fluke_net

By Connor Barlow, Caelan Booker, Adicus Finkbeiner, Logan Pashby, Dylan Thompson

The Problem

- Whale populations devastated by whaling
  - Photo identification is primary means of tracking population
- Differentiate individual whales using only a picture of its tail
  - Shape
  - Patterns/Colors
  - Unique Markings
- Data provided by Happywhale
  - Track whales worldwide
  - Massive database contributed to by scientists and citizens alike

Why Deep Learning?

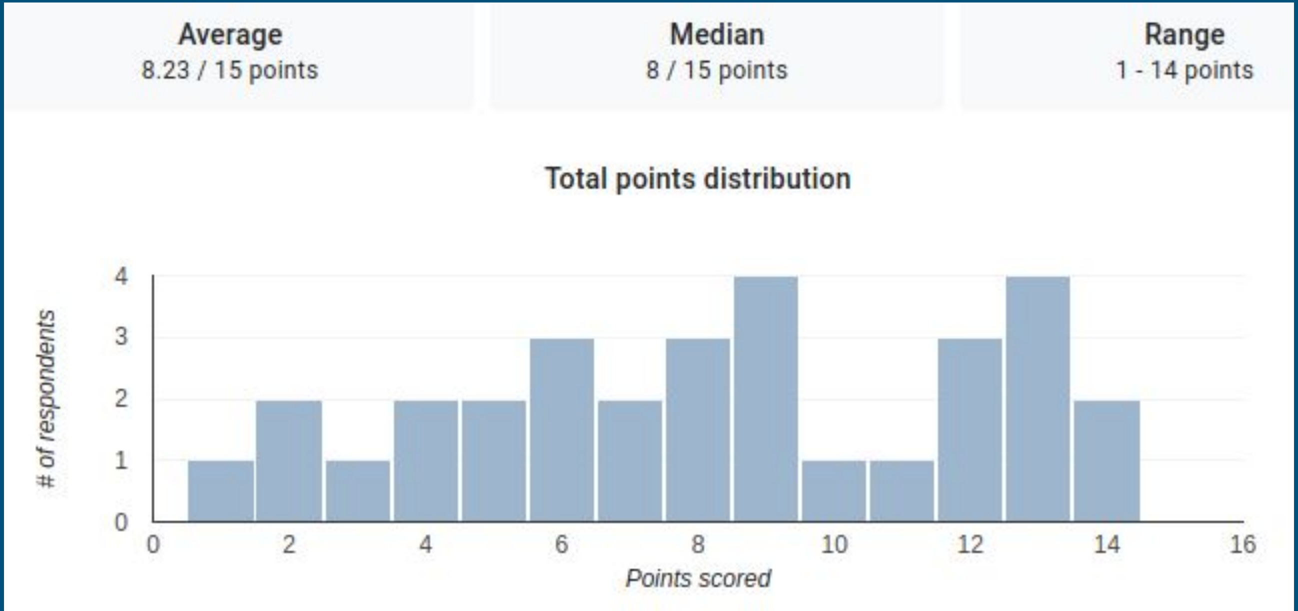
- Past 40 years it’s been manually recorded
- Massive amounts of data
    - Don’t have the manpower to process the data
    - Constantly growing dataset
- Large scale image processing

Philip Robinson, Staff Software Engineer at HappyWhale and WWU alumni

Sample Quiz Results

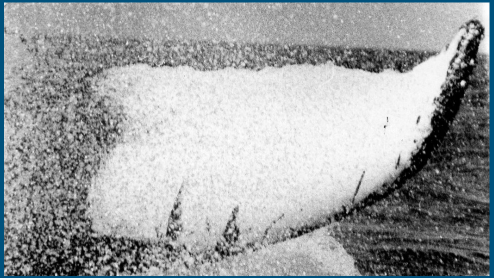
Examples from our quiz

- Surveyed 31 people
- Provided a “training set” of 15 whales broken into 5 classes and quizzed them on a 15 whale “test set”



Dataset Examples

- Varying degrees of quality
- Even with “good” images, patterns can be hard to spot
- Large variance in what makes an image “poor” quality



Dataset Structure

- 2 distinct categories of classes
  - unique_whale classes
  - new_whale class (singular)

Training Dataset (5K classes/25K examples)

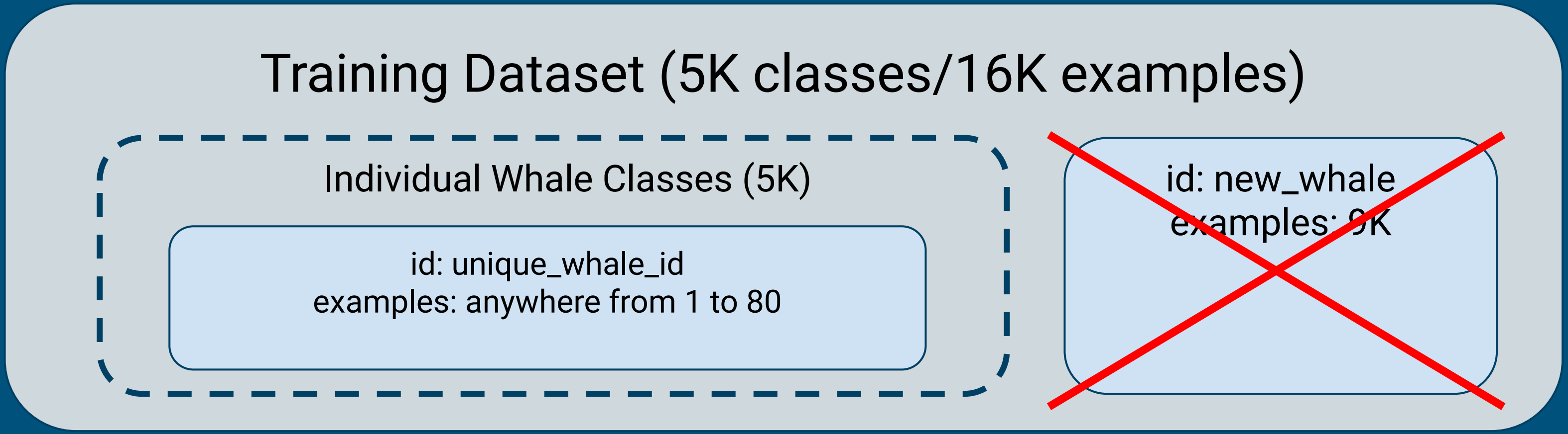
Individual Whale Classes (5K)

id: unique_whale_id
examples: anywhere from 1 to 80

id: new_whale
examples: 9K

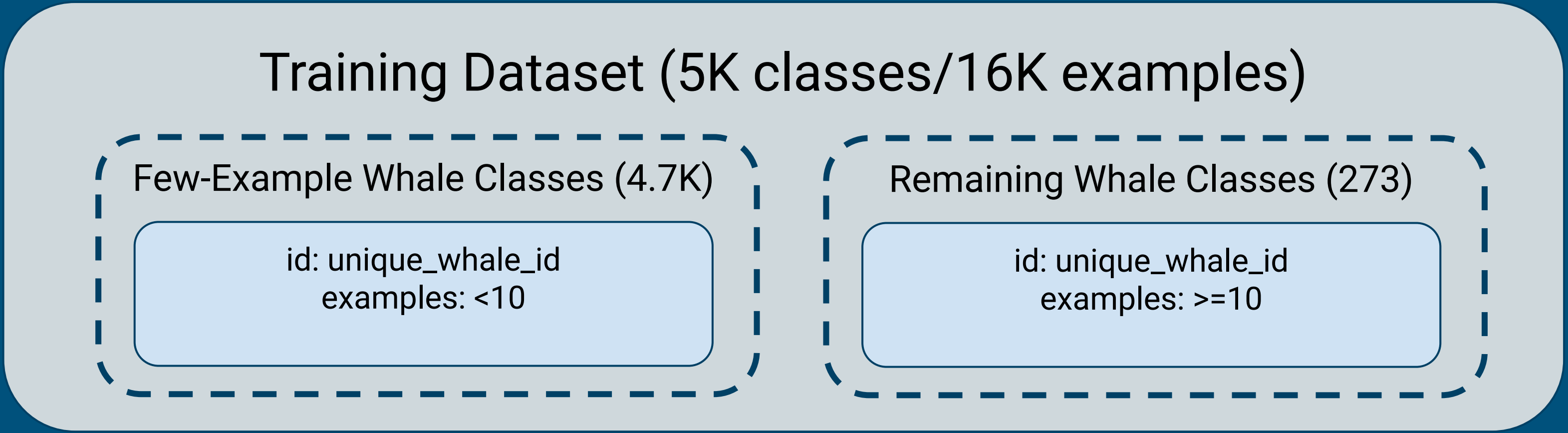
Dataset Problem #1: “new_whale” class

- Denotes “this whale is not in any other labeled class”
- Takes up 9k of our 25k total examples
- Unusable as a target for classification



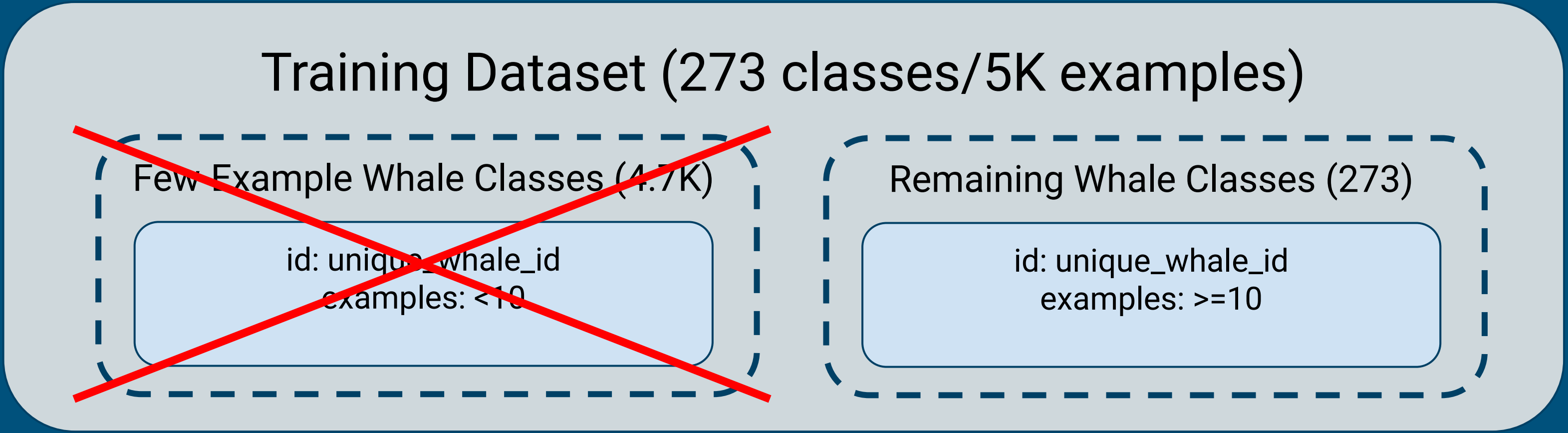
Dataset Problem #2: Huge Class Imbalance

- Spread of examples-per-class is huge
  - ~2K classes contain only one example
  - Examples/class goes up to 80



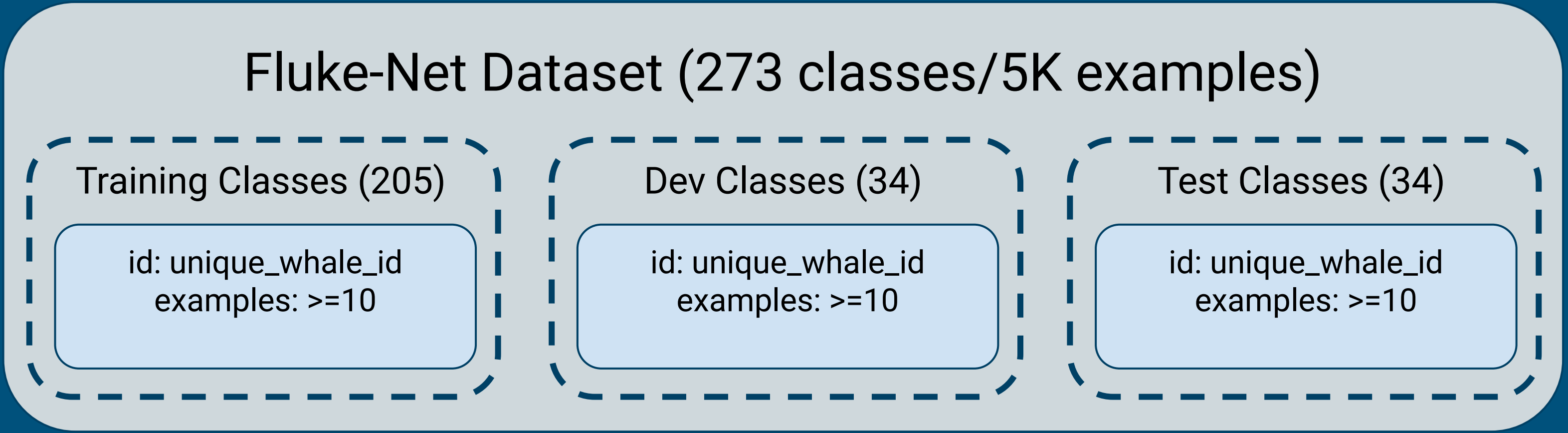
Dataset Problem #2: Huge Class Imbalance

- Spread of examples-per-class is huge
  - ~2K classes contain only one example
  - Examples/class goes up to 80
- Complexity of approach reduced, chance of overfitting increased



Resulting Dataset and Split

- Limitations in data evaluation forced further splitting
- Diverges from learning goal only in no “new_whale”
- Allows more typical FSL approaches to be applied

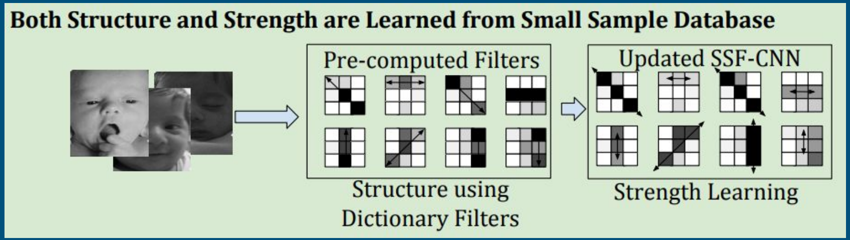


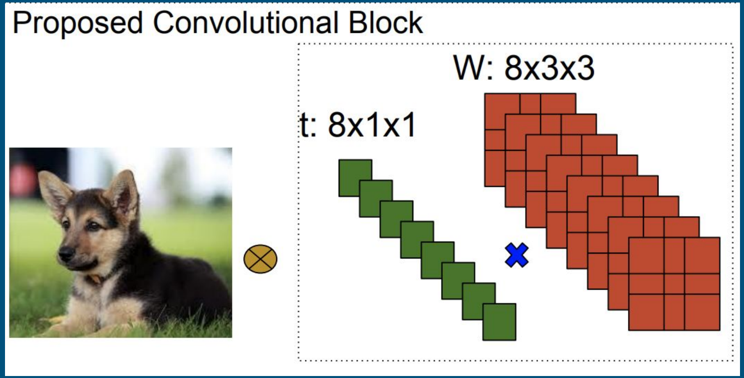
Our Approach

- Problem: few examples per class; “dirty data”
- Solution: data augmentation
  - Rotation
  - Crop
  - Horizontal reflection
  - Grayscale
  - Color jitter

Our Approach

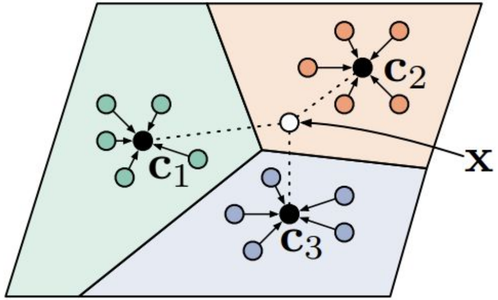
- Problem: large number of parameters relative to size of dataset
- Solution: “Learning Structure and Strength of CNN Filters for Small Sample Size Training” (Keshari et al.)





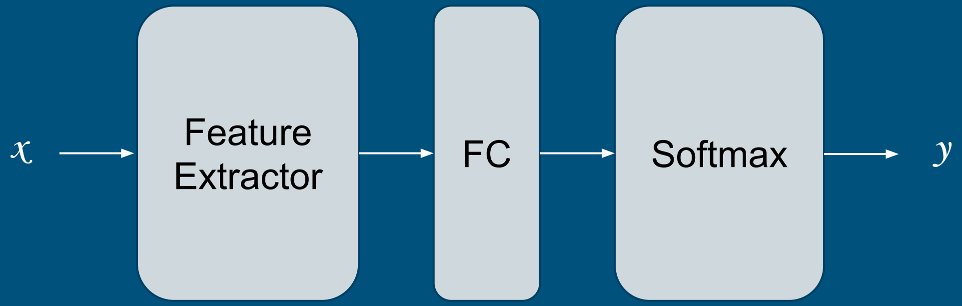
Our Approach

- Problem: few examples per class
- Solution: “Prototypical Networks for Few-shot Learning” (Snell et al.)

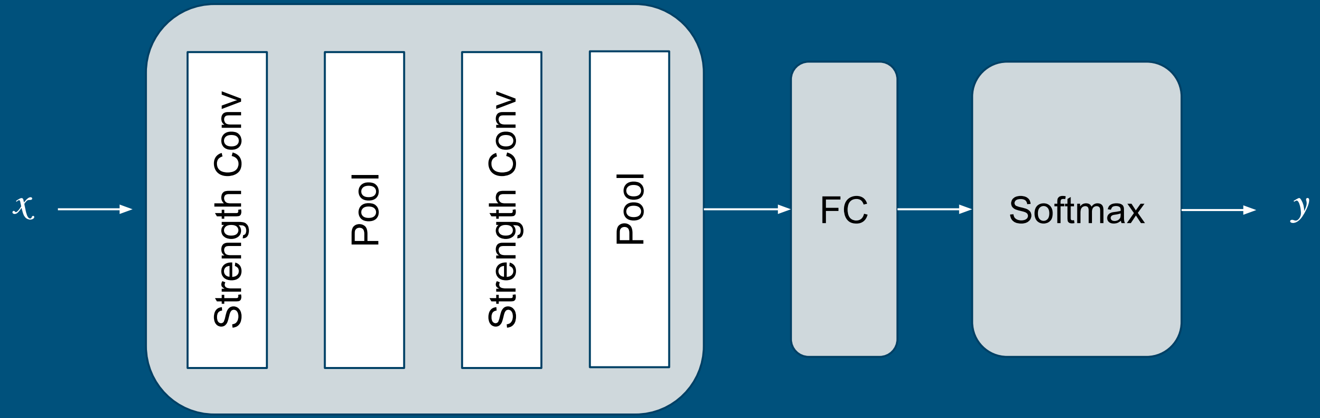


(a) Few-shot

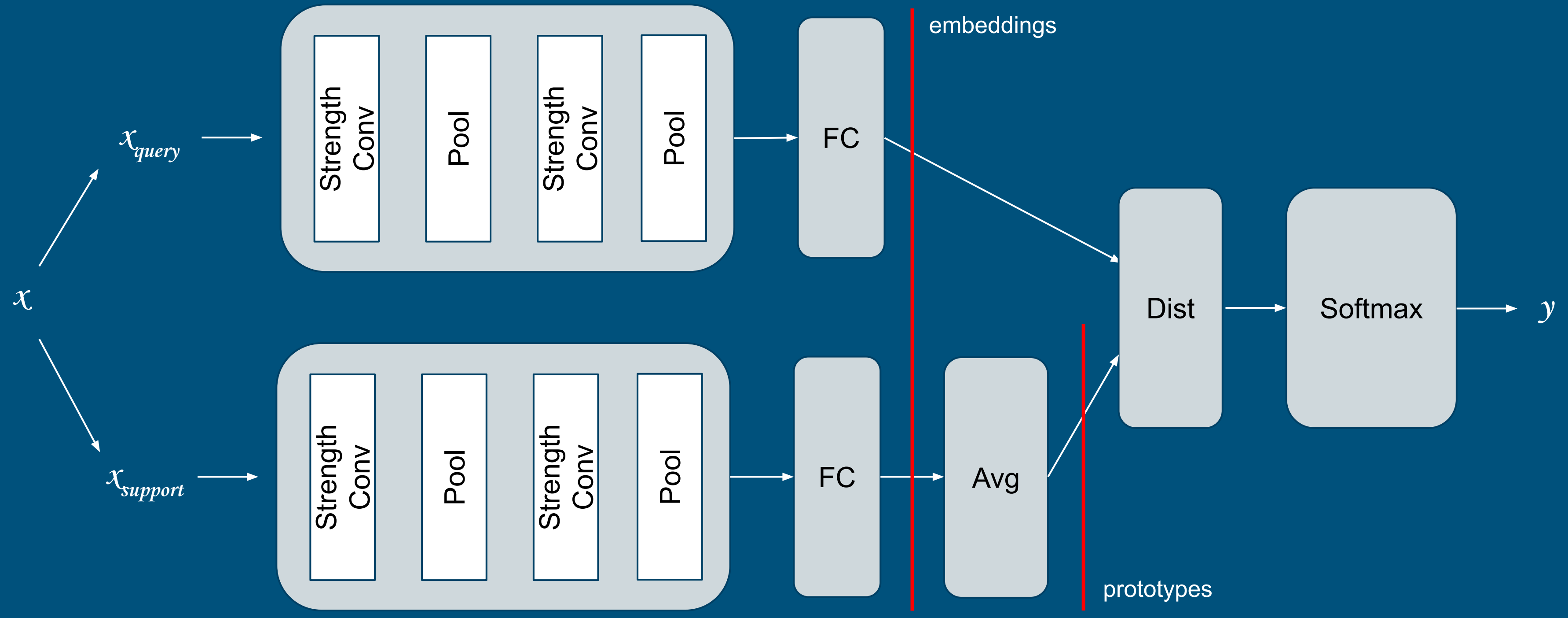
Model Architecture - Classifier Overview



Model Architecture - Feature Extraction



Model Architecture - Complete



Results – Baselines

Strategy	Test Set Accuracy
Random: Stratified	0.0053
Random: Frequency	0.0976
Random: Prior	0.0976
Random: Uniform	0.0053
K-Nearest-Neighbors (K=1, 8 examples per test class)	0.0189
Vanilla CNN	0.0224
Humans	0.55

Results - Hyperparameter Sweep

Learning Rate Sweep	
	mAP@1
1e-4	0.11
1e-3	0.15
1e-2	**0.34**
1e-1	0.26
1e0	0.24

Minibatch Size Sweep	
	mAP@1
4	**0.28**
8	0.27
16	0.23
32	0.20
64	–

Pretrained / Strength Filters			
	Pretrained		
Strength	mAP@1	Yes	**No**
	Yes	0.22	–
	No	0.25	**0.28**

# of Epochs Sweep	
	mAP@1
40	0.28
80	0.31
120	**0.38**
160	0.38

Results – Test Set Accuracy

- mAP@1: 0.33
- mAP@5: 0.44
- mAP@20: 0.46

Not bad compared to our human mAP@1 of 0.55

Potential Next Steps

Siamese Network

- Another popular few-shot technique
- Two identical neural networks on two separate images
- Compares the embedding of each image to each other via the distance
- Parameters trained to create encodings of like images to be small, and vice versa for unlike images

Semi-Supervised Few-Shot Learning

(using unlabeled examples in each episode)

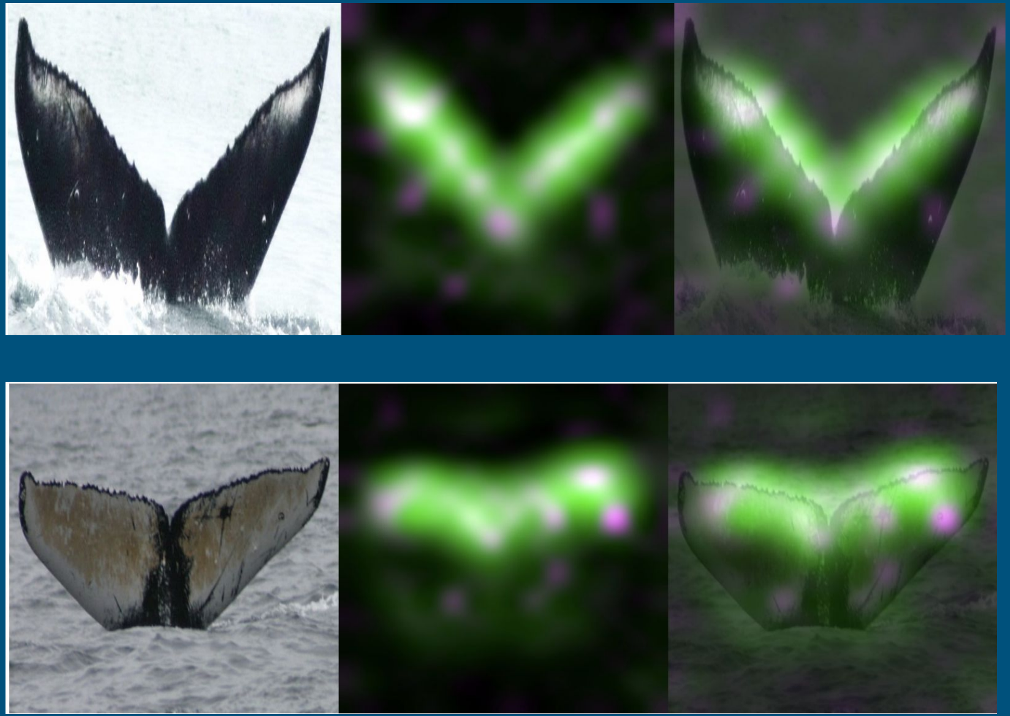
- Iterate through episodes of support, query, and unlabeled
- Refine via the unlabeled images with soft k-means
- Unlabeled can include 'distractor' images that should be masked as they get further away from the existing prototypes
  - Soft masks
  - Feed small NN stats

Source: Metalearning for Semi-Supervised Few-Shot Classification

Siamese network could add an attention span for the model

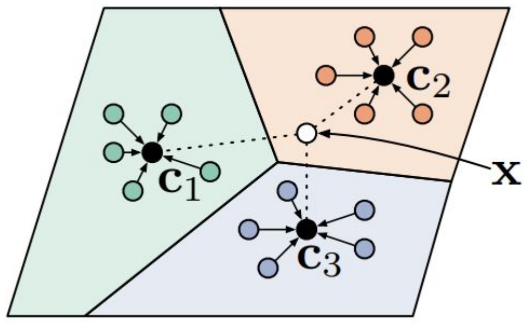
Key:

- Pink - places where value of max appears often
- Green - places where Std is large
- White - Pink + Green

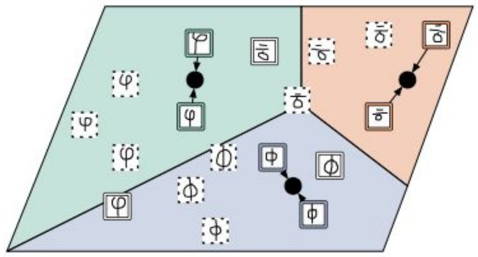


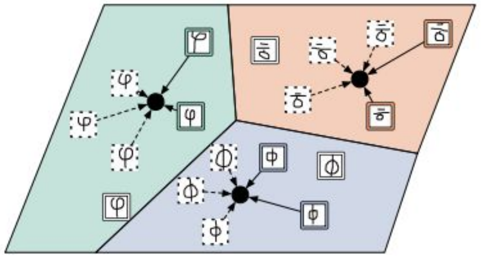
Source: Visualization of Siamese Network, Deeplearning.ai Video, Siamese Neural Networks for One-Shot Image Recognition

Semi-supervised would allow for adjustments of prototype embeddings



(a) Few-shot


Before Refinement


After Refinement

Our Model

Semi-Supervised

Source: [Prototypical Networks for Few-shot Learning" (Snell et al.)]

fin