MosquitoNet

Jay Gupta & Aliza Kajani

April 24, 2020

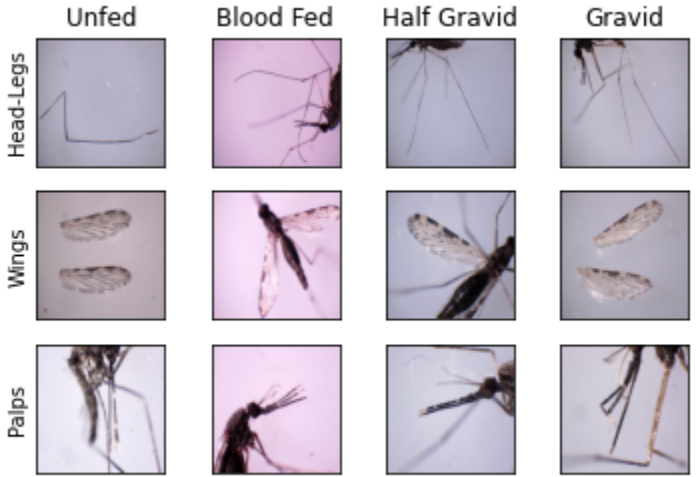
Duke
PRATT SCHOOL of ENGINEERING

Background

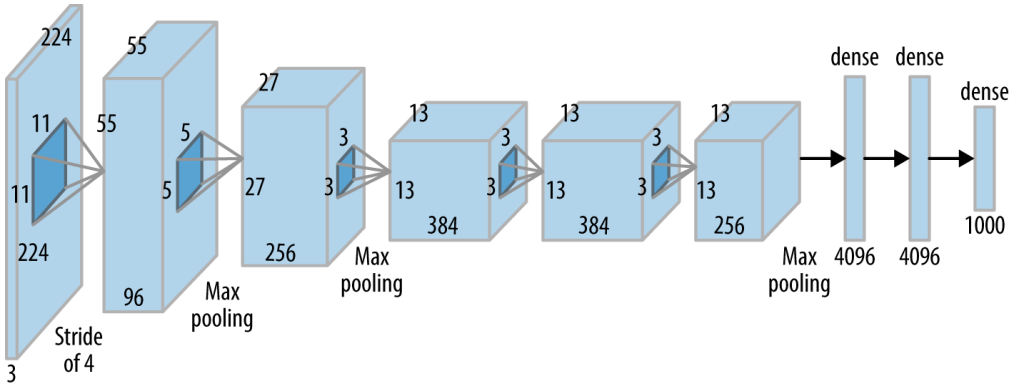
- Malaria incidence: >200M new cases per year
  - Infection caused by *Plasmodium* parasite, carried by mosquitoes
  - Global mosquito surveillance is a component of epidemiological studies
  - Gonotrophic (reproductive) cycle state can be determined through visual inspection of mosquito abdomen
- Objective:
  - Develop a deep learning-based classifier for identification of gonotrophic state from images of mosquitoes

Dataset

- Compiled by DGHI researchers
- Images of 1,327 unique mosquitoes
- Each mosquito imaged from three directions: head-legs, wings, palps
- Species: *gambiae*, *funestus*, or *demeilloni*
- Gonotrophic state: unfed, blood fed, gravid, or half gravid



Methods

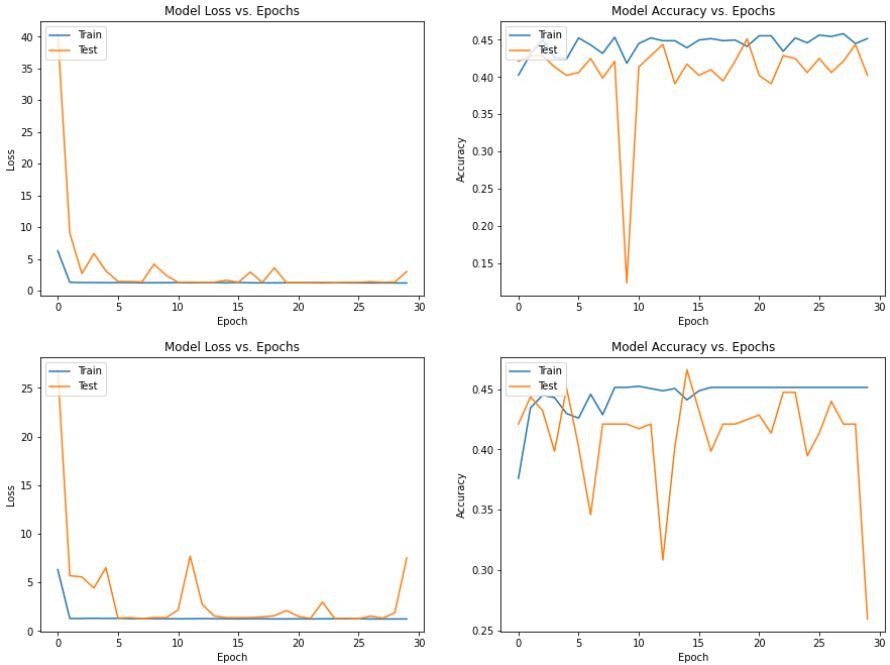


- AlexNet architecture: 5 conv layers, 3 FC layers, softmax classifier
- Physical layers
  - Camera angle (head-legs, wings, palps)
  - Pixel size (downsampled to 224x224 px)
- Simulations performed: baseline, data augmentation, and pre-training

Results: AlexNet, head-legs angle only

- Baseline
- 40.23% accuracy

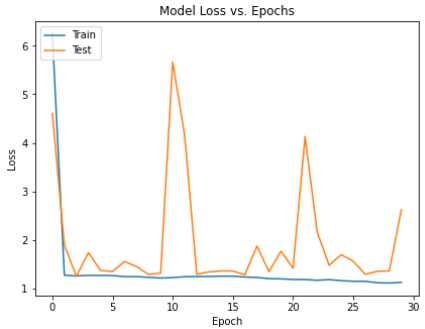
- With data augmentation
  - Rotations
  - Horizontal + Vertical flips
  - Height + Width shifts
- 42.11% accuracy
  - before overfitting

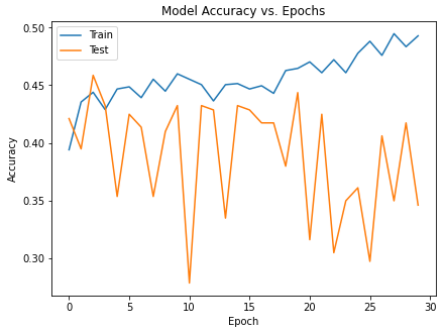


Duke
PRATT SCHOOL of ENGINEERING

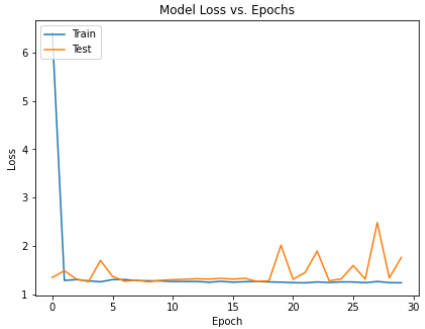
Results: AlexNet, wings angle only

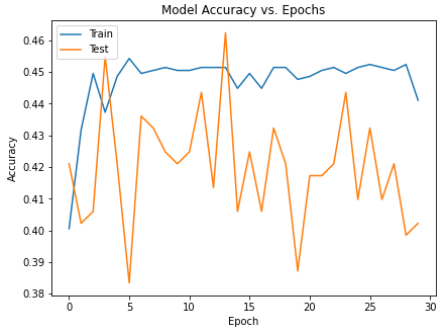
- Baseline
- 34.59% accuracy





- With data augmentation
- 40.23% accuracy

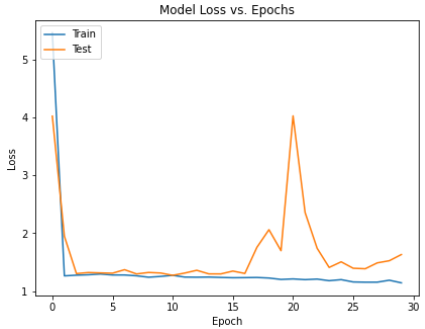


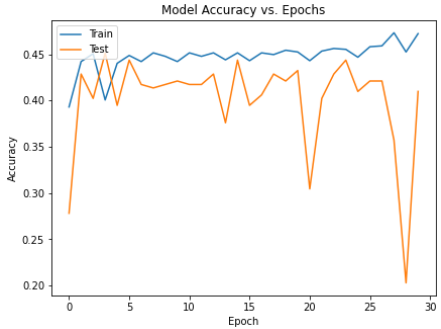


Duke
PRATT SCHOOL of ENGINEERING

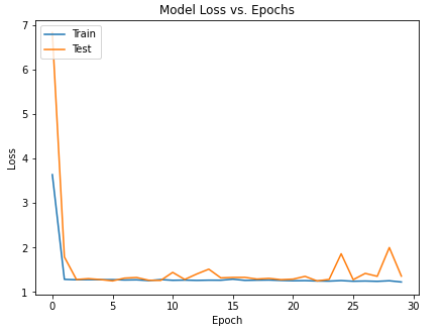
Results: AlexNet, palps angle only

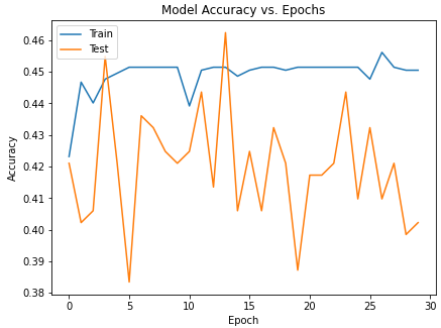
- Baseline
- 40.98% accuracy





- With data augmentation
- 39.85% accuracy

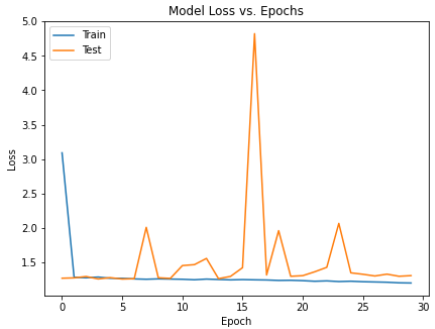


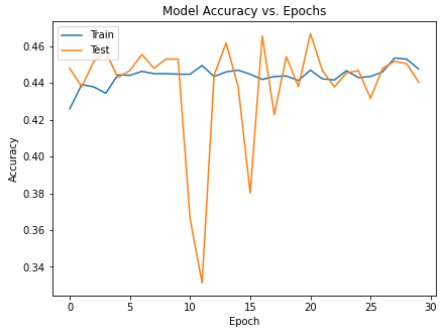


Duke
PRATT SCHOOL of ENGINEERING

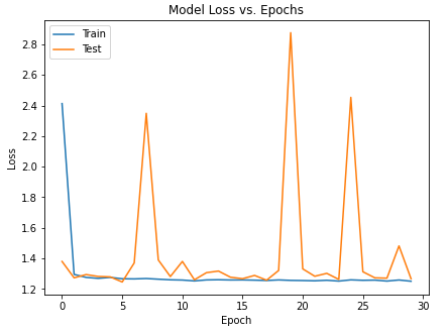
Results: AlexNet, all three angles

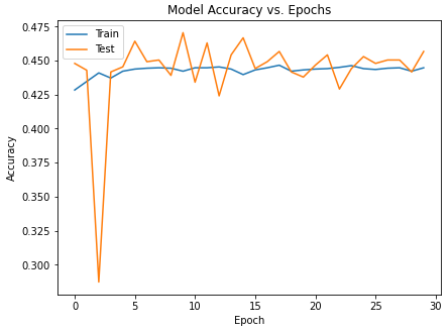
- Baseline
- 44.04% accuracy





- With data augmentation
- 45.67% accuracy

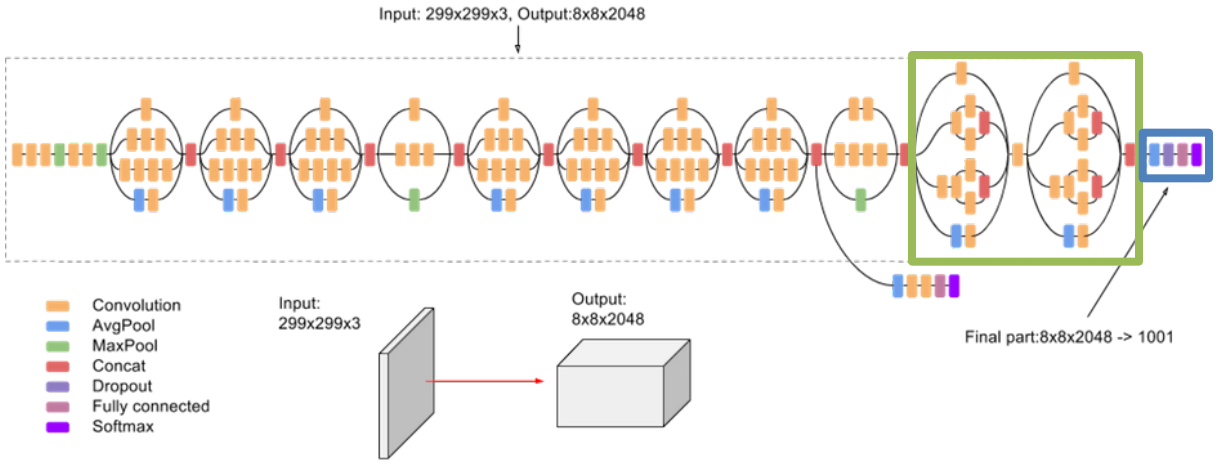




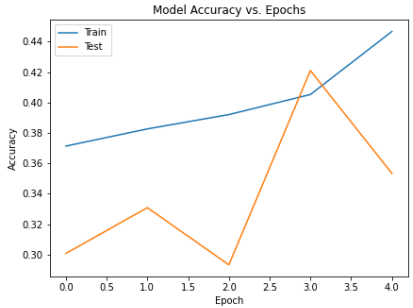
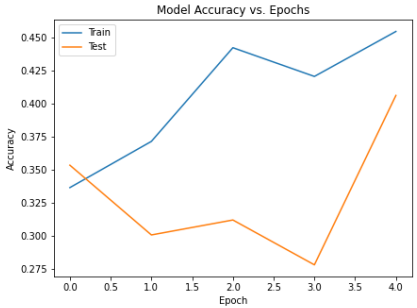
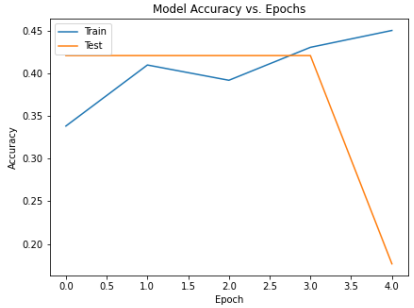
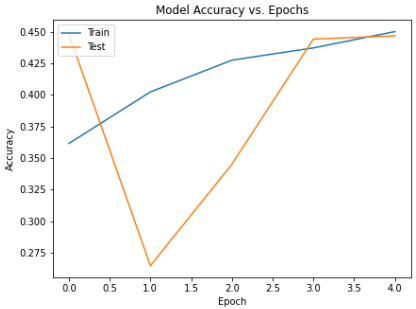
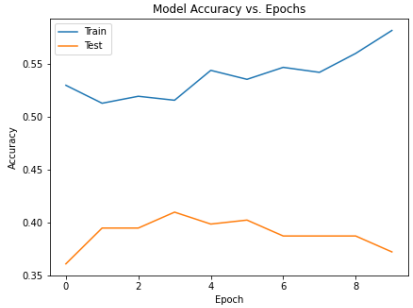
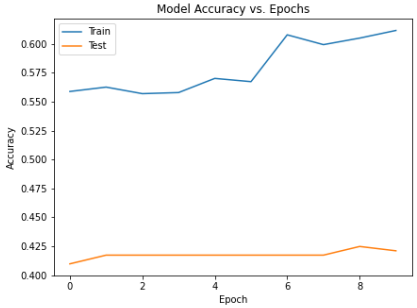
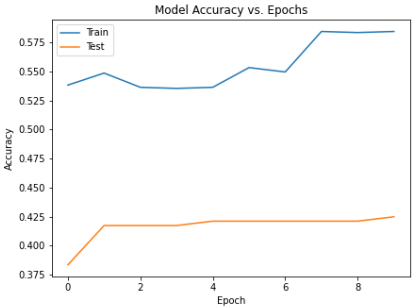
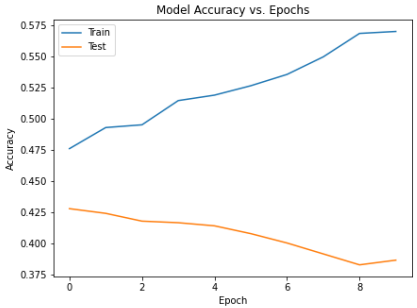
Duke
PRATT SCHOOL of ENGINEERING

Pre-trained model in Keras

- Keras Applications API contains several, downloadable image classification models pre-trained using ImageNet dataset
- InceptionV3 model – powerful but supposedly easier to train
  - Suggest fine-tuning the last 2 blocks of this model + classifier

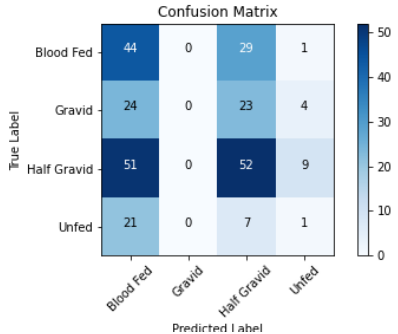


Results: InceptionV3 model, baseline

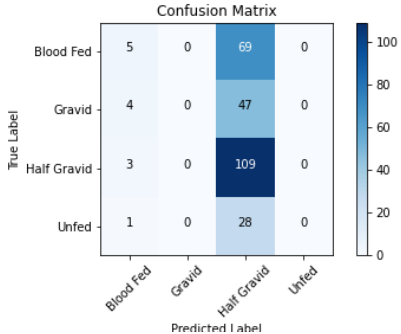
	Head-Legs	Wings	Palps	Full
Classifier only				
Fine Tuning				

Results: InceptionV3 model, confusion matrices

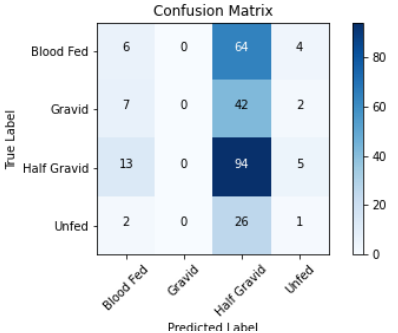
Head-Legs



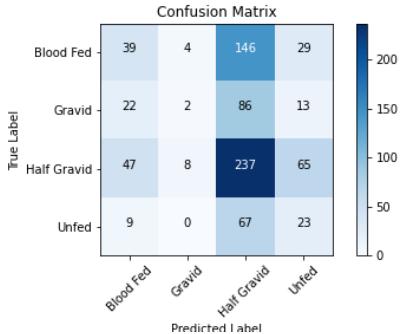
Wings

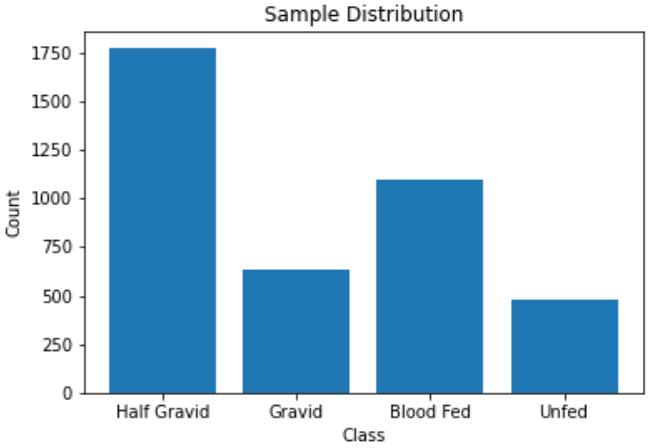


Palps



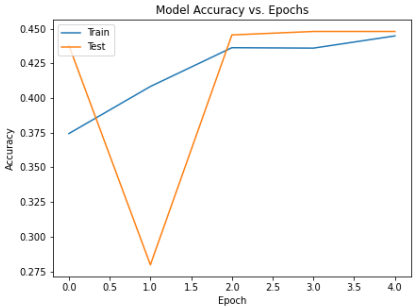
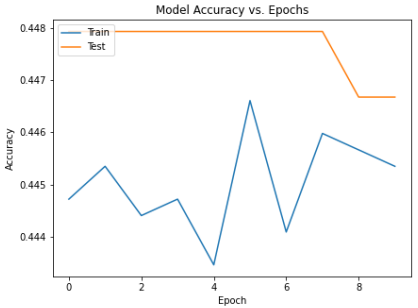
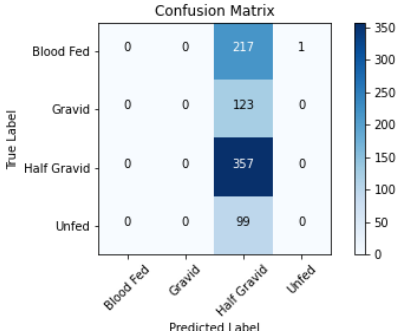
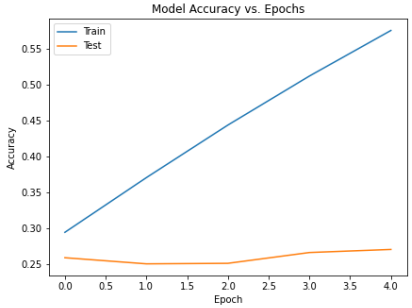
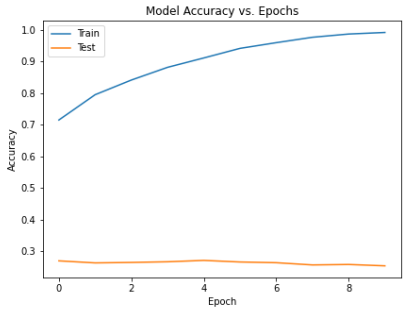
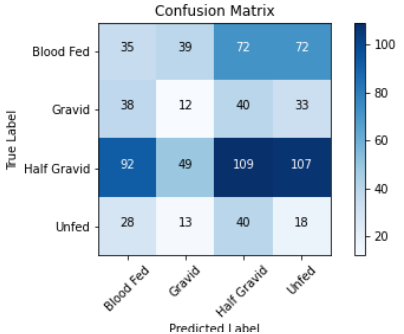
Full





Duke
PRATT SCHOOL of ENGINEERING

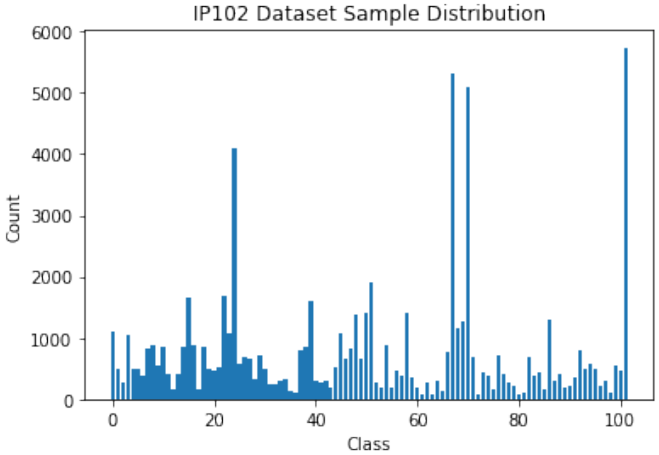
Results: InceptionV3 model, all camera angles

	Classifier Only	Fine Tuning	Confusion Matrix
Data Augmentation			
Oversampling			

Duke
PRATT SCHOOL of ENGINEERING

Pre-training with IP102 dataset

- Benchmark dataset developed for insect pest recognition tasks
- 75,222 images & 102 classes
- Top 4 classes used

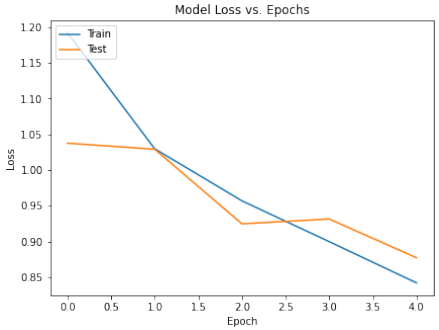


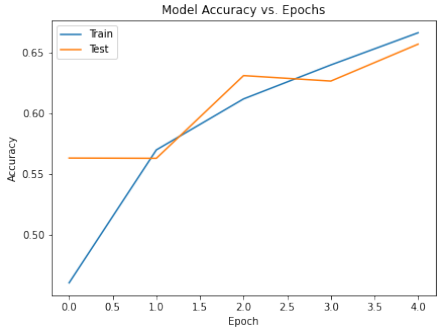
Duke
PRATT SCHOOL of ENGINEERING

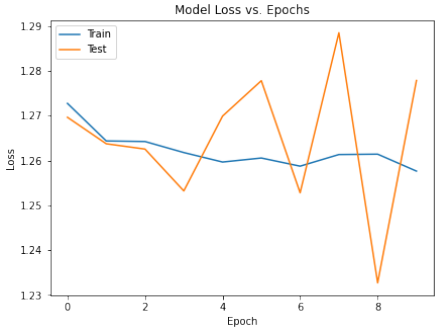
Results: pre-training with IP102 dataset

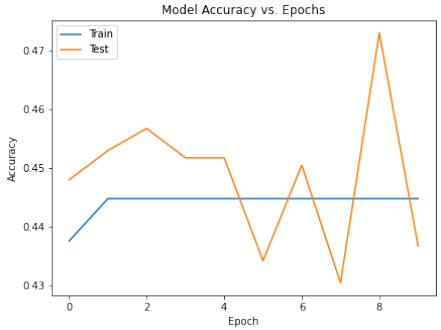
- Pre-training
- 65.68% accuracy

- Fine tuning
  - all camera angles
- 43.66% accuracy









Duke
PRATT SCHOOL of ENGINEERING

Summary

- Three major approaches considered
  - AlexNet trained on mosquito dataset
  - InceptionV3 pre-trained on ImageNet and finetuned on mosquito dataset
  - AlexNet pre-trained on IP102 and finetuned on mosquito dataset
- Challenges encountered → solutions considered
  - Overfitting → data augmentation
  - Class imbalance → random oversampling of the minority classes
- Physical layer: camera angle
  - No major differences in camera angle identified, but performance was weak for all

Future Directions

- Alternative physical layers (e.g. downsampling factor)?
- Less complex CNN architecture to reduce overfitting?
- More nuanced methods to handle class imbalances?
  - Alternative performance metrics (e.g. balanced accuracy for binary tasks)?
  - Alternatives to *random* oversampling?
- Ensemble models?
  - Is there a way to make predictions using all three images, simultaneously?