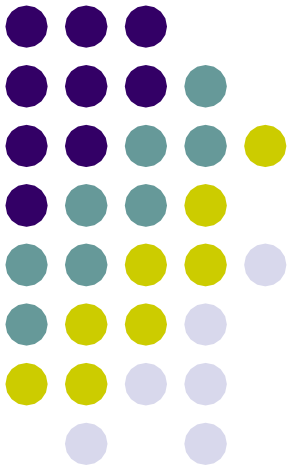
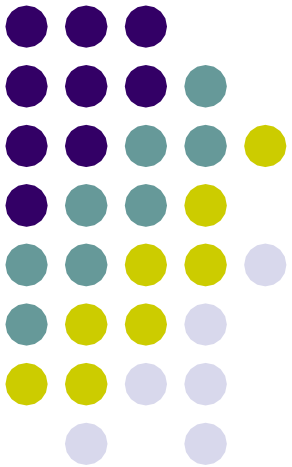
Introduction to Artificial Intelligence

Prof. Inkyu Moon

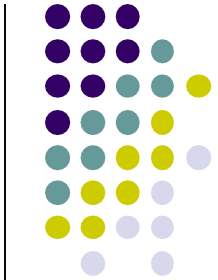
Dept. of Robotics Engineering, DGIST



CNN Designs (Assignments)



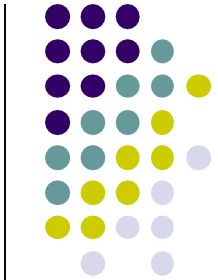
#1 Assignment



The Dogs and Cats Dataset

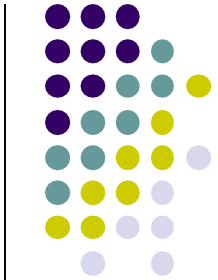
- This dataset contains 25,000 images of dogs and cats (12,500 from each class).
- After downloading and uncompressing it, we create a new dataset containing three subsets: **a training set with 500 samples of each class, a validation set with 250 samples of each class, and a test set with 250 samples of each class.**

#1 Assignment



- **We have 1,000 training images, 500 validation images, and 500 test images**.
- You should train your own CNN with 1,000 training images, evaluate the trained network with the validation dataset, and finally test the trained, evaluated network with test dataset.
- **1) Evaluate your CNN performance by varying the total number of epoch.**
- 2) Improve your CNN performance using the data augmentation and show the performance by varying the total number of epoch.
- 3) Improve your CNN performance using the pretrained convnet and show the performance by varying the total number of epoch.

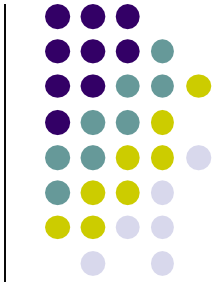
#2 Assignment



The MNIST Dataset of Handwritten Numbers

- There is a collection of images of handwritten numbers used by artificial intelligence researchers as a popular set to test their ideas and algorithms.
- Many different ideas and algorithms are tested against the same data set.
- **That data set is called the MNIST database** and is available from the neural network researcher **Yann LeCun's website http://yann.lecun.com/exdb/mnist/.**

#2 Assignment



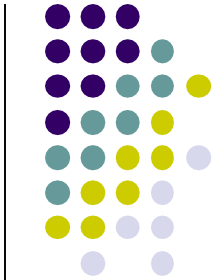
- **This website provides two CSV files**:

  A **training** set http://www.pjreddie.com/media/files/mnist_train.csv

  A **test** set http://www.pjreddie.com/media/files/mnist_test.csv

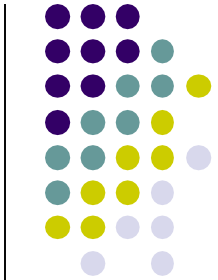
- **The training set is the set of 60,000 labelled examples used to train the CNN**: labelled means the inputs come with the desired output.
- **The smaller test set of 10,000** is used to see how well our idea or algorithm works.

#2 Assignment



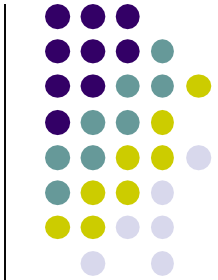
- After downloading it, we create a new dataset containing three subsets: a training set with 500 samples of each class, a validation set with 100 samples of each class, and a test set with 100 samples of each class.
- **We have 5,000 training images, 1,000 validation images, and 1,000 test images.**

#2 Assignment



- You should train your own CNN with 1,000 training images, evaluate the trained network with the validation dataset, and finally test the trained, evaluated network with test dataset.
- **1) Evaluate your CNN performance by varying the total number of epoch.**
- 2) Improve your CNN performance using the data augmentation and show the performance by varying the total number of epoch.
- 3) Improve your CNN performance using the pretrained convnet and show the performance by varying the total number of epoch.

#2 Assignment



Creating New Training Data: Rotations

- Create new your handwritten number images by rotating them clockwise or anticlockwise, by 10 degrees for example; you can create many more examples with different rotation angles.