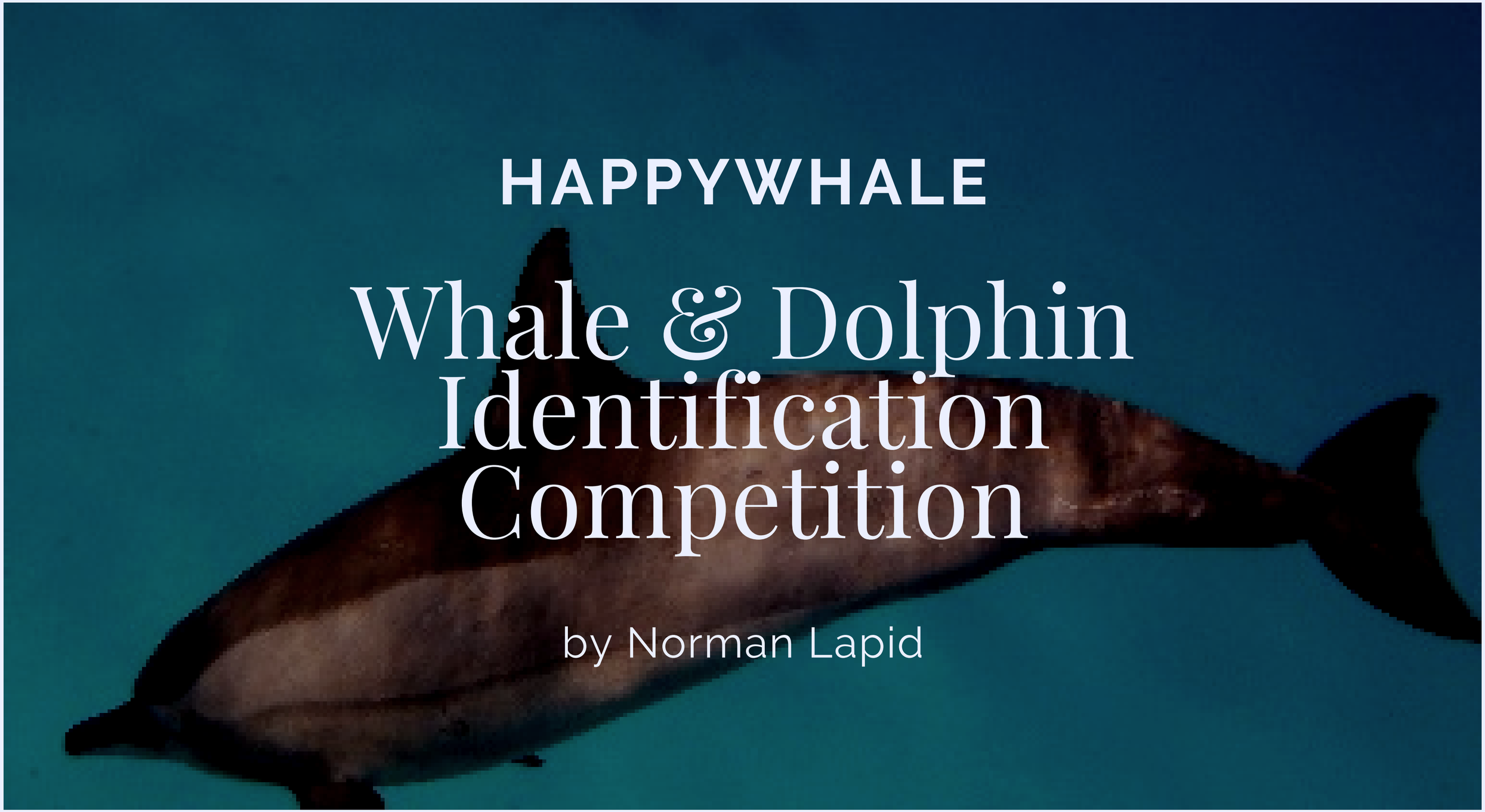

HAPPYWHALE
Whale & Dolphin Identification Competition
by Norman Lapid

Motivations

- Enjoy a cool marine science topic
- Explore computer vision architectures

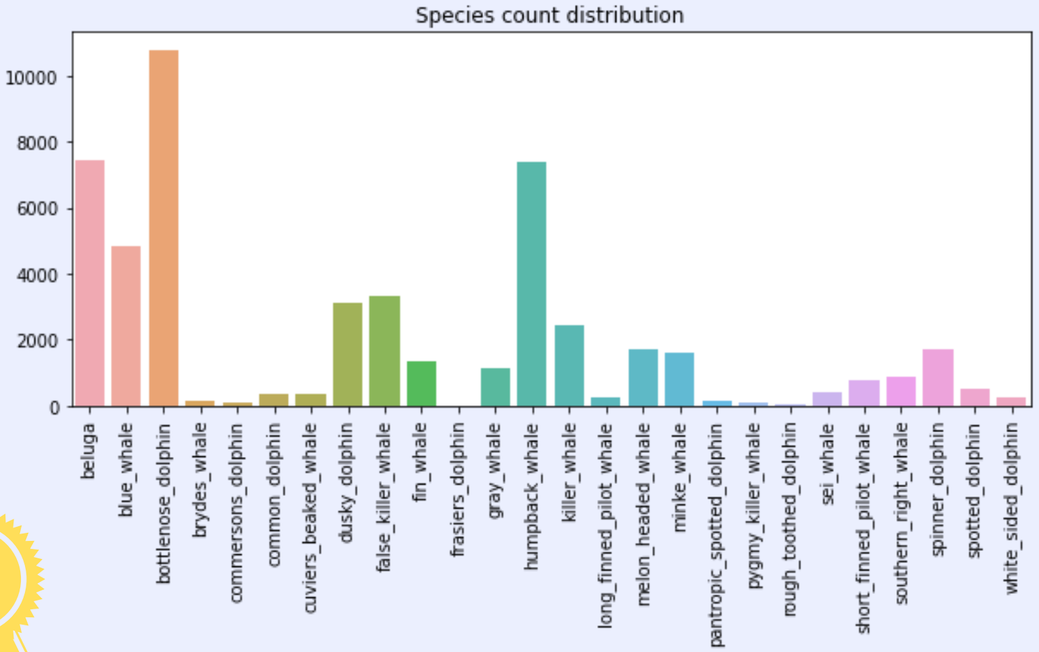
Experience joining a Kaggle competition ($25,000 in prizes)

Exploratory Data Analysis (51,033 images)

#1: Species Identification

Number of species: 26

PCC * 125: 13.9%

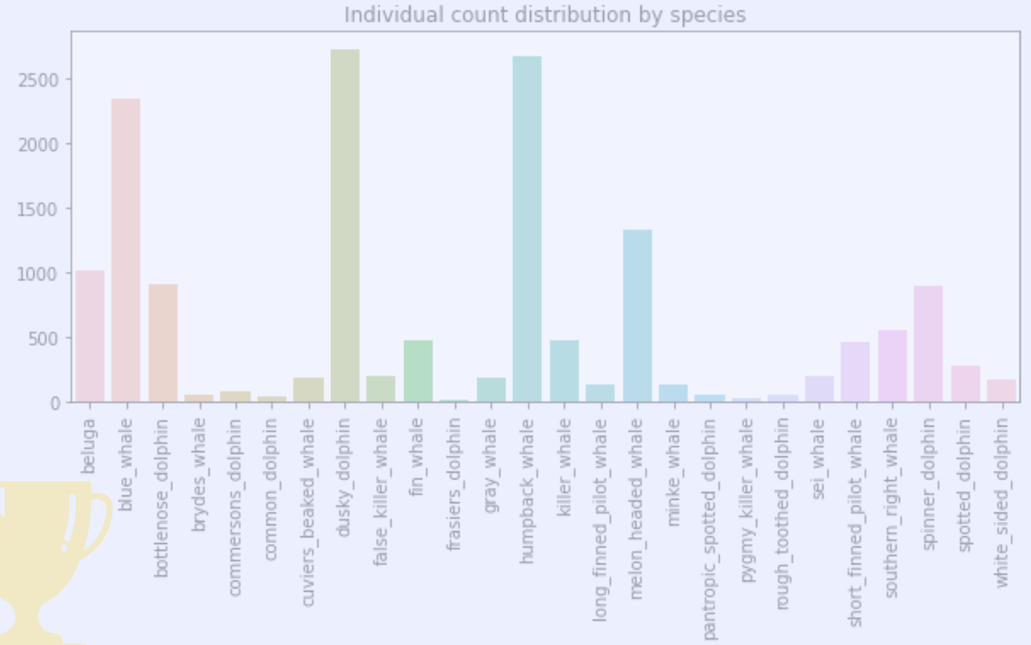


My focus

#2: Individual Identification

Number of individuals: 15,587

PCC * 125: 0.0719%



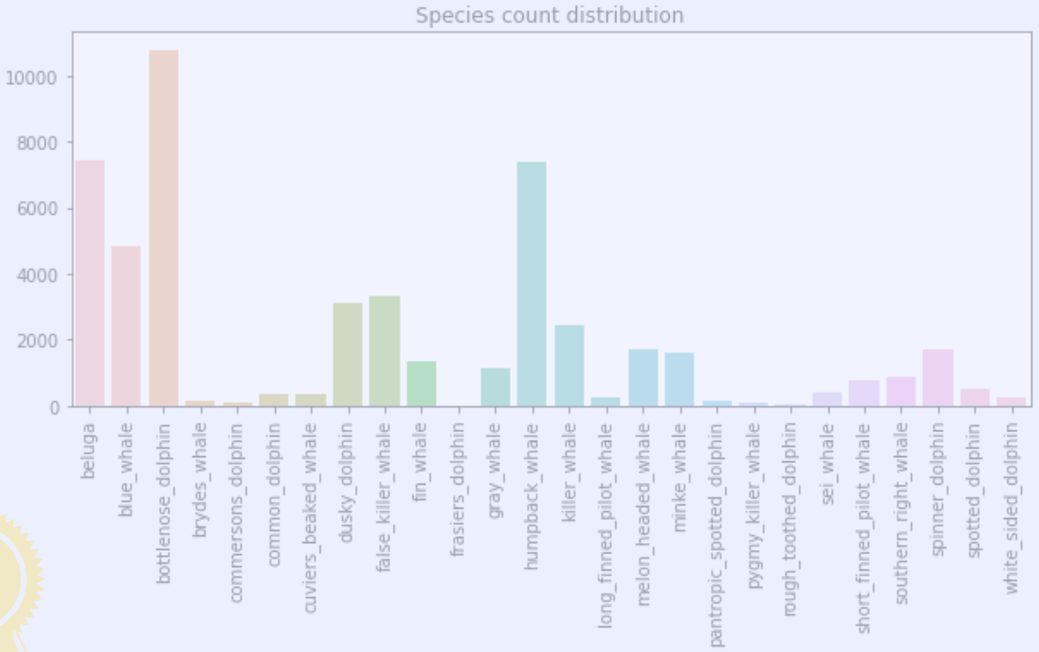
Contest focus

Exploratory Data Analysis (51,033 images)

#1: Species Identification

Number of species: 26

PCC * 125: 13.9%

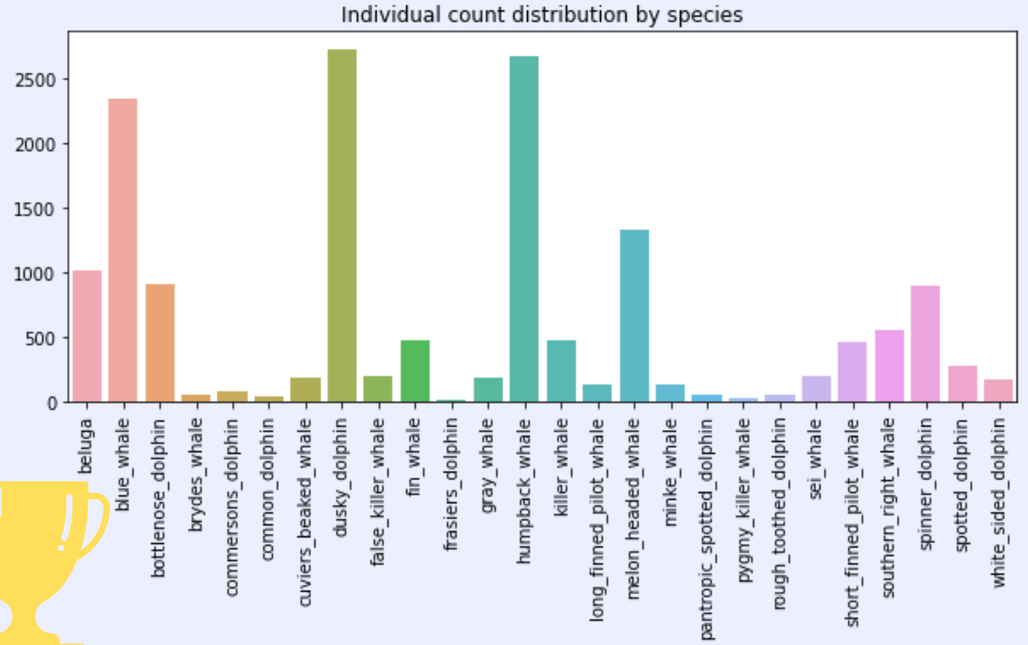


My focus

#2: Individual Identification

Number of individuals: 15,587

PCC * 125: 0.0719%



Contest focus

Base Layer

CNN or SOTA Hybrid Models

Considerations

- Competition discussion board
- Number of parameters + input resolution
- Pre-trained model availability (Keras/Github)

2014

VGG16

CNN with increased depth and small (3x3) convolution filters

2016

Inception V3

Scaled-up CNN aiming to be as efficient as possible by suitably factorized convolutions and aggressive regularization

2019

EfficientNet B0

Conditionally parameterized convolutions, which learn specialized convolutional kernels for each sample for efficient inference

May 2021

LeViT192

A hybrid using vision transformers, employing activation maps with decreasing resolutions and attention bias for fast inference

Sep 2021

CoAtNet0

Marries CNN and Transformers to improve both generalization and model capacity

Top Layer

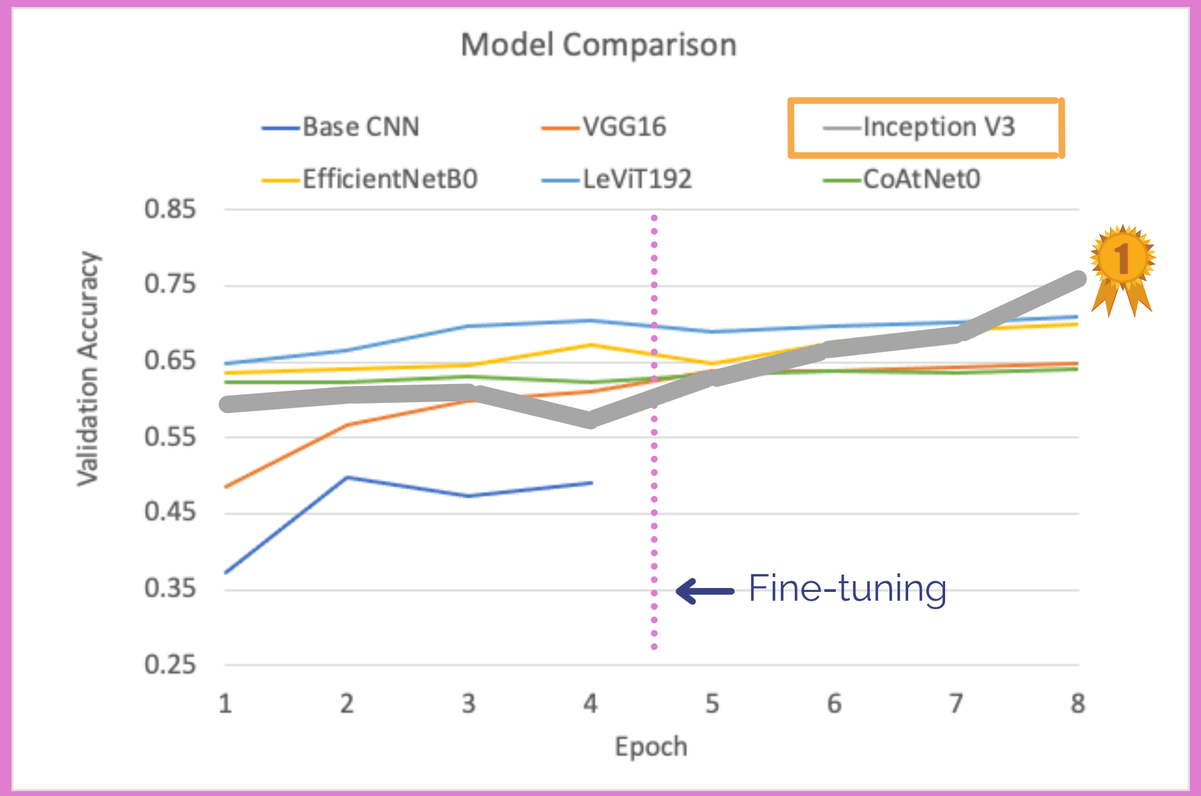
Fully-connected layers

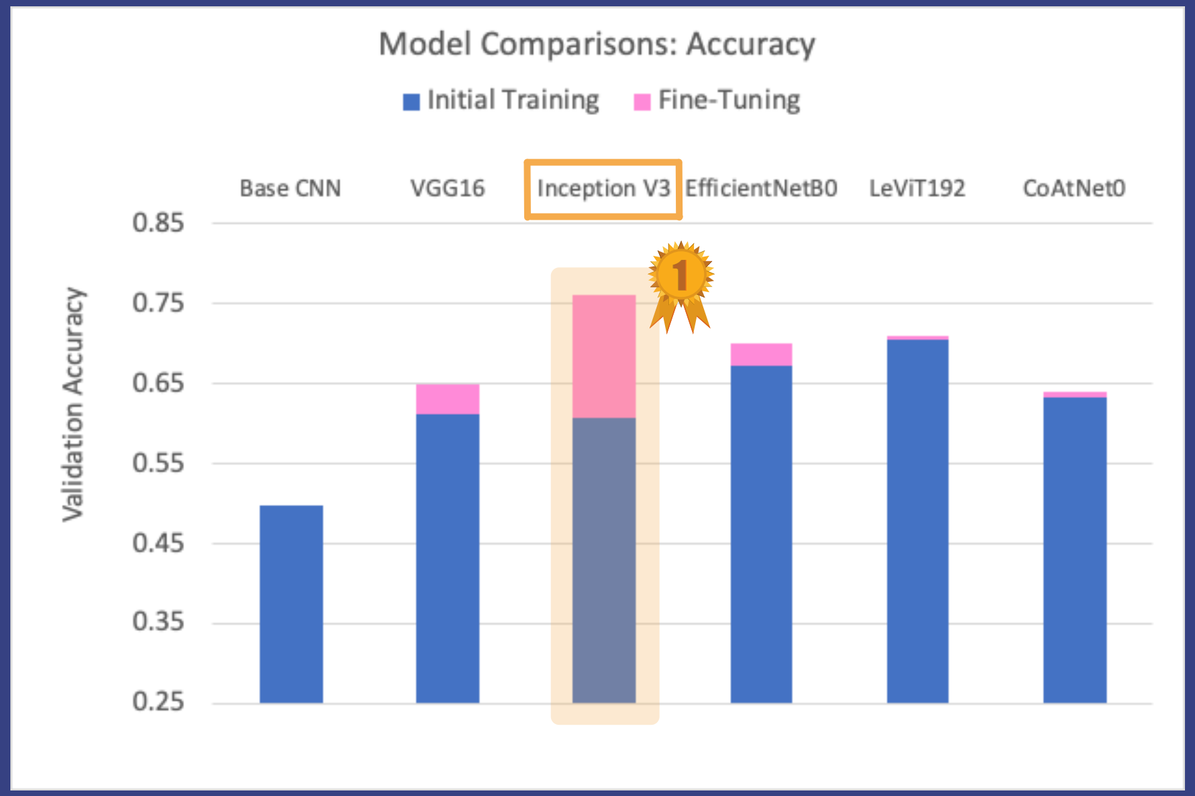
Dense layer: 256 nodes

Dropout: 0.6

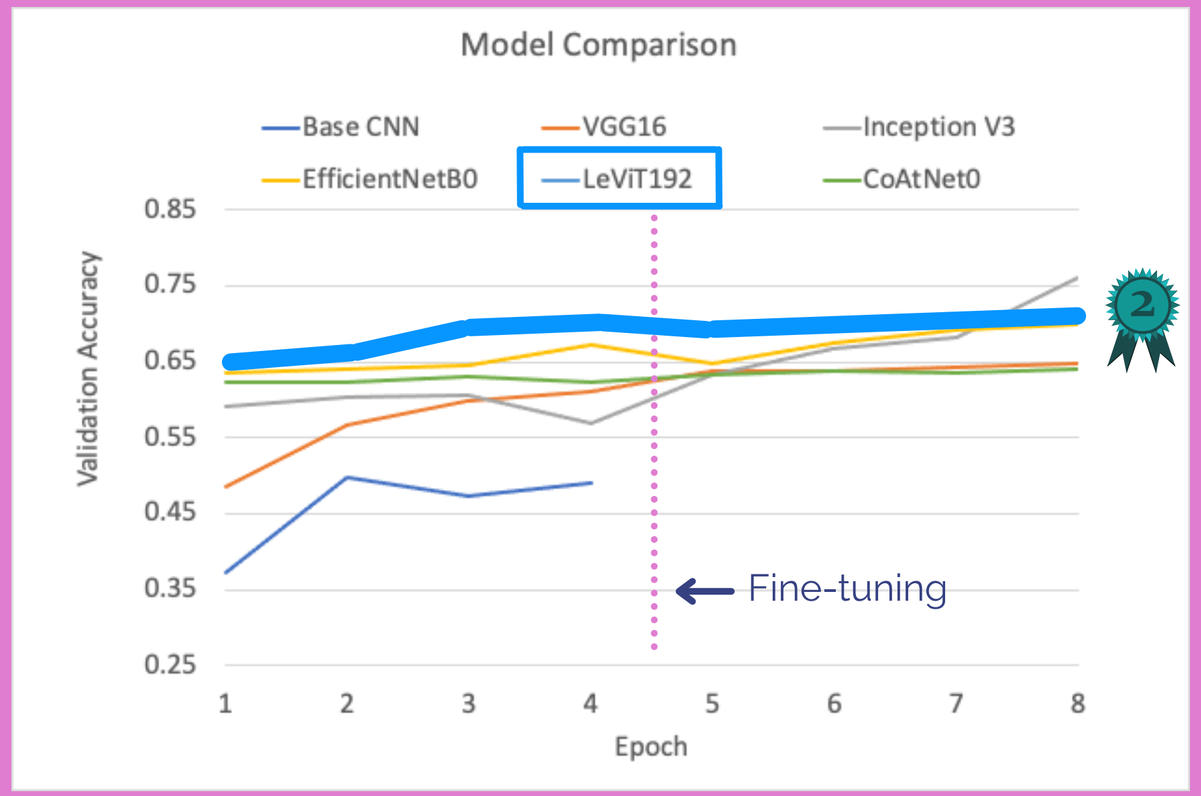
Output: Softmax activation

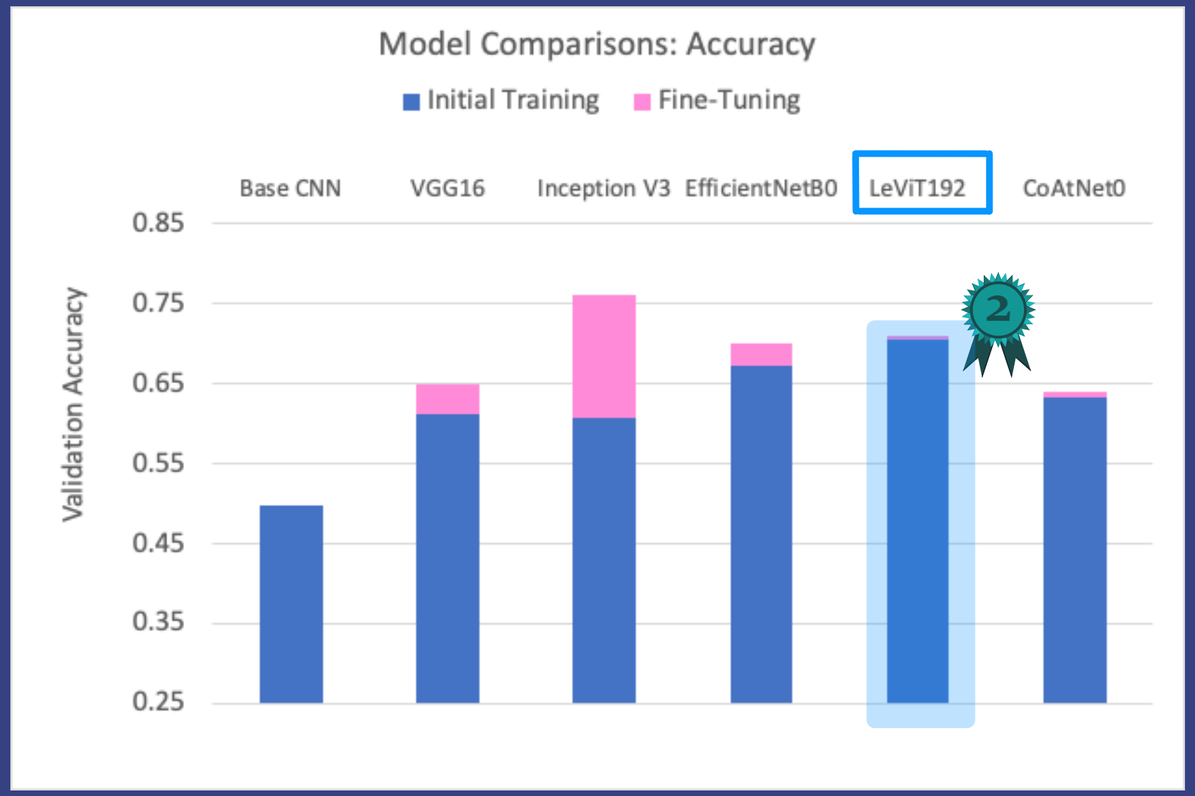
Elimination Round



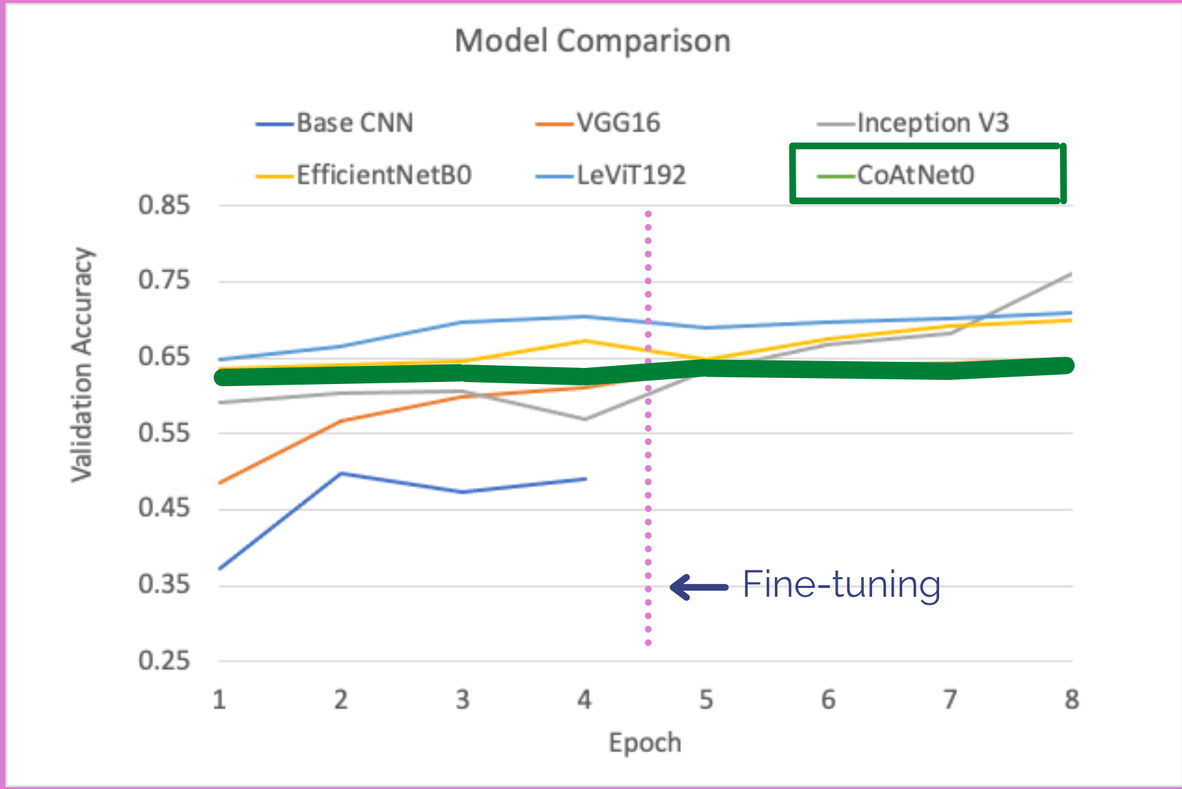


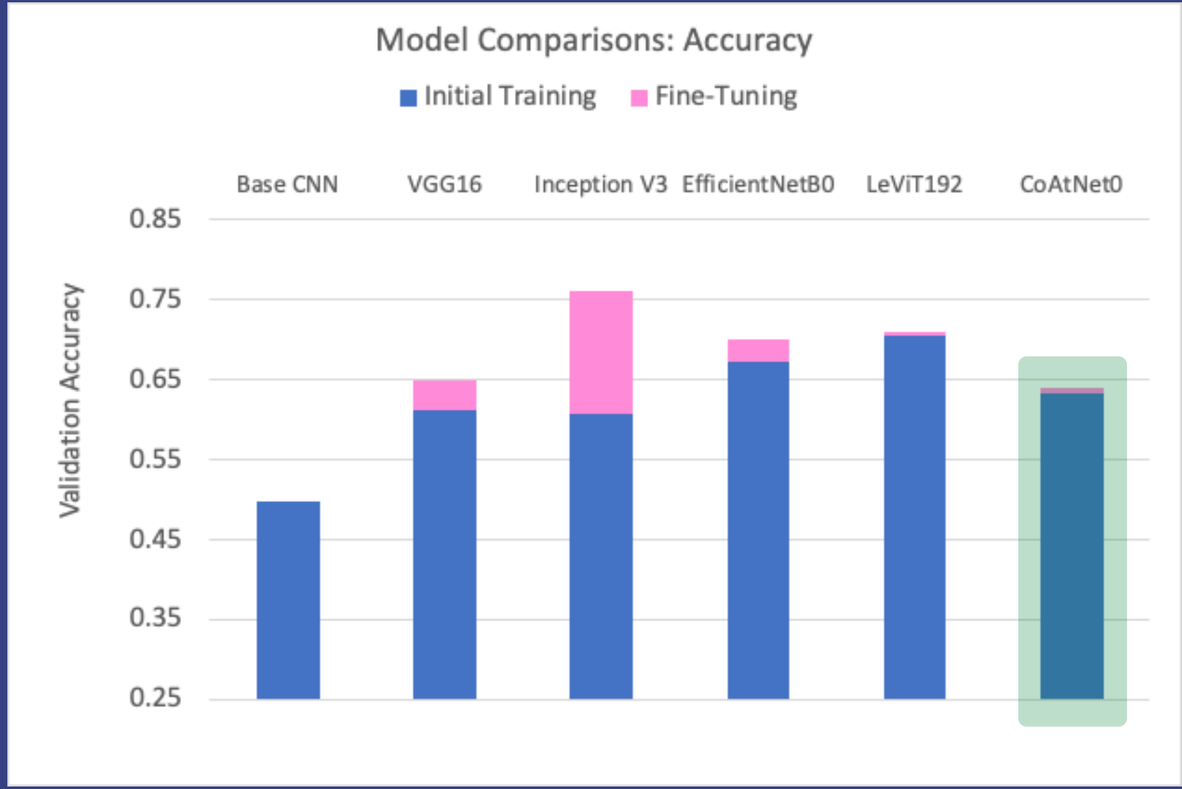
Elimination Round



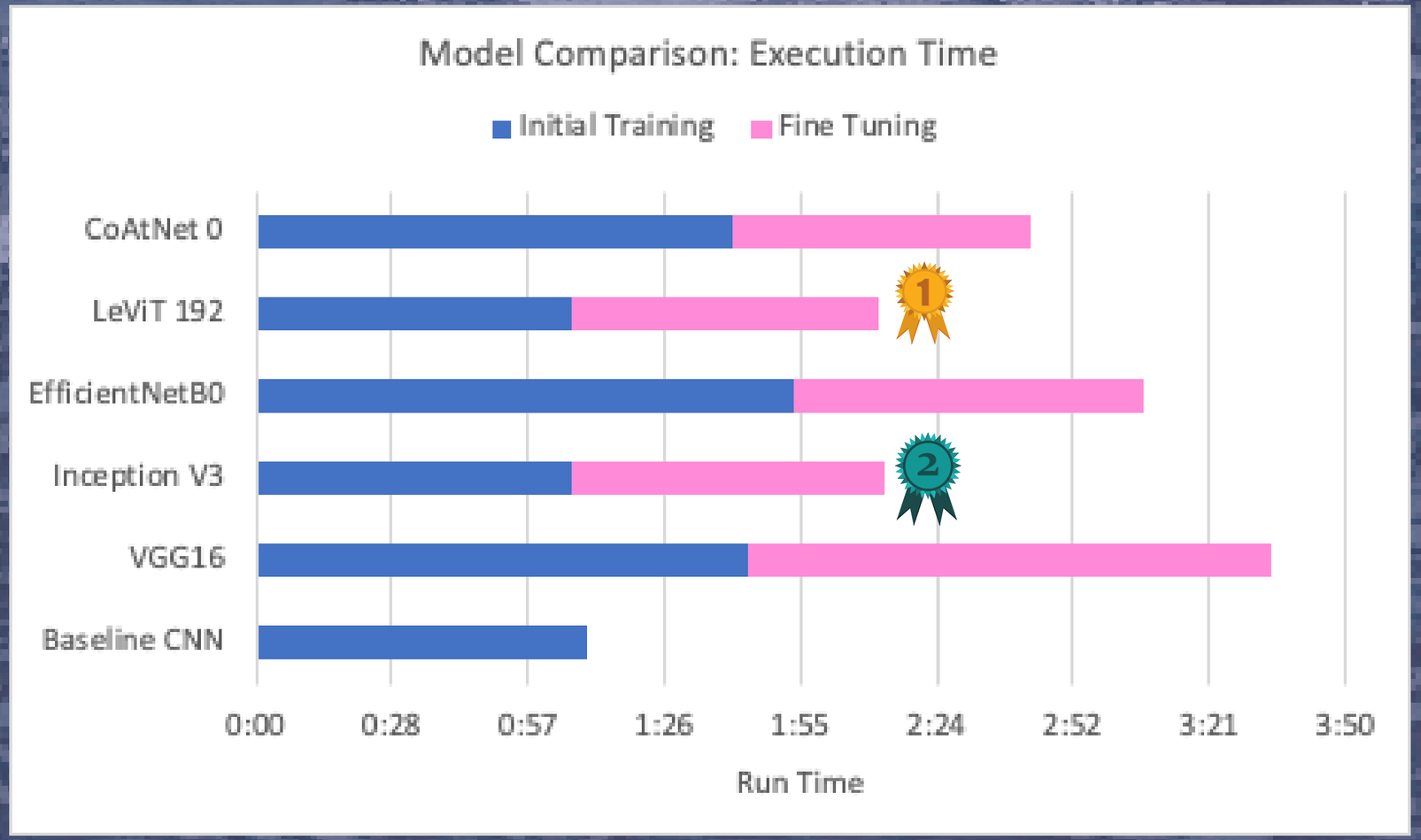


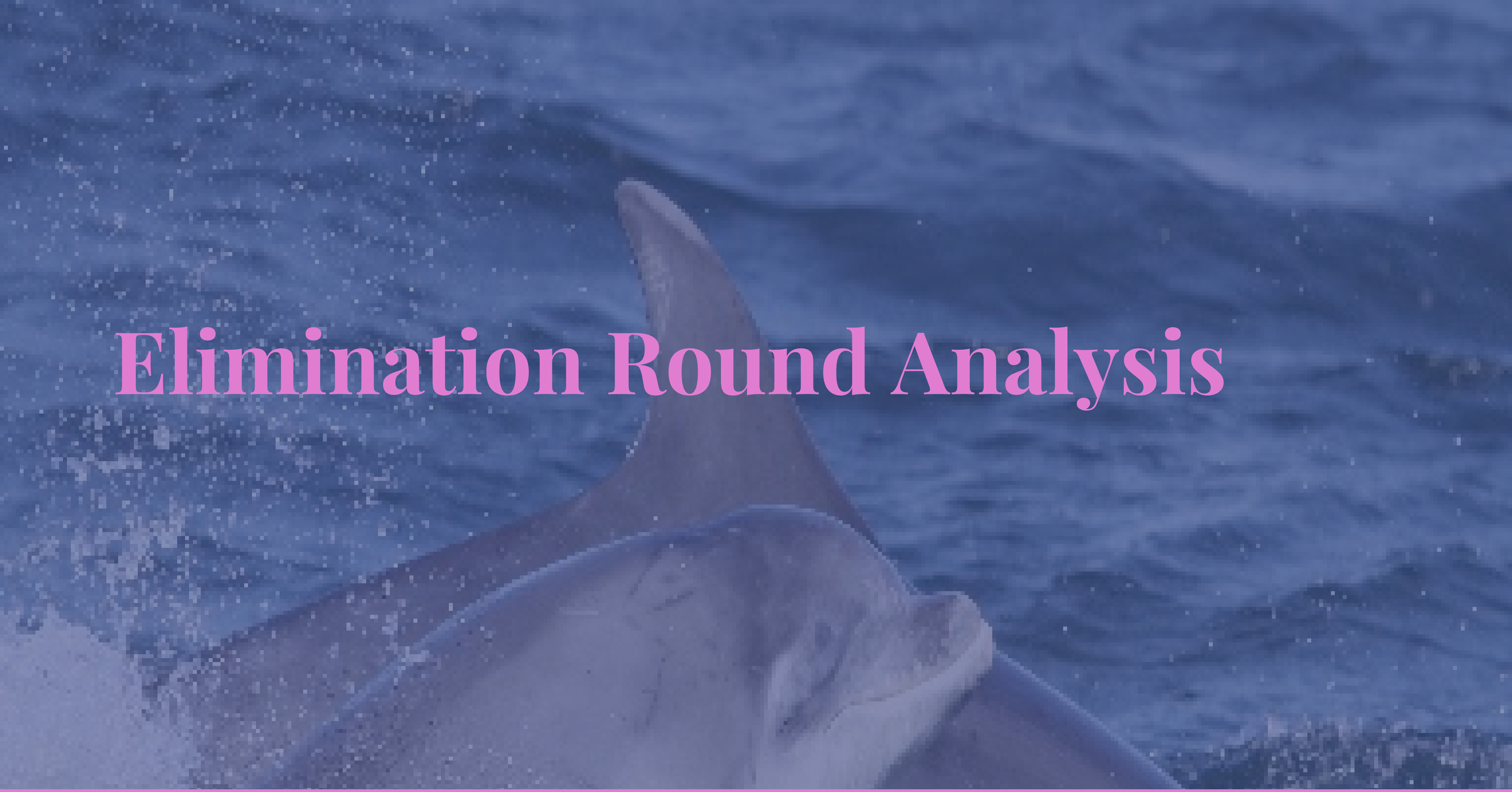
Elimination Round





Any% Speedrun Times





Elimination Round Analysis

CoatNet 0

Marrying Convolution and Attention for All Data Sizes (2021)

14.9M tunable params

17.7 non-tunable params

2 hr 44 min

64.06%

ImageNet

Unranked,
but CoatNet7
currently 1st

Intuition & Discoveries

- Not such a prodigy after all
- Perhaps needed more data to take advantage of model capacity

LeViT 192

LeViT: a Vision Transformer in ConvNet's Clothing for Faster Inference (2021)

ImageNet

Ranked 319

2.2M tunable params

7.9 non-tunable params

2 hr 12 min

70.94%

Intuition & Discoveries

- Less available community advice with SOTA 3rd-party models
- Educated guess on which layers to unfreeze
- No optimal pre-process function

EfficientNet Bo

CondConv: Conditionally Parameterized Convolutions for Efficient Inference (2019)

22.3M tunable params

2.7 non-tunable params

3 hr 8 min

69.96%

ImageNet

Ranked 312

Intuition & Discoveries

- EfficientNet not so efficient
- Too much tradeoff betwen model efficiency and capacity
- Ended up adding a lot of trainable parameters

Inception V3

Rethinking the Inception Architecture for Computer Vision (2016)

ImageNet

Ranked 371

15.8M tunable params

10.7 non-tunable params

2 hr 13 min

76.00%

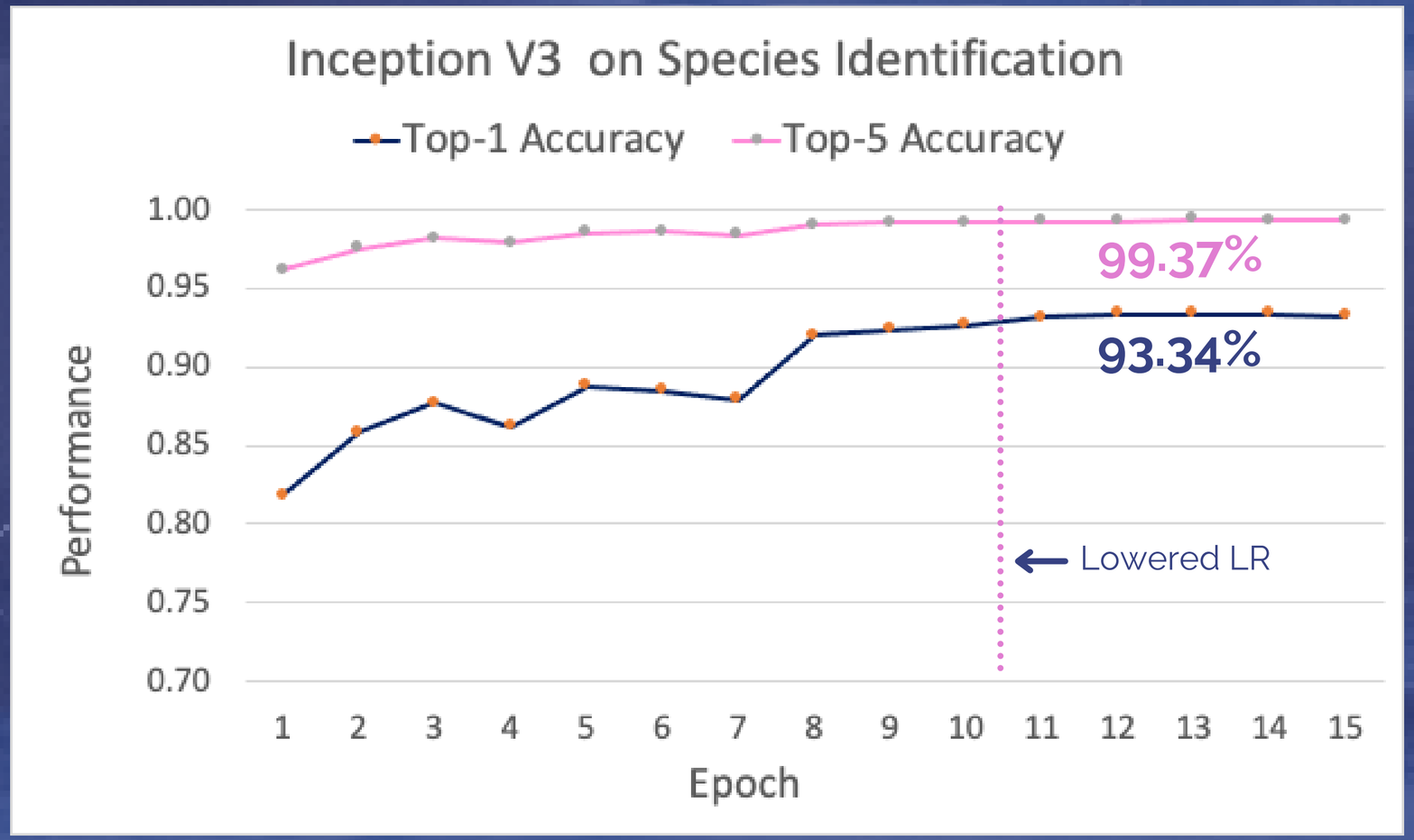
Intuition & Discoveries

- Started off modest but greatly improved each epoch
- Unfreezing the top layers had huge impact on performance
- Could keep learning rate high even during fine-tuning

The Finals Round

using Inception V3

1. Unfroze top of base layer from the onset
2. Ran for 10 epochs before fine-tuning
3. Used **ReduceLROnPlateau**



Tensorflow Tricks

1

preprocess_input()

Lambda layers to utilize optimized preprocessing functions intended for keras pre-trained models

2

ModelCheckpoint()

Lifesaver if Jojie goes down or model performance degrades

3

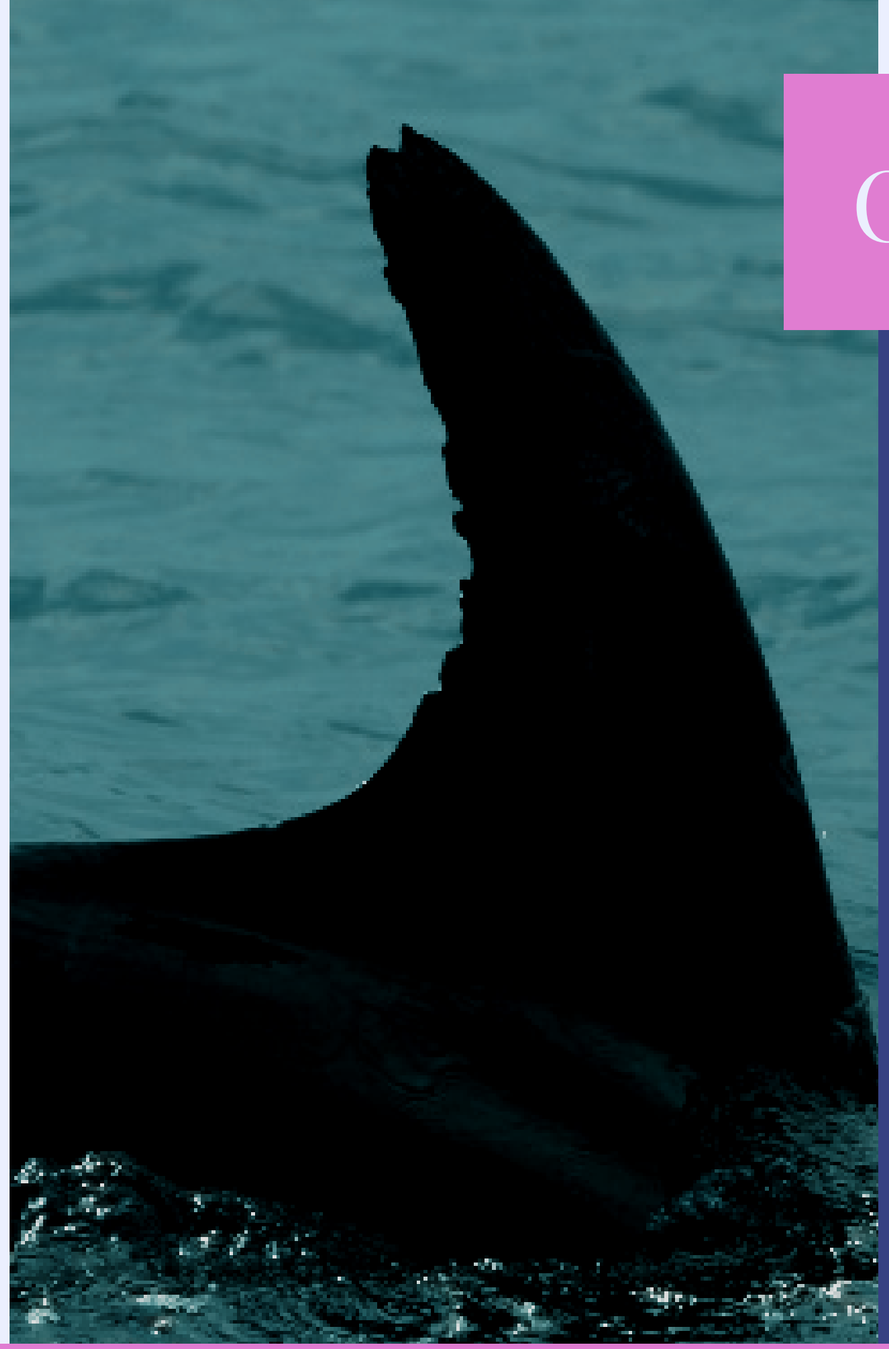
EarlyStopping()

Stop running already if model has converged

4

ReduceLRonPlateau()

Squeeze out better performance near the minima



Official Tournament

For **individual identification (15,587 classes),** might need:

- To forego transfer learning and train weights from scratch
- To utilize higher resolutions
- More time (several days) to run (hundreds of) epochs

For now, I ran only 10 epochs over 5 painful hours...

My Trial Submissions to the Competition Leaderboards

#	Team	Members	Score	Entries	Last	Code
696	**Norman Lapid**		0.002	1	1s	

🙂 Your First Entry!
Welcome to the leaderboard!

#	Team	Members	Score	Entries	Last	Code
590	**Norman Lapid**		0.105	2	18m	

🥳 Your Best Entry!
Your most recent submission scored 0.105, which is an improvement of your previous score of 0.002. Great job!

Tweet this

Ongoing & Next Steps

WAYS TO IMPROVE RANKINGS

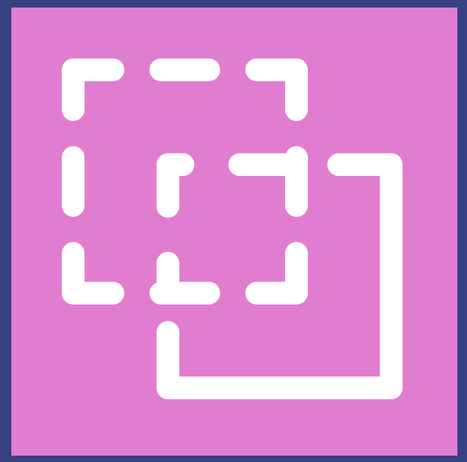
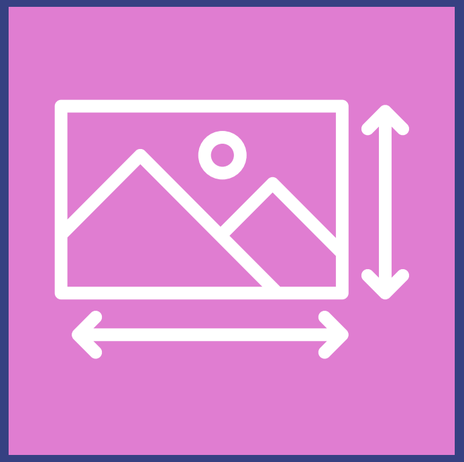
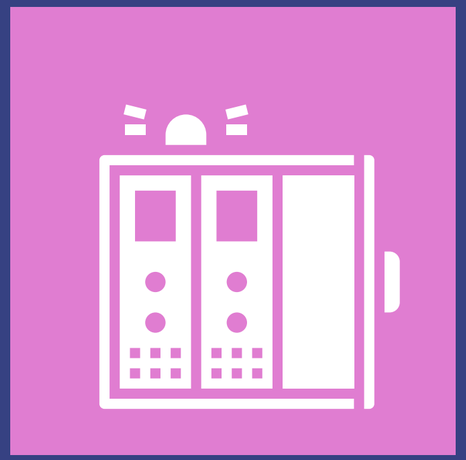


Image cropping (YOLO) without creating OOD images



Apply on higher resolution images



More time and powerful computing resources

WORDS TO PONDER

"If I have seen further it is by standing on the shoulders of giants."

- Isaac Newton

Thank you.