

Chuyên Đề Công Nghệ Convolutional Neural Networks

Giảng viên: TS. Nguyễn Thị Ngọc Diệp

Email: ngocdiep@vnu.edu.vn

Slide & Code: https://github.com/chupibk/INT3414_22

Giới thiệu môn học

Tuần 1

Big questions

WHAT: CNN là gì?

WHERE: Vị trí của CNN trong KHMT?

WHY: Tại sao học CNN?

WHEN: CNN có từ bao giờ
và đang đi tới đâu?

HOW: CNN hoạt động như thế nào?

Q1- WHAT

- Convolutional networks are simply neural networks that use convolution in place of general matrix multiplication in at least one of their layers
 - Convolution is a mathematical operation having a linear form

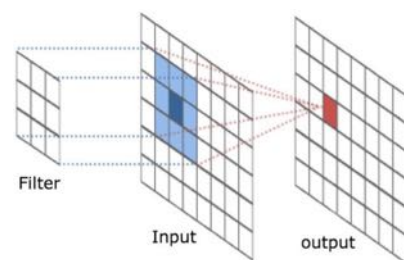


Image credit: internet

Q2- WHY

- Equip yourself with one of the most powerful and promising tools to solve various computer science tasks
- CNNs form the backbone of many modern AI systems

Q3- WHEN

- 1990s: LeCun et al. with MNIST



Graph Transformer Networks:
-> Gradient-based learning technique

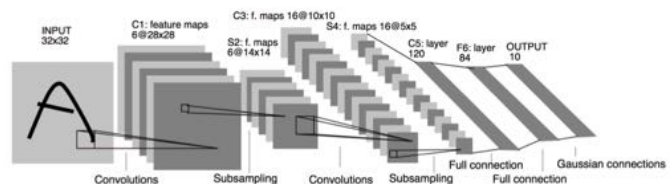


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

Ref: LeCun, Yann, et al. "Gradient-based learning applied to document recognition." *Proceedings of the IEEE* 86.11 (1998): 2278-2324.

Q3 - WHEN

- 1990s – early 2000s: further work on CNN model
- 2012: a huge surge in popularity after AlexNet
 - Achieved state-of-the-art in ImageNet LSVRC-2010 challenge (1000 classes)
 - Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.

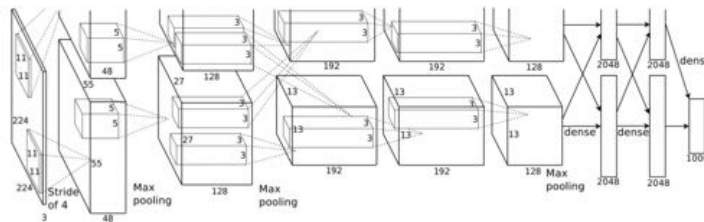
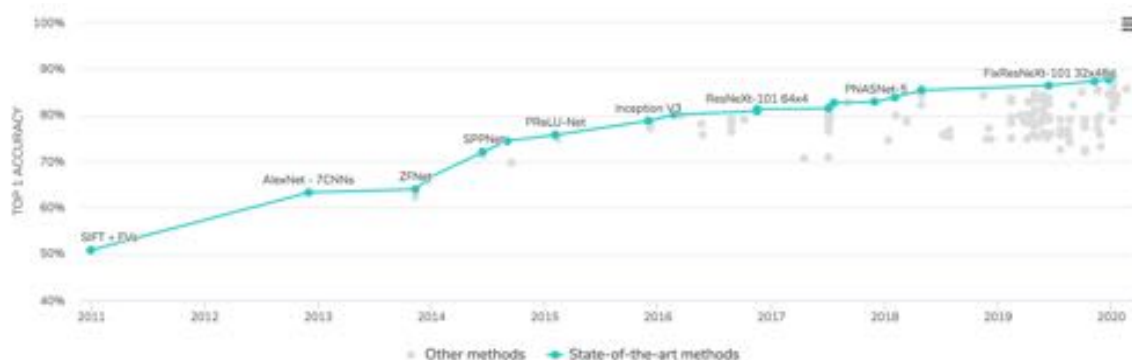


Image classification on ImageNet



Ref: <https://paperswithcode.com/sota/image-classification-on-imagenet>

Q4- WHERE

CNNs are everywhere

- Image retrieval
- Image classification
- Object detection
- Self-driving cars
- Semantic segmentation
- Face recognition
- Pose estimation
- Detect diseases
- Speech recognition
- Text processing
- Analyzing satellite data
- Music recommendation (e.g., Spotify)
- Photo Geolocation (Google's PlaNet)
- Etc...

Q5 – HOW: CNN hoạt động như thế nào?

- → là nội dung của môn học này

Mục tiêu môn học

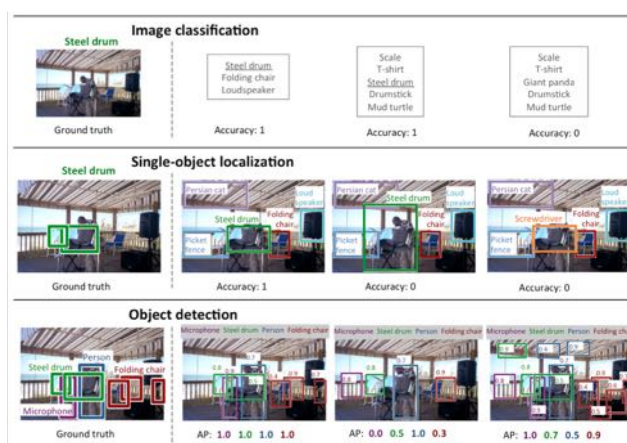
- Giới thiệu các mô hình học sâu end-to-end dùng trong thị giác máy
 - Image classification, Object detection, Segmentation
- Trang bị cho sinh viên khả năng cài đặt, huấn luyện và debug các mạng nơ-ron

Schedule

Week	Content	Class hour	Self-study hour
1	Introduction CNNs in Computer Vision	2	1
2	Foundations of CNNs Case study: Image classification problem	2	2-6
3	Training and tuning parameters Data augmentation - Data generator Model evaluation - Transfer learning	2	2-6
4	Object detection	2	2-6
5	Segmentation	2	2-6
6, 7	Mid-term presentations	2	2-6
8, 9	Advanced topics using CNNs	2	2-6
10, 11, 12	Final project presentations	1	2-6
13	Class summarization	1-3	open

Core tasks of Computer vision

Classification – Localization – Detection – Segmentation



Semantic segmentation



Instance segmentation

Ref: Russakovsky, Olga, et al. "Imagenet large scale visual recognition challenge." *International journal of computer vision* 115.3 (2015): 211-252.

Core tasks

Core CV Task	Task Description	Output	Metrics
Classification	Given an image, assign a label	Class Label	Accuracy
Localization	Determine the bounding box containing the object in the given image	Box given by (x1, y1, x2, y2)	Ratio of intersection to the union (Overlap) between the ground truth and bounding box
Object Detection	Given an image, detect all the objects and their locations in the image	For each object: (Label, Box)	Mean Avg Best Overlap (MABO,) mean Average Precision (mAP)
Semantic Segmentation	Given an image, assign each pixel to a class label, so that we can look at the image as a set of labelled segments	A set of image segments	Classification metrics, Intersection by Union overlap
Instance Segmentation	Same as semantic segmentation, but each instance of a segment class is determined uniquely	A set of image segments	

Classification + Localization task

Classification: C classes

Input: Image

Output: Class label

Evaluation metric: Accuracy



→ CAT

Localization:

Input: Image

Output: Box in the image (x, y, w, h)

Evaluation metric: Intersection over Union

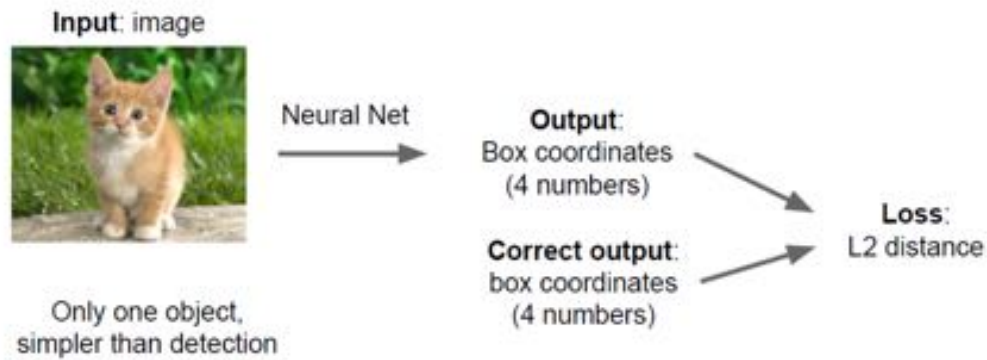


→ (x, y, w, h)

Classification + Localization: Do both

How to do localization?

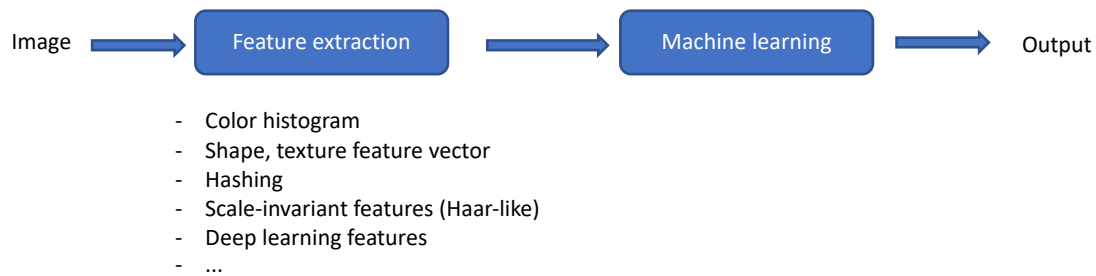
Idea: Localization as regression



Slide Credit: A Karpathy, CS231n

Học sâu vs học máy

Phương pháp cổ điển trong thị giác máy



Phương pháp học sâu: end-to-end

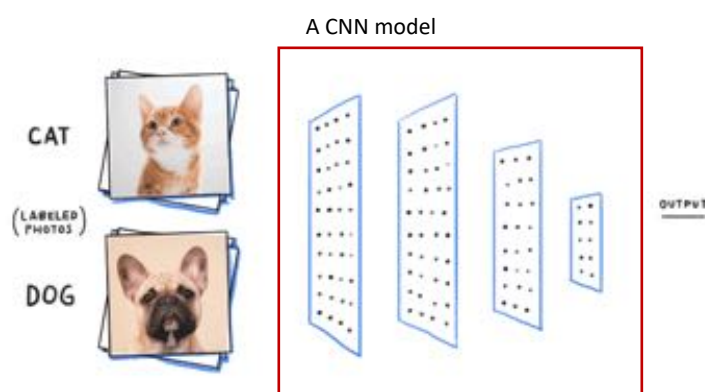


Image credit: <https://becominghuman.ai/building-an-image-classifier-using-deep-learning-in-python-totally-from-a-beginners-perspective-be8dbaf22dd8>

A quick demonstration

<https://playground.tensorflow.org/>



Some applications

Image Classification: grayscale

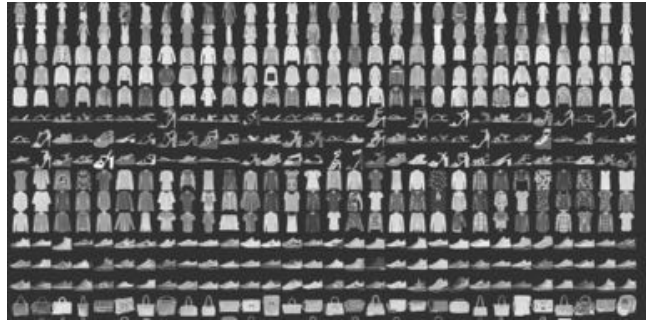
MNIST (LeCun et al., 1998)



10 classes

<http://yann.lecun.com/exdb/mnist/>

Fashion-MNIST (2017)



10 classes

<https://github.com/zalandoresearch/fashion-mnist>

Eyeglass detection



MeGlass, 2018

<https://github.com/cleardusk/MeGlass>

DiepNg 2020 – INT3414_22

Dog breed identification

120 breeds

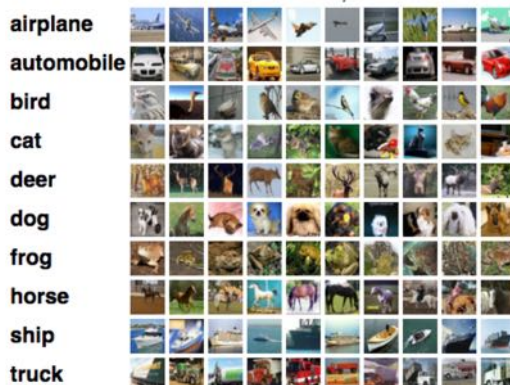


<https://www.kaggle.com/c/dog-breed-identification>

DiepNg 2020 – INT3414_22

Image classification: tiny images

CIFAR-10 (Krizhevsky and Hinton, 2009)
32x32 size



CIFAR-100 (Krizhevsky and Hinton, 2009)

Superclass

aquatic mammals
fish
flowers
food containers
fruit and vegetables
household electrical devices
household furniture
insects
large carnivores
large man-made outdoor things
large natural outdoor scenes
large omnivores and herbivores
medium-sized mammals
non-insect invertebrates
people
reptiles
small mammals
trees
vehicles 1
vehicles 2

Classes

beaver, dolphin, otter, seal, whale
aquarium fish, flatfish, ray, shark, trout
orchids, poppies, roses, sunflowers, tulips
bottles, bowls, cans, cups, plates
apples, mushrooms, oranges, pears, sweet peppers
clock, computer keyboard, lamp, telephone, television
bed, chair, couch, table, wardrobe
bee, beetle, butterfly, caterpillar, cockroach
bear, leopard, lion, tiger, wolf
bridge, castle, house, road, skyscraper
cloud, forest, mountain, plain, sea
camel, cattle, chimpanzee, elephant, kangaroo
fox, porcupine, possum, raccoon, skunk
crab, lobster, snail, spider, worm
baby, boy, girl, man, woman
crocodile, dinosaur, lizard, snake, turtle
hamster, mouse, rabbit, shrew, squirrel
maple, oak, palm, pine, willow
bicycle, bus, motorcycle, pickup truck, train
lawn-mower, rocket, streetcar, tank, tractor

<https://www.cs.toronto.edu/~kriz/cifar.html>

Image classification: large data

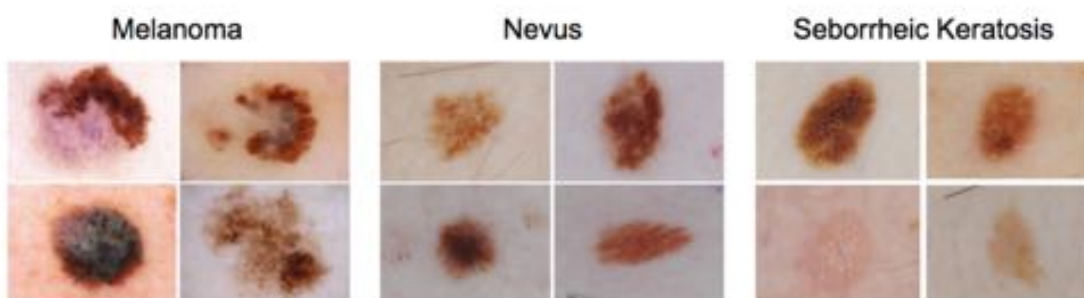
ImageNet (1000 classes) - Deng et al., 2009



<http://www.image-net.org/>

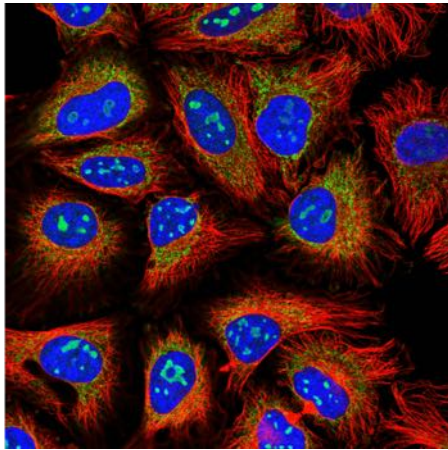
Image classification: medical image

ISIC2017, Skin lesion analysis toward melanoma detection



<https://challenge.kitware.com/#phase/5840f53ccad3a51cc66c8dab>

Human protein atlas image classification



28 classes:

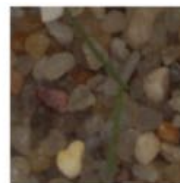
- | | |
|------------------------------|-----------------------------------|
| 0. Nucleoplasm | 14. Microtubules |
| 1. Nuclear membrane | 15. Microtubule ends |
| 2. Nucleoli | 16. Cytokinetic bridge |
| 3. Nucleoli fibrillar center | 17. Mitotic spindle |
| 4. Nuclear speckles | 18. Microtubule organizing center |
| 5. Nuclear bodies | 19. Centrosome |
| 6. Endoplasmic reticulum | 20. Lipid droplets |
| 7. Golgi apparatus | 21. Plasma membrane |
| 8. Peroxisomes | 22. Cell junctions |
| 9. Endosomes | 23. Mitochondria |
| 10. Lysosomes | 24. Aggresome |
| 11. Intermediate filaments | 25. Cytosol |
| 12. Actin filaments | 26. Cytoplasmic bodies |
| 13. Focal adhesion sites | 27. Rods & rings |

<https://www.kaggle.com/c/human-protein-atlas-image-classification/overview>

Plant seedlings classification

Differentiate a weed from a crop seedling
960 plants, 12 species

1. Black-grass
2. Charlock
3. Cleavers
4. Common Chickweed
5. Common wheat
6. Fat Hen
7. Loose Silky-bent
8. Maize
9. Scentless Mayweed
10. Shepherds Purse
11. Small-flowered Cranesbill
12. Sugar beet



<https://www.kaggle.com/c/plant-seedlings-classification/data>

PlantDisease classification



38 classes

<https://www.crowdai.org/challenges/1>

Image classification: satellite images

DeepSat SAT-4, four classes: barren land, trees, grassland, all land → red, green, blue and near infrared bands



<https://www.kaggle.com/crawford/deepsat-sat4>

And many others...


















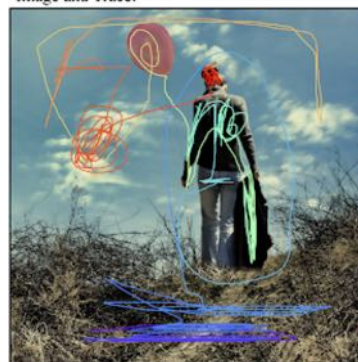
All Categories - image classification			Prize: \$25,000		INT3414_22			
21 Competitions								
	Recursion Cellular Image Classification CellSignal: Characterizing biological signal from experimental noise in cellular images Released: 1 month to go - % research settings, classification, image data	\$13,000 858 teams		CofeCount's Image Classification Challenge Categories: e-commerce photos Released: 2 years ago - % multiclass classification	\$35,000 827 teams		Human Protein Atlas Image Classification Classify subcellular protein patterns in human cells Released: 7 months ago - % classification, image data	\$37,000 2,189 teams
	Materialist Challenge (Fashion) at FGVC5 Image classification of fashion products Released: 4 years ago	\$2,500 212 teams		Materialist Challenge (Furniture) at FGVC5 Image Classification of Furniture & Home Goods Released: 4 years ago	\$2,500 428 teams		Digit Recognizer Learn computer vision fundamentals with the famous MNIST data Testing Started: Ongoing - % federate data, image data, multiclass classification, object identification	Knowledge 2,860 teams
	Kuzushiji Recognition Opening the door to a thousand years of Japanese culture Released: 5 months ago - % history, image data, multiclass classification, paper	\$15,000 124 teams		APTOS 2019 Blindness Detection Detect diabetic retinopathy in eye fundus before it's too late Released: Code Competition - 10 days to go - % healthcare, medicine, image data, multiclass classification	\$10,000 2,810 teams		IWildCam 2019 - FGVC6 Categorize animals in the wild Released: 10 months ago - % image data, multiclass classification	Kudros 338 teams
	Inclusive Images Challenge Stress test image classifiers across new geographic distributions Released: 8 months ago - % image data, multiclass classification	\$25,000 468 teams		dstl Satellite Imagery Feature Detection Can you train an eye in the sky? Released: 2 years ago - % image data, multiclass classification, object segmentation	\$100,000 419 teams		Leaf Classification Can you see the random forest for the leaves? Released: 2 years ago - % image data, multiclass classification, object identification	Knowledge 1,098 teams
	National Data Science Bowl Predict weapon health, and planckton at a time Released: 4 years ago - % geospatiality, image data, multiclass classification	\$175,000 1,049 teams		Data Science Bowl 2017 Can you improve lung cancer detection? Released: 2 years ago - % healthcare, image data, binary classification	\$1,000,000 1,972 teams		Plant Seedlings Classification Determine the species of a seedling from an image Released: 1 year ago - % plants, image data, multiclass classification	Kudros 834 teams
	Diabetic Retinopathy Detection Identify signs of diabetic retinopathy in eye images Released: 4 years ago - % symptoms, health sciences, image data, binary classification	\$100,000 661 teams		Dog Breed Identification Determine the breed of a dog in an image Released: 4 years ago - % animals, image data, multiclass classification, object identification	Kudros 1,285 teams		Dogs vs. Cats Redux: Kernel's Edition Disagregate images of dogs from cats Released: 2 years ago - % animals, image data, binary classification, object identification	Knowledge 1,314 teams
	Statoli/C-CORE Iceberg Classifier Challenge Ship or iceberg, can you decide from space? Released: 2 years ago - % weather, shipping, image data, binary classification	\$50,000 3,341 teams		The Nature Conservancy Fisheries Monitoring Can you detect and classify species of fish? Released: 2 years ago - % fishing, image data, multiclass classification, object detection	\$150,000 2,293 teams		Intel & MobileODT Cervical Cancer Screening Whose cancer treatment will be most effective? Released: 2 years ago - % healthcare, image data, multiclass classification, object identification	\$100,000 848 teams

Image captioning

Image and Trace:



Caption:

In the front portion of the picture we can see a dried grass area with dried twigs. There is a woman standing wearing light blue jeans and ash colour long sleeve length shirt. This woman is holding a black jacket in her hand. On the other hand she is holding a balloon which is peach in colour. On the top of the picture we see a clear blue sky with clouds. The hair colour of the woman is brownish.

Voice:



Source: <https://ai.googleblog.com/>

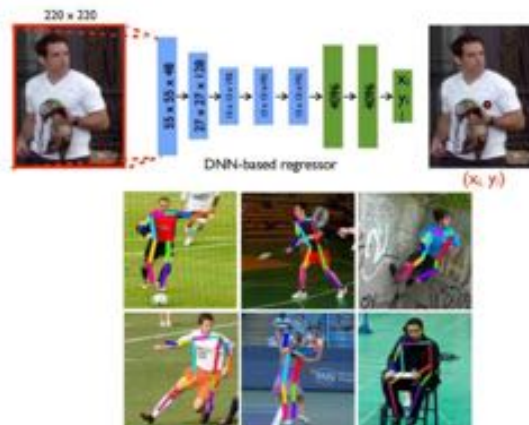
Human Pose Estimation

Represent a person by K joints

Regress (x, y) for each joint from last fully-connected layer of AlexNet

(Details: Normalized coordinates, iterative refinement)

Toshev and Szegedy, "DeepPose: Human Pose Estimation via Deep Neural Networks", CVPR 2014



Object detection

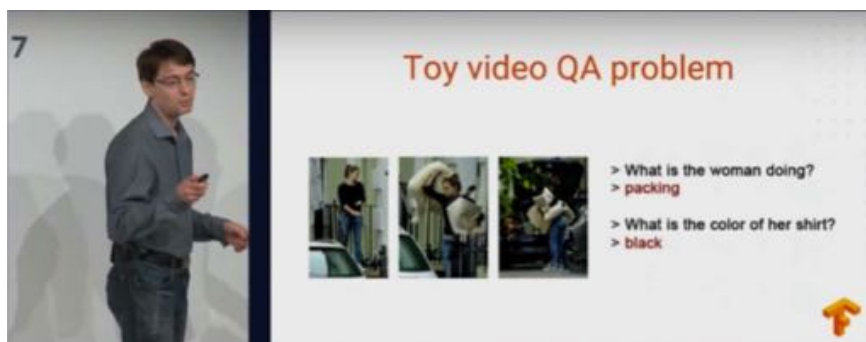


Image Q&A



<https://arxiv.org/pdf/1612.00837.pdf>

Video Q&A



<https://www.youtube.com>

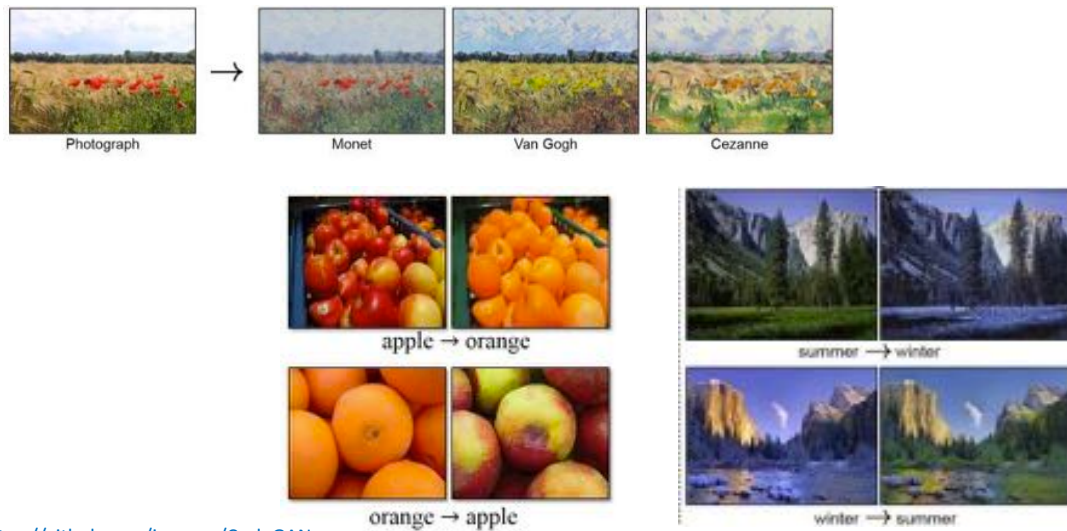
Image-to-Image Translation



Image generation



Style transfer



Thông tin về môn học

Chính sách đối với môn học

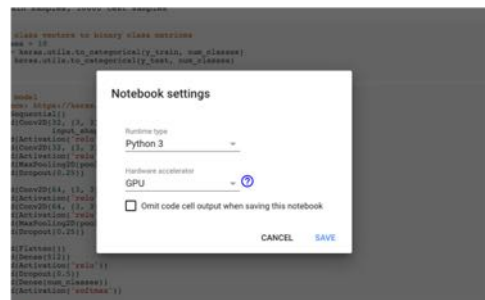
- Sinh viên nghỉ quá 20% số buổi học lý thuyết (3 buổi học) sẽ không được thi cuối kỳ
 - ~~Mỗi buổi học sẽ có điểm danh~~
- Sinh viên tích cực làm bài tập, tham gia thảo luận, trả lời câu hỏi sẽ được xem xét cộng điểm
- Với các nội dung liên quan tới bài tập giữa kỳ và đồ án môn học nếu sinh viên gian lận thì sẽ bị điểm môn học là 0

Trọng số điểm

Hình thức	Phương pháp	Mục đích	Trọng số
Đồ án môn học	Dự án nhỏ làm việc theo nhóm	Đánh giá kỹ năng lập trình, xây dựng hệ thống dịch vận dụng kiến thức đã học	50%
Chuyên cần	Điểm danh	Đánh giá tính chuyên cần của sinh viên	+
Bài tập giữa kỳ	Nộp bài luận	Kiểm tra khả năng tự học của sinh viên	50%
Tổng			100%

Thư viện và ngôn ngữ lập trình sử dụng

- Ngôn ngữ: Python
- Thư viện chính: Pytorch
- Tuy nhiên sinh viên có thể lựa chọn các thư viện mạng học sâu khác
 - Keras, Tensorflow, Theano, v.v...
- Khuyến khích
 - Sử dụng Google Colab
 - Lập trình với Jupyter Notebook



Tài liệu tham khảo

- CS231n Convolutional Neural Networks for Visual Recognition
 - <http://cs231n.github.io/>
- Jeremy Howard & Rachel Thomas
 - <http://course.fast.ai>
- Deeplearning.ai Course 4
 - https://www.youtube.com/watch?v=ArPaAX_PhIs&list=PLkDaE6sCZn6GI29AoE31iwdVwSG-KnDzF
 - Or <https://www.coursera.org/learn/convolutional-neural-networks>
- <https://pytorch.org/tutorials/>