

TP3: Convolutional Neural Networks and Computer Vision

DONGMO Prince Williams
Supervisor: Dr. Louis Fippo Fitime

October 2025

1 Introduction

This report presents the work done for TP3 in the Deep Learning course at ENSPY. The main objectives are to understand and implement Convolutional Neural Networks (CNNs), residual networks (ResNet), and neural style transfer, using the CIFAR-10 dataset and Keras.

2 Part 1: Theoretical Concepts

- **Convolution:** The filter (kernel) extracts local patterns from the image. The stride controls the movement of the filter. The main goal of a convolutional layer is to extract relevant spatial features while reducing dimensionality.
- **Pooling:** Max Pooling keeps the maximum value in a region (robust to noise). Average Pooling computes the mean (smooths features). Both reduce data size and help extract invariant features.
- **Flattening:** Convolutional layers produce feature maps, which are flattened into vectors for fully connected (Dense) layers for classification.
- **ResNet:** Residual connections (skip connections) allow gradients to flow more easily in deep networks, solving the vanishing gradient problem.
- **Style Transfer:** Combines the content of one image and the style of another using a pre-trained CNN (VGG16).

3 Part 2: Classic CNN Implementation

3.1 CIFAR-10 Data Preparation

- Dataset: 10 classes of 32x32 color images.
- Data normalized to $[0, 1]$, labels one-hot encoded.

3.2 Model Architecture

- 2 convolutional layers (32 and 64 filters, 3x3, ReLU)
- 2 max pooling layers (2x2)
- Flatten, Dense (512, ReLU), Output (Softmax)

3.3 Results

Model trained for 10 epochs, accuracy evaluated on test set.

4 Part 3: ResNet Implementation

4.1 Residual Block

- Main path: 2 Conv2D layers
- Skip connection: input added to output

4.2 Training

ResNet trained on CIFAR-10 for 10 epochs. Test accuracy reported.

5 Part 4: Neural Style Transfer

5.1 Principle

- Content image: provides structure
- Style image: provides artistic style
- VGG16 extracts features; optimization combines both

5.2 Result

Generated image combines content and style. Output saved as `output_style_transfer.jpg`.

6 Conclusion

This TP provided hands-on experience with CNNs, ResNet architectures, and neural style transfer. The practical implementation and theoretical analysis deepen understanding of modern computer vision techniques.

Annexes

Lien du depot Github : https://github.com/DPW2005/deep_learning.git

Lien du document latex : <https://www.overleaf.com/project/68d392e5a266d8f7159dd874>