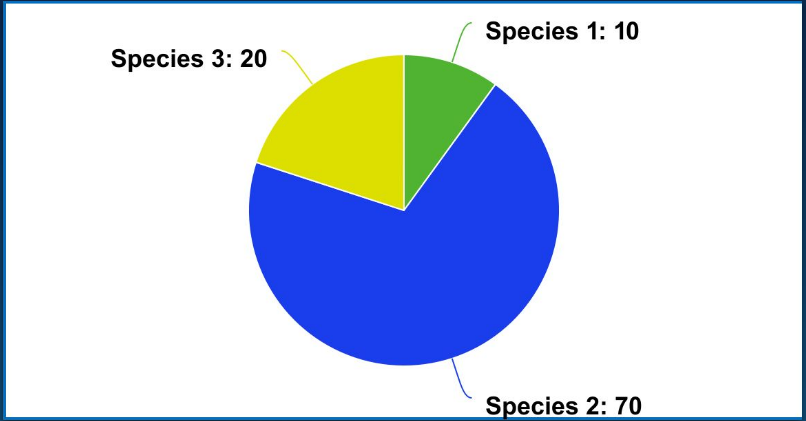


- **Mean absolute error:** (|40 - 10| + |60 - 70| + |0 - 20|) ÷ 3 = 20%

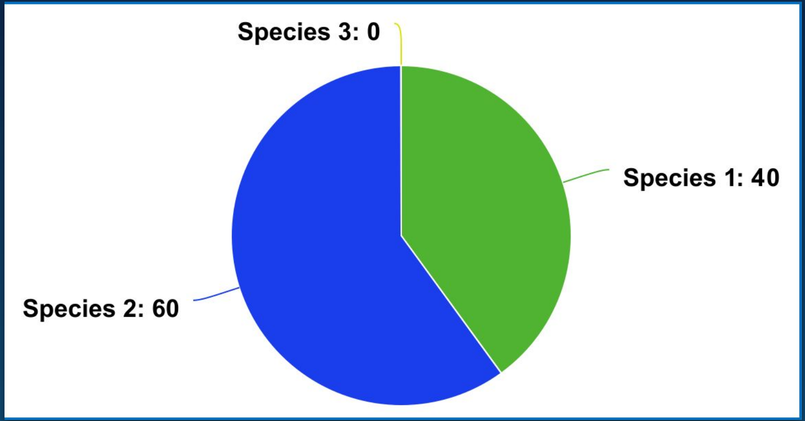
Metrics

Weighted Classification Error

Ground Truth



Predicted

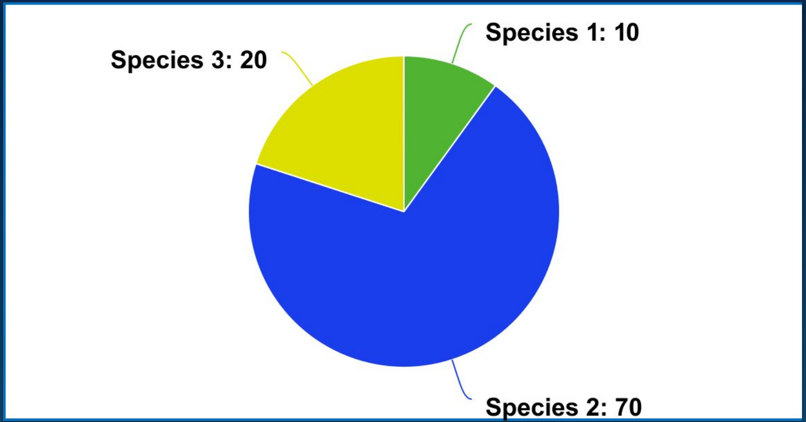


- **Mean absolute error:** (|40 - 10| + |60 - 70| + |0 - 20|) ÷ 3 = 20%
- But this gives equal weight to all classes, which might not be appropriate.
  - E.g. there are 5 classes total, but only 3 present; now divisor is 5, so error is artificially reduced

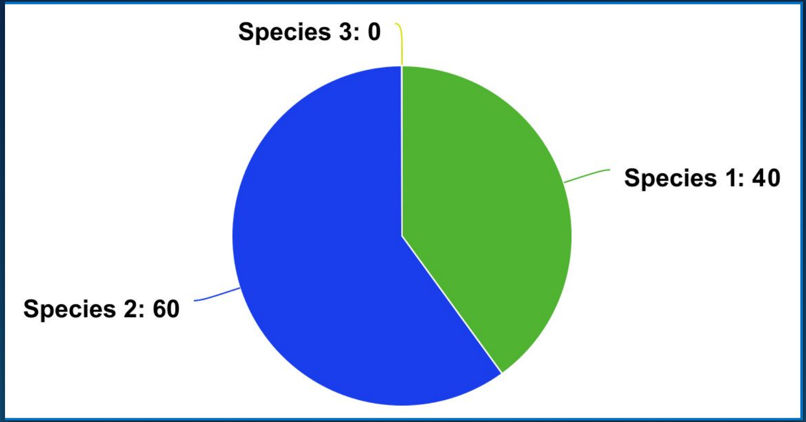
Metrics

Weighted Classification Error

Ground Truth



Predicted



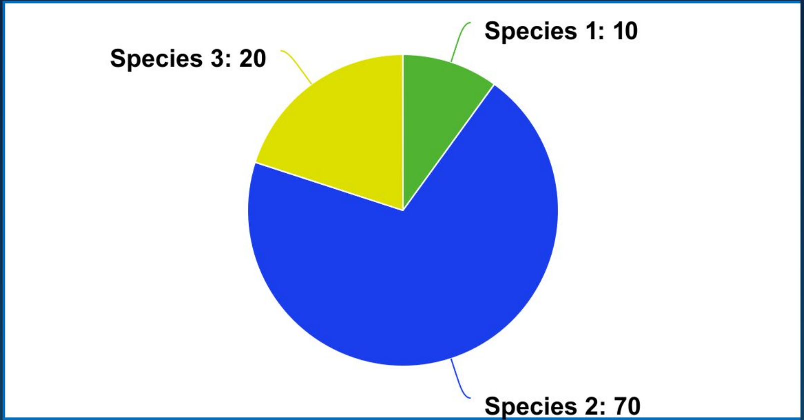
- **Weight by ground truth:** 0.1*|40 - 10| + 0.7*|60 - 70| + 0.2*|0 - 20| = 14%

Metrics

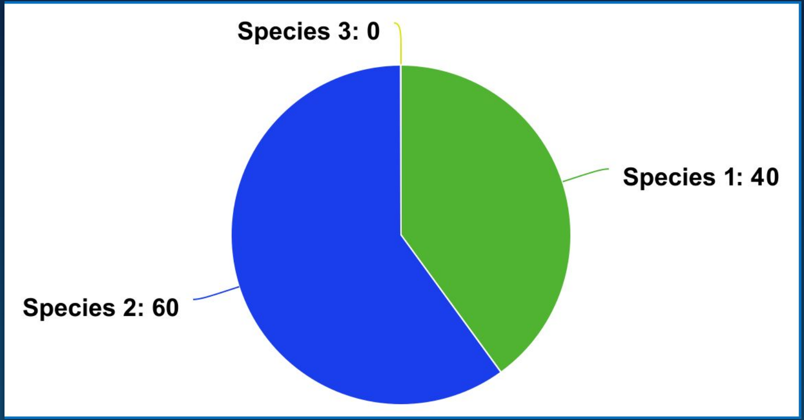
AI.FISH

Weighted Classification Error

Ground Truth



Predicted

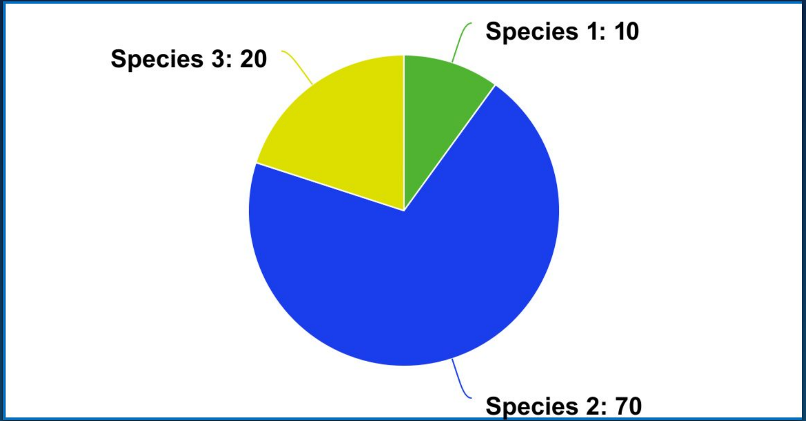


- **Weight by ground truth:** 0.1*|40 - 10| + 0.7*|60 - 70| + 0.2*|0 - 20| = 14%
- But what if a ground truth class is not present, but you predict it? Error for that class would be 0.

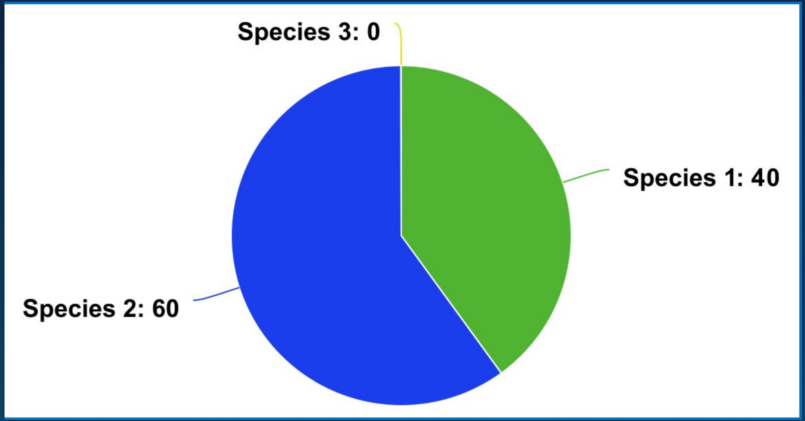
Metrics

Weighted Classification Error

Ground Truth



Predicted

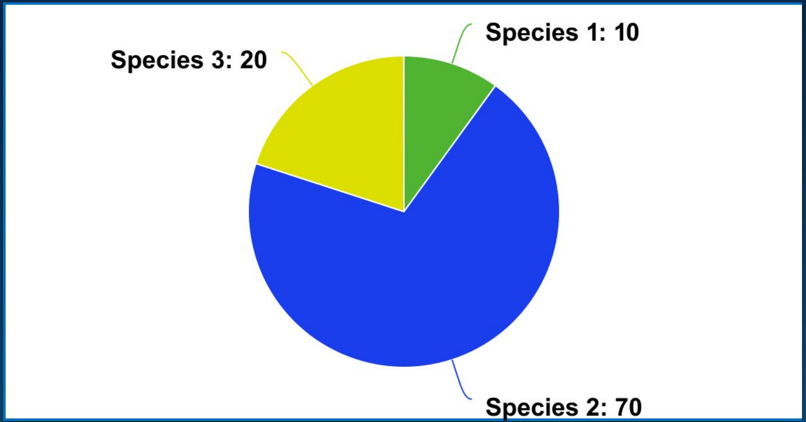


- **Weight by avg of GT + predicted:** 0.25*|40 - 10| + 0.65*|60 - 70| + 0.1*|0 - 20| = 19.2%

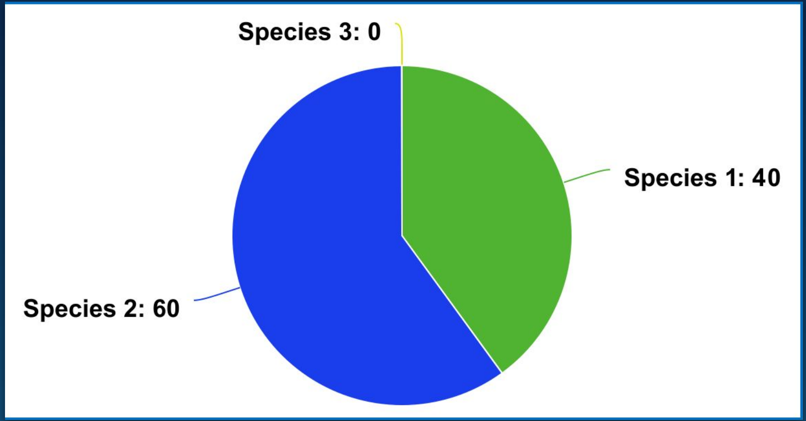
Metrics

Weighted Classification Error

Ground Truth



Predicted



- **Weight by avg of GT + predicted:** 0.25*|40 - 10| + 0.65*|60 - 70| + 0.1*|0 - 20| = 19.2%
- Bonus: Now the measure is symmetric, so neither needs to be considered the "ground truth": we can compare the discrepancy of human reviewers, for example.

Results

AI.FISH

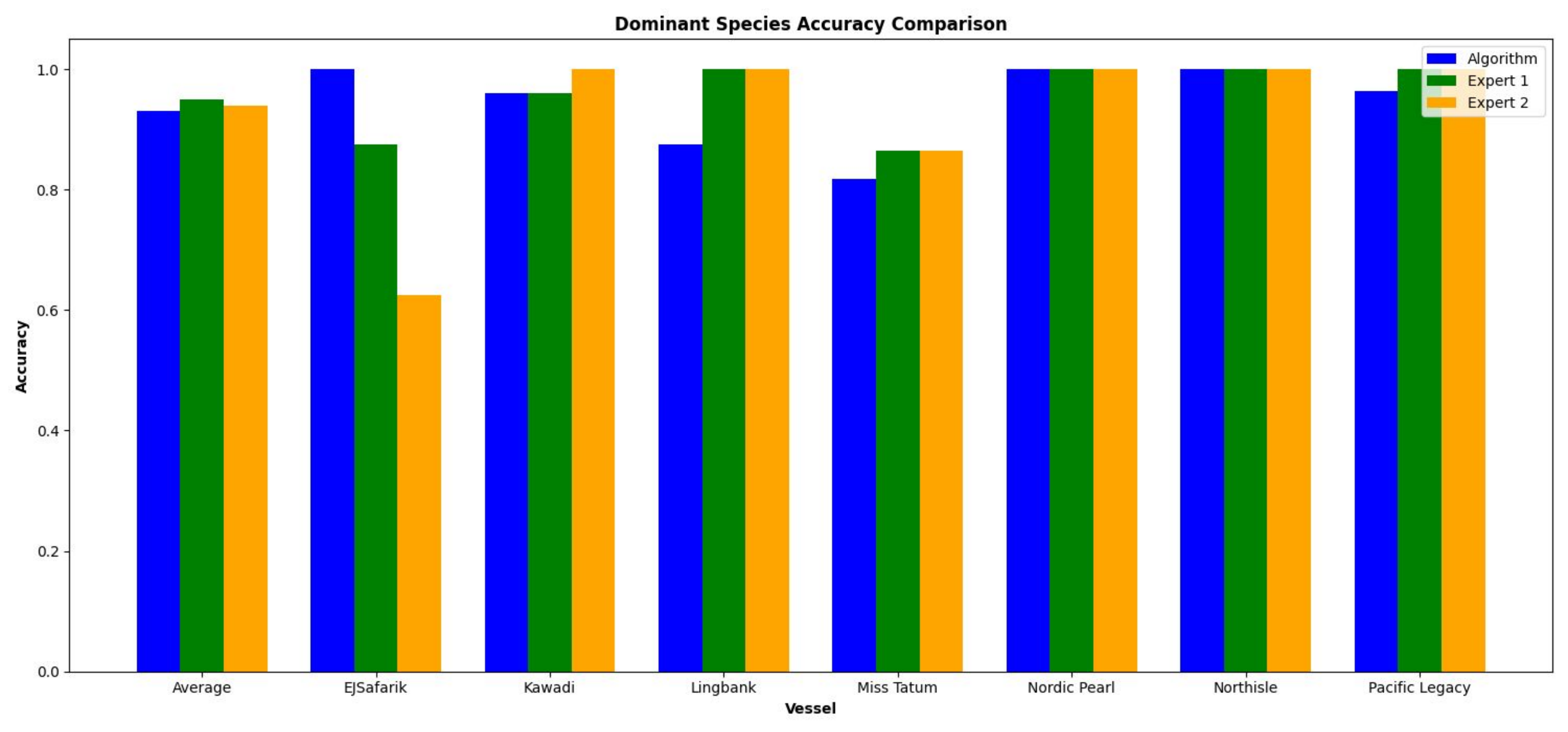
Evaluate:

- Model performance
- Human expert performance

Using:

- Test set of 100 held-out examples
  - Sampled from *tows not present in training data*

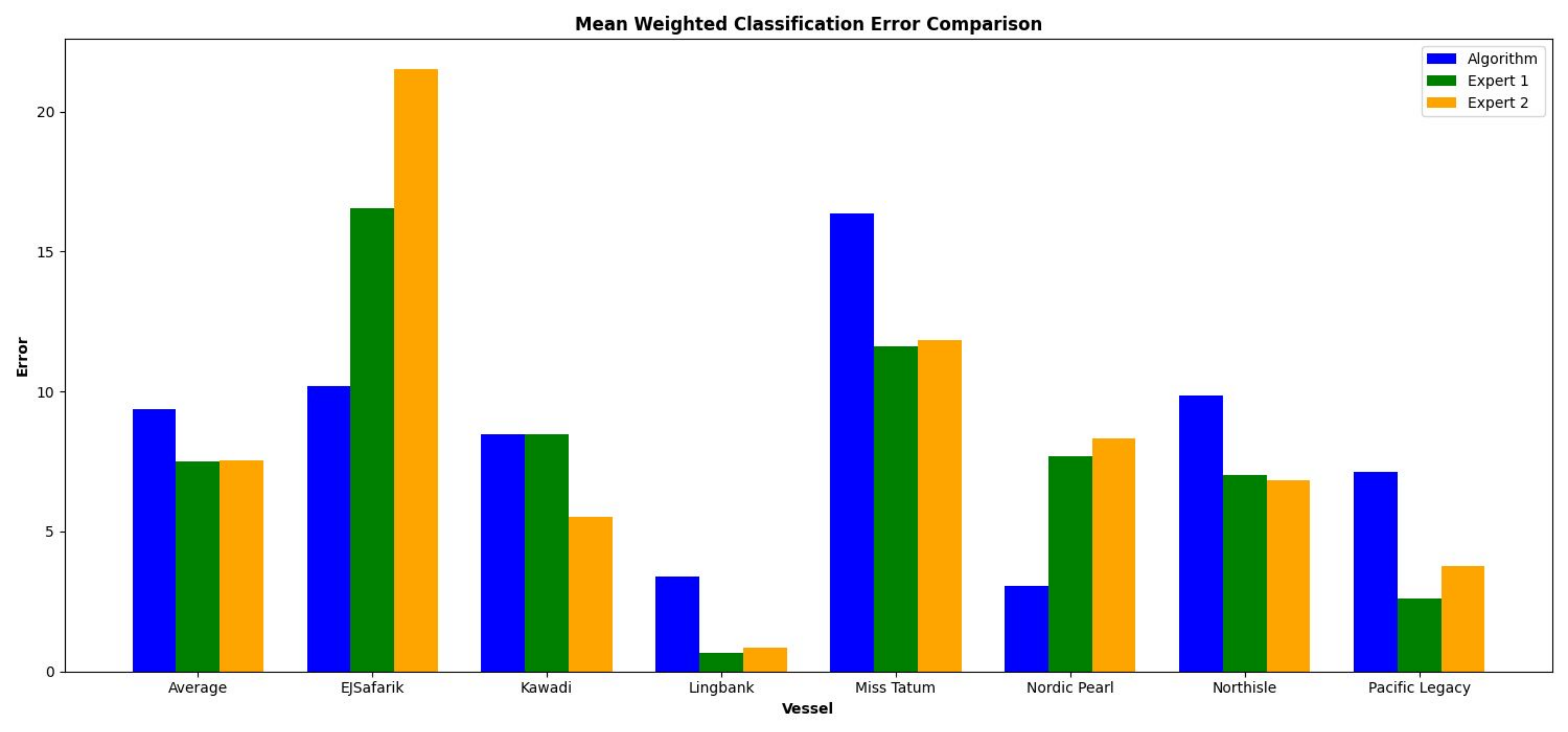
Results



Dominant Species Accuracy Comparison
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
Algorithm
Expert 1
Expert 2
Average
EJSafarik
Kawadi
Lingbank
Miss Tatum
Nordic Pearl
Northisle
Pacific Legacy
Vessel

Results

AI.FISH



Results

AI.FISH

Metric	Algorithm	Expert 1	Expert 2
Dominant Species Accuracy	94%	95%	94%
Mean Weighted Classification Error	9.4%	7.5%	7.5%

- Algorithm achieves human expert-level performance on dominant species classification
- Algorithm is within 2% of human expert performance when considering mean weighted classification error
- Demonstrates the **feasibility of our approach** for producing **accurate automated apportionment estimates**


Next Steps
Proof of Value Proposal Outline

Access to Data

AI.FISH

- Work with Archipelago to obtain access to at least 20 boats to participate in the Proof of Value phase

- NOTE: Data will not be streamed from the vessels but will be analyzed as part of the typical Program flow alongside all standard review that is normally conducted

POV Analysis

- Reviewers will follow their usual routine
- During review clips from relevant tows on participating boats will be extracted from footage loaded into FishVue Interpret and loaded into the POC solution
- The reviewer will select at least 3 - 5 images from each tow for catch apportionment analysis
- The reviewer will review catch apportionment results and then utilize these results to inform catch estimates.
- A sample of tows will have full manual review performed so that we can assess the impact on accuracy of catch estimates

Timeline

AI.FISH

- The POV phase will be conducted in 2 iterations to facilitate an opportunity for incorporation of feedback on either algorithm or user interface performance
- The first iteration will run from **Oct 16 - Dec 29, 2023**
- The second iteration will run from **Jan 15 - Feb 29, 2024**
- Time prior to Oct 16, 2023 will be used to line up permissions with vessel owners and to finalize the software implementation to operate at scale.
- Aim to present results to DFO team by **March 15, 2024**.

KPIs for POV Phase

AI.FISH

- Number of catch apportionment results rejected by reviewers (single images)
- Number of total catch apportionment results rejected by reviewers (tows)
- Accuracy of catch apportionment (deviation compared with manual review)
- Delta in results between first and second iterations
- Change in administrative user feedback between iterations if any UI/UX adjustments are made