KTH Royal Institute of Technology
Data Science
Spring P4
DD2424 Deep Learning in
Data Science

Project Proposal

Glenn Wei Zheng Tan
Sophia Nicole Chua
Yerui Huang

Project Title

Furry Classification

Project Direction

Default project 1 + some extension

Project Description

A brief description of the problem that you will work on and how you will try to solve it. Reference to at least one paper (or webpage) that inspired as regards the problem statement and/or proposed approach to tackling the problem. (It is completely fine if you just want to replicate results of some paper and explore the influence of parameter settings, training conditions, see if the method can be transferred to another dataset etc...)

1. Felis or Canis: Image Classification Using Deep Learning | IEEE Conference Publication | IEEE Xplore
2. Research on Cat and Dog Image Recognition Based on Several Classic Neural Networks | IEEE Conference Publication | IEEE Xplore
3. A Case Study: Cat-Dog Face Detector Based on YOLOv5 | IEEE Conference Publication | IEEE Xplore
4. The Application and Expansion of the YOLO Algorithm in the Field of Campus Cat and Dog Identification | IEEE Conference Publication | IEEE Xplore
5. A Small Sample Image Recognition Method Based on ResNet and Transfer Learning | IEEE Conference Publication | IEEE Xplore

KTH Royal Institute of Technology

Data Science

Spring P4

DD2424 Deep Learning in Data Science

The objective of this project is to explore and implement transfer learning using modern convolutional neural networks (CNNs) to classify the images of pets — specifically cats and dogs. In particular, there are two levels of classification that we aim to achieve:

1. **Binary Classification:** Distinguish between pictures of cats and dogs
2. **Multi-class Classification:** Identify the dog or cat breed from 37 possible categories

Following the project description, this project will investigate how to adapt a pre-trained model, such as ResNet18 or ResNet34, to the aforementioned tasks by fine-tuning its architecture and optimizing its performance. Additionally, we will experiment with strategies to address challenges like class imbalance, data augmentation, and overfitting.

Each task will roughly be approached through the following steps:

1. **Binary Classification**:

    - Replace the final layer of a pre-trained ResNet with a binary classifier.
    - Fine-tune the final layer using the Adam optimizer on the pet dataset.
    - Achieve ≥99% test accuracy by leveraging transfer learning's power.

2. **Multi-class Classification**:

    - Modify the final layer to output probabilities for 37 classes.
    - Explore two fine-tuning strategies:
        - Fine-tuning the last *l* layers of the network progressively.
        - Gradual unfreezing, starting with the top layers and progressively unlocking earlier layers.
    - Experiment with learning rate schedulers, L2 regularization, and data augmentation techniques (e.g., flips, rotations, random crops) to improve performance.

3. **Class Imbalance**:

    - Simulate class imbalance by reducing training samples for certain breeds.
    - Evaluate the impact on classification performance using standard cross-entropy loss.
    - Implement strategies such as weighted cross-entropy or oversampling to mitigate performance degradation.

4. **Advanced Investigations**:

    - Study catastrophic forgetting by fine-tuning the pet classifier on a different dataset (e.g., 102 Flower Dataset) and then re-evaluating the original task.
    - Experiment with batch normalization fine-tuning, sophisticated data augmentations (e.g., CutMix, MixUp), and deeper networks (e.g., ResNet50 or ResNet101).

5. **Metrics and Validation**:

    - Use accuracy, precision, recall, and F1-score to evaluate performance.
    - Employ a validation set to tune hyperparameters and select models.

Dataset: Training, Validation and Testing

The Oxford-IIIT Pet Dataset

Deep Learning Software Packages

PyTorch, TensorFlow

KTH Royal Institute of Technology
Data Science
Spring P4
DD2424 Deep Learning in
Data Science

Software Implementation

How much of the software implementation your group will write and how much you will rely on open-source implementations. It is not acceptable though to rely completely on open-source implementations and just change parameter settings. One goal of the project is to learn some deep learning package to allow you to set up architectures and training procedures.

We will download a pre-trained modern ConvNet and change the first and final layer and fine tune them as required in the project 1 pdf. We will write codes to explore which strategy has a better performance on the validation dataset and how to have a stable performance on an imbalance training data.

Baselines and Initial Experiments

Baselines:

The model should be able to solve the binary classification problem of recognising pictures of Dog Vs Cat with an accuracy of ~ 95%, and solve the multi-class classification problem of recognising the breed of 37 outputs of cat or dog, with a final test accuracy.

Initial Experiments:

- **Dataset:** Approximately 7400 images of 37 distinct pet breeds of cats and dogs, with roughly 200 images for each breed
- **Model Architecture:** Defines a custom convolutional neural network (CNN) with two convolutional layers followed by three fully connected layers.
- **Training Approach:** Manually implements the training loop, including forward pass, loss computation, backward pass, and weight updates.
- **Data Augmentation:** Applies normalization to the dataset to standardize the input features.

KTH Royal Institute of Technology

Data Science

Spring P4

DD2424 Deep Learning in Data Science

Project Milestones

E grade:

1. Replace the final layer of the pre-trained ConvNet to solve the binary classification problem of recognising pictures of Dog Vs Cat.
2. Solve the multi-class classification problem of recognising the breed of cat or dog.
3. Fine-tuning with imbalanced classes

D-C range:

1. Test and explore if deeper networks would significantly improve the classifier. We will observe the challenges of using a deeper network along the way.
2. Explore the idea of catastrophic forgetting
3. Fine-tune the batch norm mean and standard deviation and keep the weights of the layers (except the final layer) frozen and see if you can improve results.
4. Add more sophisticated data-augmentations such as random erasing, CutMix or MixUp to help with regularization.

Project Learning Goals

Specify for each group member the skills/knowledge w.r.t. deep learning "theory" and practice they aim to acquire from completing the project.

Yerui Huang

Through this project, I hope to gain a deeper understanding of deep learning theory and practice including how to design and train deeper neural networks, apply transfer learning, and explore the effects of catastrophic forgetting. I also want to implement advanced data augmentation techniques such as CutMix to improve generalization. Practically, I can enhance my skills in using PyTorch to build custom models and training pipelines.

Glenn Wei Zheng Tan

Through this project, I aim to deepen my understanding of how to apply deep learning theories in practice by building a model that accurately classifies real-world

objects. I'm particularly focused on learning how to effectively fine-tune parameters across multiple network layers. Additionally, I want to gain hands-on experience with PyTorch, as it is a critical tool for my future career aspirations in data science.

Sophia Nicole Chua

Through this project, I aim to strengthen my understanding of the underlying principles of transfer learning, particularly, how pre-trained networks generalize to new tasks, through implementation. I want to delve into optimizing neural network training by experimenting with different optimizers, learning rate schedulers, and regularization techniques. On the practical side, I hope to enhance my expertise in PyTorch by developing end-to-end training workflows, incorporating advanced data augmentation strategies, and effectively managing computational resources for large-scale image classification tasks.

Project Aim

A grade of D-C