

Assignment-1

Compare performance of 10 CNNs pre-trained on ImageNet dataset using images of 20 classes of CIFAR-100 dataset.

- Students need to submit individual hand-written report via CR on 25.5.2025 (Sunday) before CSE4261 class.
- Code + digital report should be submitted to the Assignment box by scheduled time.

Students may compare models based on accuracy, hardware requirements, size of the model, inference time, architecture etc.

Assignment-2

Write a report:

- describing the effect of different types of activation function of your chosen 10 CNNs
- mentioning at least three CNNs which use regular kernel, deformable kernel, dilated kernel, depthwise separable kernel, modified depthwise-separable kernel, pointwise kernel
- describing different layers' feature map of your chosen CNN

Assignment-3

Discuss feature extraction power of your favorite CNN pretrained by ImageNet dataset before and after transfer learning by MNIST digit dataset by plotting high dimensional feature vectors on 2D plane using three dimension reduction techniques

Assignment-4

.Manually draw a neural network having 3 hidden layers for classifying images of 10 English digits.

- Manually draw Forward Propagation Computation Graph
- Manually draw Backpropagation Computation Graph
- Manually derive equations for updating weights of each hidden layer and output layer.
- Train model using the MNIST digit dataset with the help of Tensorflow's `tf.GradientTape()`.
- Compare the performance of the model trained by `tf.GradientTape()` with Tensorflow's `model.fit()`

Assignment-5

Implement Fast Gradient Signed Method for adversarial attack to your favorite pre-trained CNN. You can take help from:

https://www.tensorflow.org/tutorials/generative/adversarial_fgsm

Choose an image belong to any class supported by ImageNet 1000 classes mentioned in the following link:

<https://deeplearning.cms.waikato.ac.nz/user-guide/class-maps/IMAGENET/>

Discuss whether this kind of attack can be done if you add a noise generated by Gaussian distribution.

Assignment-6

- By the help of Grad-CAM and Integrated Gradients, investigate which area of an adversarial example of an image makes your favorite neural network confused.
- Compare most important area of an adversarial example with the important area of the original image.
- Investigate what will happen if you consider softmax layer for Grad-CAM and the previous-layer of the softmax layer for IG for estimating gradients.

Assignment-7

- Capture a video without any human face using a smart phone and detect object using Ultralytics provided YOLOv8, YOLOv11 and YOLOv12 models. Compare their object detection capability, speed.
- Fine-tune any Ultralytics YOLOv8 model to prepare it as a face detector using WIDER FACE training set by splitting it into training set and validation set.
- Compare fine-tuned YOLOv8 based face detector with face detector available in the following GitHub directory using the validation set of WIDER FACE dataset.
- <https://github.com/Yusepp/YOLOv8-Face>
- Build a face detector having the architecture of YOLOv1, train it using WIDERFACE WIDER FACE training set by splitting it into training set and validation set.

Assignment-8

Write a report after:

- Training and evaluating a U_Net by your favorite dataset publicly available for Segmantion task.
- Training and evaluating a U-Net by your favorite dataset publicly available for Crowd Counting task.
- Training and evaluating an MCNN by your favorite dataset publicly available for Crowd Counting task.
- Comparing MCNN-based and U-Net based crowd counters.

Assignment-9

Write a report based on your findings and analysis after:

- training an autoencoder as a 2D feature generator and displaying CIFAR10 dataset's features
- comparing autoencoder generated features with features extracted by a pre-trained CNN and reduced by dimension reduction techniques like PCA, t-SNE.
- training a denoising autoencoder for CIFAR10 dataset
- training a CNN based CIFAR-10 classifier without any single-image data augmentation techniques
- training a CNN based CIFAR-10 classifier with a single-image data augmentation techniques

Assignment-10

Write a report after:

- Generating five images passing five noise vectors/matrices drawn from a normal distribution with mean 5 and variance through the Decoder of your normal Autoencoder.
- Generating five images passing five noise vectors/matrices drawn from a normal distribution with mean 5 and variance through the Decoder of your Denoising Autoencoder.
- Training and evaluating a Variational Autoencoder
- Generating five images passing five noise vectors/matrices drawn from a normal distribution with mean 5 and variance through the Decoder of your Variational Autoencoder.

Assignment-11

Write a report:

- comparing three face verifiers which are trained using Binary Cross-Entropy loss, contrastive loss and triplet loss, receptively.
- comparing effect of Binary Cross-Entropy (BCE) loss and Mean Squared Error (MSE) as a reconstruction loss during training of a Variational Autoencoder (VAE).

Assignment-12

Write a report by comparing performances of four 10 class classifiers after:

- building a small CNN based classifier, and training it by your favorite dataset (except the MNIST digit or fashion dataset) having images of 10 classes.
- preparing your two favorite big CNNs, pre-trained by ImageNet dataset for 1000 classes, for classifying 10 classes of your chosen dataset by fine tuning last 1/2 layers
- transferring knowledge from your big fine-tuned one classifier to your own small CNN
- transferring knowledge from your big fine-tuned two classifiers to your own small CNN

Assignment-13

Write a report based on your findings after

- generating synthetic faces using Deep Convolutional GAN
- generating synthetic faces with long hair using Conditional GAN
- generating painted form of faces using CycleGAN
- figuring out the reason behind using the BCE loss instead of the MiniMax loss in Tensorflow implementation (mathematical derivation).

Assignment-14

Write a report after:

- training a ViT based classifier for 20 classes of ImageNet dataset
- comparing ViT based classifier with a fully connected feed forward (FCFNN) network and CNN based classifier for the same dataset
- showing the effect of the number of heads on the performance of ViT
- showing the effect of the choice of patch embedding on the performance of ViT
- showing the effect of the choice of positional embedding on the performance of ViT
- discussing the difficulties of handling ViT over FCFNN and CNN

Assignment-15

Write a report after building a BERT model from scratch, training it with the masked language modeling task, and then fine-tune this model for a:

- sentiment classification task using
- question-answering task using SQuAD (Stanford Question-Answering) Dataset
- semantic similarity to determine how similar two sentences are using SNLI (Stanford Natural Language Inference) Dataset

You can take help from

- https://keras.io/examples/nlp/masked_language_modeling/
- https://keras.io/examples/nlp/text_extraction_with_bert/
- https://keras.io/examples/nlp/semantic_similarity_with_bert/

Compare your BERT's performance with publicly available BERT model which can be run on your available GPU.

For your own BERT model, you may think to tune the embedding's dimension, number of encoder blocks.

TheoryAssignment-1

Write your opinion about whether you want to have a system with artificial intelligence as:

- a teacher for solving your Homework-Problems/Assignments and helping you in coding.
- an assistant at your working place
- an assistant at your home which will help you handling household chores such as regular cleaning, shopping even deciding your food menu analyzing your financial status, desire and many other things.
- a companion for the rest of your life with whom you can share your every thought without any hesitation in your mother language, who will understand your emotional and psychological status, who will be a great listener and an advisor, who will be aware of your health and so on.
- a boss at your working place

Write down the reasons why you want or do not want any or all of the above mentioned AI-based systems in your life.

Do you think you will be too much dependent on any of the above mentioned system within 20 years? Justify your opinion.

Note: This Assignment + ClassTest will be graded for Continuous Assessment of Theory Class.
Submission Deadline: Two days before your theory exam

