Deep Learning and Neural Networks

Demetrio Labate

March 8, 2025

Final Projects

Project Guidelines

You are given some flexibility in the choice of your final project, within the following guidelines:

1. You are encouraged to form a group of 2 people (no more than 3).
2. The project **must include some original work** such as the comparison of different methods, the analysis of new/improved training; it is not sufficient to run an existing model on new data.
3. The team will deliver a 20 min Powepoint/Latex/Keynote presentation in class.
4. Your project and presentation must include a critical discussion of the results.

Project Guidelines

- The project can be based on a paper or a few papers.
- The project can be based on a "challenge"

Possible topics:

1. Transfer Learning;
2. Object detection;
3. Implicit Neural Representations;
4. Emotion Recognition;
5. Adversarial training, adversarial attack;
6. Text generation;
7. Interpolation in latent space

Project ideas

Transfer Learning is a technique in which knowledge learned from a task is re-used in order to boost performance on a related task. It may be especially useful when there are insufficient data to train a model. A related problem is **domain adaptation**, where we aim at learning a model from a source data distribution and applying that model on a different (but related) target data distribution.

- *Top-tuning: A study on transfer learning for an efficient alternative to fine tuning for image classification with fast kernel methods*, by Alfano et al.
- *Combining subclassifiers in text categorization: A dst-based solution and a case study*, by K. Sarinnapakorn and M. Kubat
- Semi-supervised Domain Adaptation via Minimax Entropy, by saito et al.

Project ideas

Object Detection

Object detection is a computer vision task that involves identifying and locating objects in images or videos. One of the earliest successful attempts to address the object detection problem using deep learning was the R-CNN. You Only Look Once (YOLO) proposes using an end-to-end neural network that makes predictions of bounding boxes and class probabilities all at once. It differs from the approach taken by previous object detection algorithms, which repurposed classifiers to perform detection. YOLO v5, v8 are implemented in PyTorch.

- *Kidney Recognition in CT using YOLOv3*, by Lemay.
- *Tomato detection based on modified YOLOv3 framework*, by Lawal.
- *Small-object detection based on YOLOv5 in autonomous driving systems*, by Mahaur and Mishra.

Project ideas

Implicit Neural Representations

Neural implicit representations are neural networks (e.g., MLPs) that estimate the function f that represents a signal continuously (1d, 2D, 3D), by training on discretely represented samples of the same signal. They learn how to estimate the underlying (continuous) function f.

- *Learning Implicit Fields for Generative Shape Modeling*, by Zhiqin Chen and Hao Zhang
- *Deep Learning on Implicit Neural Representations of Shapes*, by Luca De Luigi, Adriano Cardace, Riccardo Spezialetti, Pierluigi Zama Ramirez, Samuele Salti, Luigi Di Stefano

Project ideas

Emotion Recognition or **Emotion detection** is a field within the realm of artificial intelligence and computer vision that involves the identification and interpretation of human emotions from facial expressions or from the voice or from other individual appearance. Accurate emotion detection has numerous practical applications, including human-computer interaction, customer feedback analysis, and mental health monitoring. CNNs have emerged as a powerful tool in this domain.

- Speech Emotion Recognition Using CNN, by Huang et al
- Facial emotion recognition using convolution neural network, by Modi et al
- Vision Transformer With Attentive Pooling for Robust Facial Expression Recognition, by Xue et al
- Text-based emotion detection: Advances, challenges, and opportunities, by Acheampong et al

Project ideas

Adversarial attacks

Neural networks consistently misclassify adversarial examples - inputs formed by applying small but intentionally worst-case perturbations to examples from the dataset, such that the perturbed input results in the model outputting an incorrect answer with high confidence.

- *Explaining and Harnessing Adversarial Examples*, by Ian J. Goodfellow, Jonathon Shlens, Christian Szegedy
- *Practical Black-Box Attacks against Machine Learning*, by Papernot et al.

Project ideas

Text generation is one of the most interesting and challenging applications of deep learning techniques. Sequential data like the text is often difficult to generate because it involves an understanding of context. Even if a model generates words that are often closely associated with each other based on past datasets, one wrong word can render the entire sentence meaningless.

- Using Deep Learning for Text-Generation of Wikipedia Articles
- Generating text using a Recurrent Neural Network
- Generating Text with Recurrent Neural Networks, by Sutskever, Marten and Hinton

Project ideas

Interpolation in latent space. Autoencoders represent an effective approach to computing the underlying factors that characterize datasets of different types. The latent representation of autoencoders has been studied in the context of enabling interpolation between data points by decoding convex combinations of latent vectors.

- Generating In-Between Images Through Learned Latent Space Representation Using Variational Autoencoders, by Critovao et al
- Understanding and Improving Interpolation in Autoencoders via an Adversarial Regularizer, by Berthelot et al
- Autoencoder Image Interpolation by Shaping the Latent Space, by Oring et al
- VConv-DAE: Deep Volumetric Shape Learning Without Object Labels, by Sharma et al

Project ideas

- **Kaggle Competitions** Kaggle Competitions are designed to provide challenges for competitors at all different stages of their machine learning careers. As a result, they are very diverse, with a range of broad types
- **Kaggle Competition Challenges**
- Kaggle: Emotion Detection
- Kaggle: Emotion Detection - faces

- Useful link with State-of-the-art applications organized by topics and related datasets