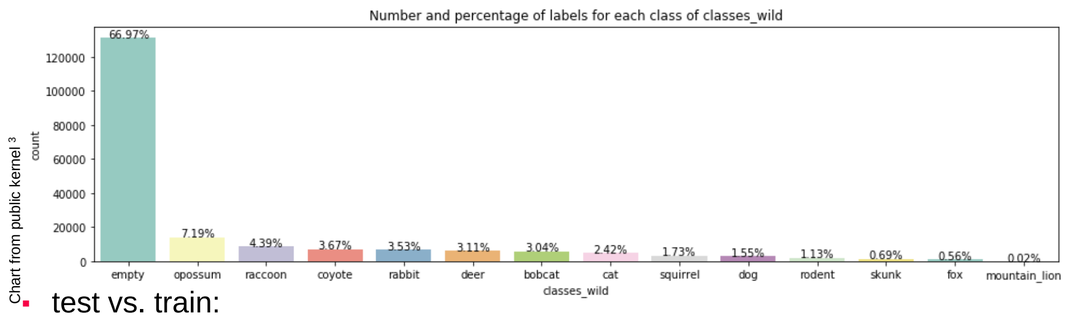
Animal classification from trap cameras - from iWildCam 2019 Kaggle challenge

Adrian Spataru, Hussain Hussain, Nikita Lvov
Institute of Interactive Systems and Data Science

Image from public kernel [3]

Problem and Data

- Trap images
- Training set taken in the American Southwest
- Test set taken in the American Northwest
- 14 classes: 13 animals + empty
- Unevenly distributed



- test vs. train:
  - nature/environment
  - class distribution (noticed during experiments)

Samples from training set

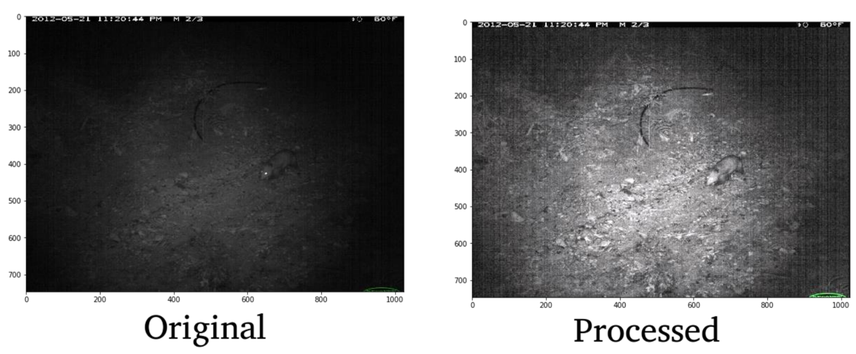
Samples from test set

Approaches

Image pre-processing [1]

For better image recognition experience the next techniques were used:

- CLAHE (Contrast Limited Adaptive Histogram Equalization)
- Simple WB (Algorithm stretching the image channels to the specified range)
- Rotation and Warping



An example that shows the result of the image preprocessing stage

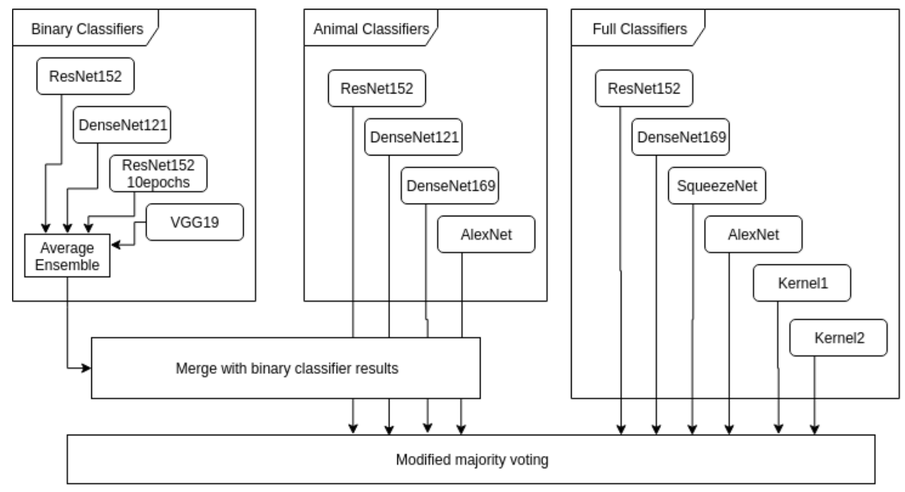
Deep learning models

- **Two-stage Pipeline Model** (Idea from Sadegh, et al., 2018[2])
  - The binary classifier: **Is the animal presented or not?**
  - The animal classifier: **Which animal is it?**

    ➡ Override the second classifiers result by the first one
- **Full model** classify all of the 14 classes including the *empty* class
  - includes two implemented Kaggle kernels [3,4]

Average ensemble of 2-stage methods

- Separate approach
- Average ensemble **× 4 binary classifiers**
- Average ensemble **× 4 animal classifiers**
- Merge ➡ 1 classifier



Ensemble

- Average ensemble **× 4 binary classifiers**
- Merge result **× 4 animal classifier** ➡ 4 complete models
- Finally: MAJORITY VOTING ENSEMBLE **× 10 models**

Filtered voting

Empty → **AT LEAST 6 models** voted for empty

Misclassified deer as dog/coyote → replace dog/coyote votes with deer

Results

- Private/public leaderboard on kaggle.com
- Included late submissions (denoted with *****)

Approach	Private	Public
Time metadata	0.80	0.90
DenseNet169 (all in one)	0.103	0.106
DenseNet121 (2-stage)	0.108	0.113
Average ensemble (on 2 stages)	0.107	0.110
*Majority vote (not filtered)	0.106	0.118
*Majority vote + reassign votes	0.152	0.165
*Majority vote + empty vote majority	0.116	0.124
Maj. vote - fully filtered - all models	**0.159**	**0.166**
*Maj. vote filtered (only 2-stage)	0.154	0.162
*Maj. vote filtered (only full)	0.148	0.161

Insights

- Filtered majority voting: ~50% improvement over a single model
- Average ensemble didn't actually improve
- 2-stage models > full models ➡ divide and conquer
- Necessity of taking empty class on majority
- Reassigning votes helped the most. Empty majority: not as much

Rank

- 21st/336 on public leaderboard
- 22nd/336 on private leaderboard **(in top 7%)**

Literatur / Zitat

[1] Chanran, K. (2019, March). ***Image Pre-processing for Wild Images***. Retrieved from kaggle.com:https://www.kaggle.com/seriousran/image-pre-processing-for-wild-images#CLAHE-&-Simple-WB

[2] Sadegh, M. N., Nguyenb, A., Kosmalac, M., Swansond, A., Palmere, M. S., Packere, C., & Clunea, J. (2018). **Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning**. *PNAS Journal.*

[3] Preda, G. (2019, May). ***iWildCam 2019 EDA and Prediction.*** Retrieved from https://www.kaggle.com/gpreda/iwildcam-2019-eda-and-prediction

[4] (2019, March). ***DenseNet Transfer Learning - iWildCam 2019.*** Retrieved from https://www.kaggle.com/xhlulu/densenet-transfer-learning-iwildcam-2019

TU Graz –*Institute of Interactive Systems and Data Science*
8010 Graz, Inffeldgasse 16/C, Austria