L1 – introduction + history of Computer Vision

L2 – image classification talking about popular datasets like :

-MNIST

-CIFAR10

-ImageNet

-MIT Places

-Omniglot

-k-NN Algorithm

-train: prepare data set (Download + Preprocess)

-predict: let’s talk about image classification: u take each Test image and compute the Distance(Euclidian/Manhattan) between test and each image from the training set, sort it descendently and pick the top k distances and choose the predicted label via the most frequency class label.

-also use cross validation folding to choose the “best k”

- in case of a tie choose a strategy (e.g. pick the smallest label)

-curse of dimensionality

-distances between points in high dimensions are increasing exponentially

-computation inneficient(calculating distances between all points)

L3 – linear classifier(SVM)

-bias trick(not used in Neural Networks) – getting rid of the bias vector by concatenating it to the Weight matrices and adding a “1” to the X features Vector

-SVM (finding a hyperplan that best divide each class in sections)

-talking about CIFAR10 data set (50000 images) of (32,32,3) images of 10 classes(cat,car,frog,…)

- z = W \* X

W shape: (10,3072)

10 – 10 classes

3072 – image flattened(32 \* 32 \* 3 pixels)

X shape : (3072,1)

z Shape: (10,1) – 10 predicted scores for each class

-Visual viewpoint of the classifier: the classifier tries to understand the template (shapes and colors of the background).

-ex: If it learns that most photos of horses are on a green backround and the horse is usually brown and tries to predict a car(not necessarily brown on a greenish backround it will most likely fail)

-Geometric point of the classifier: each row of the W matrix defines a hyperplan and the prediction consists in which hyperplan does the unseen image belong to.

-Loss functions: SVM loss – Multiclass SVM loss function – one vs all Approach

A black background with white text

AI-generated content may be incorrect.

Sj – score of each incorrect class

Syi – score of the correct class

Δ – 1

Total loss:

A black background with white symbols

AI-generated content may be incorrect.

Λ – lambda (L2 regularization term) prevents overfitting of data

Cross-Entropy Loss – interpret classifier scores as probabilities using SoftMax function

A black background with white text

AI-generated content may be incorrect.

L4

Batch Gradient Descent – determine the gradient by computing the derivative of the whole training set

-can be inneficient when having lots of data with lots of parameters to learn

-hyperparameters:

-learning rate

-number of steps(epochs)

-Weight initialiatization(currently important?!)

A close-up of a computer code

AI-generated content may be incorrect.

Stochastic Gradient Descent(SGD) – determining gradient by computing the derivative of minibatches of training set

-hyperparameters:

-learning rate

-number of steps(epochs)

-Weight initialiatization(currently important?!)

-Batch size(32/64/128)

-Data Sampling(not that important of image classification)

A close-up of a computer code

AI-generated content may be incorrect.

-choosing the learning rate is important because:

- if choosing a larger learning rate the model can overshoot over the globabl minimum and might never reach it

- if choosing a smaller learning rate the model can be slow and the number of epochs to converge could be so big that the computation power will not handle it

SGD + Momentum:

A screenshot of a computer

AI-generated content may be incorrect.

Adam – combines Momentum + Adaptive Learning Rate for faster convergence