BIRZEIT UNIVERSITY

Department of Electrical and Computer Engineering
Second Semester, 2024/2025
Intelligent Systems Lab, ENCS5141

Assignment#2: Comparative Analysis of CNN and Patch-based LSTM Architectures for Image Classification

Objective

To evaluate and compare the performance of **Convolutional Neural Networks (CNNs)** and **Long Short-Term Memory networks (LSTMs)** for image classification, where the LSTM processes **sequential image patches** instead of whole images. This study investigates how temporal modeling of spatial patches compares to spatial feature extraction in image recognition tasks.

Datasets

Test your approaches on both of the following datasets.
- CIFAR-10 (10 classes, 32×32 RGB images)
- MNIST (28×28 grayscale images)

Methodology

1. **Data Preprocessing**
   - Normalize all images to the range [0, 1].
   - For LSTM: divide each image into non-overlapping **patches** (e.g., 4×4 or 8×8).
     - Flatten each patch to form a sequence of vectors.
2. **Architecture A – CNN Baseline**
   - Standard CNN model with 2–3 convolutional layers, followed by pooling, dense layers, and softmax.
   - Example: Conv → ReLU → Pool → Conv → ReLU → Flatten → Dense → Softmax
3. **Architecture B – AlexNet**
   - Use the pre-trained model on ImageNet dataset.
4. **Architecture C – Patch-based LSTM**
   - Each image is divided into sequential patches (e.g., 16 patches from 32×32 images).
   - Each patch is flattened and passed in sequence to an LSTM.
   - Output of the final LSTM cell is connected to dense + softmax layers for classification.
5. **Training**
   - Use the same loss function (CrossEntropy) and optimizer (e.g., Adam).
   - Train all models for the same number of epochs with identical batch sizes.

6. **Evaluation Metrics**
    - Accuracy
    - Training time
    - Inference time
    - Confusion matrix

Submissions:

- You need to submit the code in .ipynb format. You can obtain this file in Google Colab by navigating to the File menu and selecting Download > Download .ipynb.
- Comparative performance plots (training curves, accuracy, confusion matrix).
- Final report summarizing:
    - Methodology and experimental setup
    - Quantitative results
    - Discussion of advantages and disadvantages of each architecture

Important notes:
- Deadline: **Sunday, 8 June 2025** at 11:59 pm. Please submit your solution through Ritaj as a reply to this message.
- Ensure adherence to the report preparation guidelines outlined in the "ENCS5141 Case Study Report Guidelines.pdf" document. If you opt to write the report using LaTeX, utilize the provided report template "ENCS5141 Sample Report.tex".
- Late Submission Policy: One mark (out of the 30 marks assigned to case studies in the course outline) will be deducted for every day of late submission of the case study report. No submissions will be accepted beyond the third day past the due date.