# Classification of Butterfly Species

Thomas Martin Berger
thberger@stanford.edu

Stanford University

CS231n: Deep Learning for Computer Vision
Stanford University

June 4, 2022

# Overview

# Introduction & Problem Statement

**Introduction & Related Work**

- Automatic classifier as an useful tool to facilitate the work of scientists
- Usage of "realistic" dataset compared to other datasets considered in related papers (project data includes background noise, butterfly image taken from several angles, imbalanced data, ...)

**Problem Statement**

- **Task**: Multi-class classification (110 different species)
- **Input**: Butterfly image
- **Output**: Labeled image
- **Metrics**: Accuracy, Macro F1-score

# Dataset

Provided by the association "Arbeitsgemeinschaft bayerischer Entomologen e.V.", enriched with images from Kaggle

| **Species name** | **Number of observations** |
|---|---|
| Polygonia c-album | 3633 |
| Gonepteryx rhamni | 3096 |
| Aglais urticae | 2487 |
| Anthocharis cardamines | 2346 |
| Iphiclides podalirius | 2273 |
| Argynnis paphia | 2210 |
| ... | ... |
| Euphydryas cynthia | 258 |
| Boloria thore | 199 |
| Oeneis glacialis | 119 |
| Plebejus argyrognomon | 47 |
| Maculinea telejus | 45 |
| Maculinea rebeli | 26 |

# Dataset (cont.)

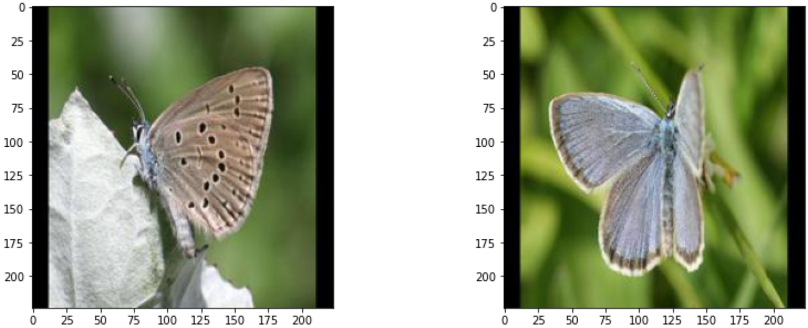



Figure 2: Maculinea Rebeli examples

# Methods

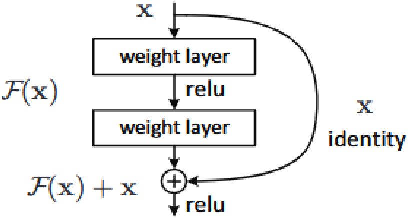
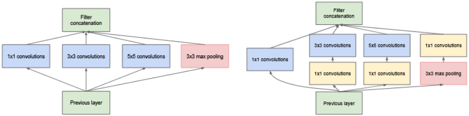



Figure 3: Residual learning, as described in He *et al.* [1]




(a) Inception module, naïve version (b) Inception module with dimension reductions

Figure 4: Inception architectures, as described in Szegedy *et al.* [2]

- **Baseline Model**: Includes convolutional layer followed by batch-normalization before ReLU; additionally includes a Residual Layer
- **Residual-11**: Extension of baseline
- **ResNet50**: Pretrained model (ImageNet) with residual layers
- **Inception-ResNet-v2**: Pretrained model including concept of inception and residual layers

# Methods (cont.)

**Losses**

- Categorical Cross Entropy

$$L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{ik} \log f_k(x_i)$$

- Weighted Categorical Cross Entropy

$$L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} w_k y_{ik} \log f_k(x_i),$$

## Class weights

$$w_k = 1 + \log\left(\frac{\#\text{ of observations in most frequent class}}{\#\text{ of observations in class } k}\right)$$

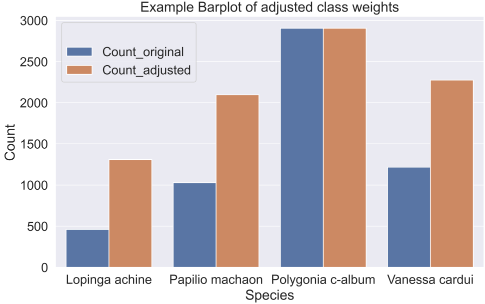



Figure 5: Number of observations scaled with class weights

| | Baseline Model | Residual-11 | ResNet50 | Inception-ResNet-v2 |
|---|---|---|---|---|
| Train Acc | 64.22% | 69.55% | 81.75% | 85.07% |
| Train F1 | 57.81% | 63.19% | 76.98% | 81.05% |
| Valid Acc | 63.70% | 67.52% | 79.70% | 82.39% |
| Valid F1 | 58.13% | 62.56% | 74.55% | 77.57% |
| Test Acc | 64.10% | 67.92% | 79.82% | **81.93%** |
| Test F1 | 57.93% | 62.70% | 73.77% | **77.56%** |

Table 1: Results without class weights

| | Baseline Model | Residual-11 | ResNet50 | Inception-ResNet-v2 |
|---|---|---|---|---|
| Train Acc | 44.58% | 78.50% | 84.31% | 83.51% |
| Train F1 | 37.87% | 74.76% | 81.10% | 79.94% |
| Valid Acc | 40.32% | 77.13% | 80.92% | 80.81% |
| Valid F1 | 34.44% | 74.56% | 76.88% | 76.44% |
| Test Acc | 40.91% | 77.25% | 80.92% | **81.35**% |
| Test F1 | 34.79% | 73.34% | 76.88% | **76.94**% |

Table 2: Results with class weights

# Conclusion

**Main results:**

- Inception-ResNet-v2 model delivered the best results probably due to its deep and robust architecture
- Considering pretrained weights were useful
- Mixed results for applying class weights

**Next Steps:**

- Collecting more data for species with low number of observations
- Tuning
- Testing other models

# Bibliography


[1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[2] C. Szegedy, W. Liu, Y. Jia, *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[3] K. R. M. Fernando and C. P. Tsokos, "Dynamically weighted balanced loss: Class imbalanced learning and confidence calibration of deep neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.