

Project I - Image Classification

Team: Filip Kołodziejczyk, Jerzy Kraszewski

Dataset

[Dataset - CINIC 10 Public DataSet](#)

The dataset contains ~180k images of 10 classes. Those are 32x32 color (RGB) images. The dataset comes with established division into training and test sets.

General approach

Since the task is open-ended, we want to tackle it with an iterative approach. We will start with a simple code with one model end-to-end (from data loading to evaluation). Then we will modularize the code to be able to easily swap models. We will start with a simple model (e.g. VGG) and then move to more complex ones (e.g. ResNet, EfficientNet). On top of that we will check the influence of hyperparameters, possibly the same for all models, on the obtained results (both training and regularization). We will also investigate the influence of data augmentation techniques on the obtained results. Finally, we will apply ensemble methods (hard/soft voting, stacking) to the models. If there will be time left, we would like to use CapsNet and Vision Transformer.

Plan

We want to split the project into 3 phases (each phase we want to discuss with supervisor). The plan is as follows:

1. **Phase 1** (deadline: 3rd project meeting)

- ☐ Load the dataset
- ☐ Preprocess the dataset
- ☐ Implement VGG
- ☐ Train the model (without hyperparameter tuning)
- ☐ Evaluate the model
- ☐ Add data augmentation techniques
- ☐ Evaluate the model
- ☐ Add hyperparameter tuning
- ☐ Evaluate the model
- ☐ Add regularization techniques
- ☐ Evaluate the model

2. **Phase 2** (deadline: 5th project meeting)

- ☐ Eventual code fixes
- ☐ Modularize the code for easy swapping of models and hyperparameters (automatic process)
- ☐ Add ResNet and EfficientNet
- ☐ (optional) Add Inception CNN - GoogLeNet
- ☐ (optional) Add CapsNet
- ☐ (optional) Add Vision Transformer
- ☐ Prepare initial report

- ☐ Prepare initial presentation
- 3. **Phase 3** (deadline: 6th project meeting, project deadline)
 - ☐ Eventual code fixes
 - ☐ Finishing the report
 - ☐ Presentation preparation

Report outline:

1. Problem introduction (image classification)
2. Dataset description
3. Very short review of literature (available architectures)
4. Justified choice of architectures
5. Methodology description:
 - Data preprocessing
 - Training process
 - Hyperparameter tuning
 - Evaluation
6. Short code description (used libraries, etc., how to use it)
7. Results
8. Conclusions

Notes:

- at least 2 hyper-parameters related to training process
- at least 2 hyper-parameters related to regularization
- at least X data augmentation techniques from the following groups:
 - standard operations (where $x=3$)
 - more advanced data augmentation techniques like mixup, cutmix, cutout (where $x=1$)
- constant seed for reproducibility
- only given dataset
- preferred pre-trained models
- compare to achieved accuracy [CIFAR-10 accuracy for methods](#) and
- adhere to statistical significance
- variance not less important than mean