4-6 Wednesday – 210-GD3

Special topics in Computer Science INT3121 20

Lecturer: Nguyen Thi Ngoc Diep, Ph.D.

Email: ngocdiep@vnu.edu.vn

Slide & Code: https://github.com/chupibk/INT3121-20

Image Classification with Convolutional Neural Networks

Week 1: Introduction

Why is image classification important?

- Classification is significant in order to understand characteristics of groups
- Image classification is essential in other important tasks of computer vision → the core task in CV
  - Semantic/Instance segmentation
  - Object detection
- Large variety of practical applications

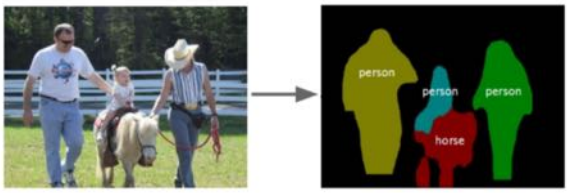


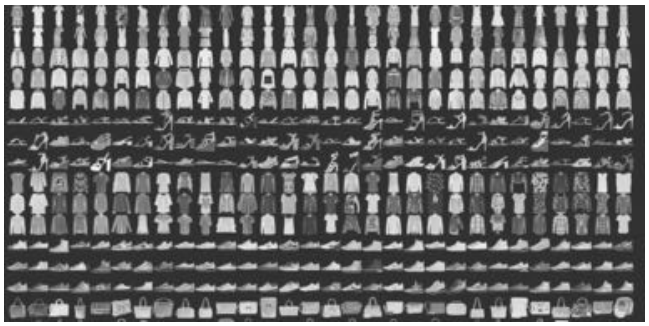
Image Classification: grayscale

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

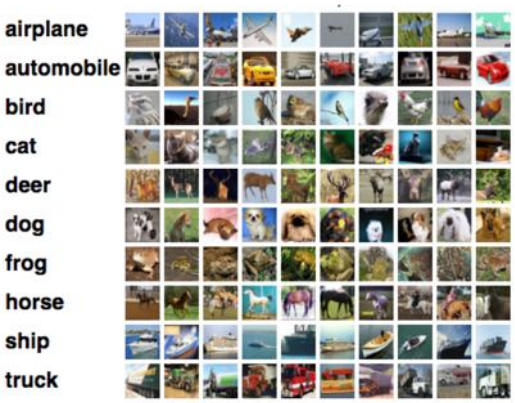
Dog breed identification

120 breeds

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

Image classification: tiny images

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



CIFAR-100 (Krizhevsky and Hinton, 2009)

Superclass	Classes
aquatic mammals	beaver, dolphin, otter, seal, whale
fish	aquarium fish, flatfish, ray, shark, trout
flowers	orchids, poppies, roses, sunflowers, tulips
food containers	bottles, bowls, cans, cups, plates
fruit and vegetables	apples, mushrooms, oranges, pears, sweet peppers
household electrical devices	clock, computer keyboard, lamp, telephone, television
household furniture	bed, chair, couch, table, wardrobe
insects	bee, beetle, butterfly, caterpillar, cockroach
large carnivores	bear, leopard, lion, tiger, wolf
large man-made outdoor things	bridge, castle, house, road, skyscraper
large natural outdoor scenes	cloud, forest, mountain, plain, sea
large omnivores and herbivores	camel, cattle, chimpanzee, elephant, kangaroo
medium-sized mammals	fox, porcupine, possum, raccoon, skunk
non-insect invertebrates	crab, lobster, snail, spider, worm
people	baby, boy, girl, man, woman
reptiles	crocodile, dinosaur, lizard, snake, turtle
small mammals	hamster, mouse, rabbit, shrew, squirrel
trees	maple, oak, palm, pine, willow
vehicles 1	bicycle, bus, motorcycle, pickup truck, train
vehicles 2	lawn-mower, rocket, streetcar, tank, tractor

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

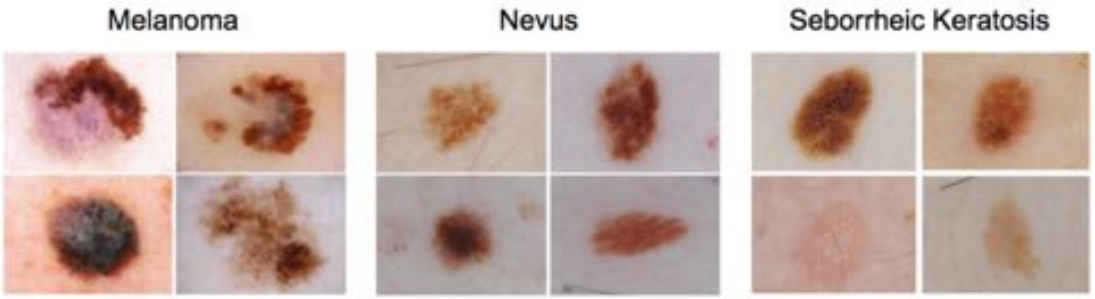
Image classification: large data

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

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

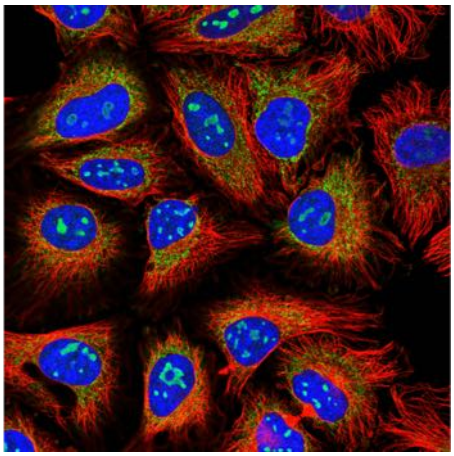
Image classification: medical image

ISIC2017, Skin lesion analysis toward melanoma detection



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

Human protein atlas image classification



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

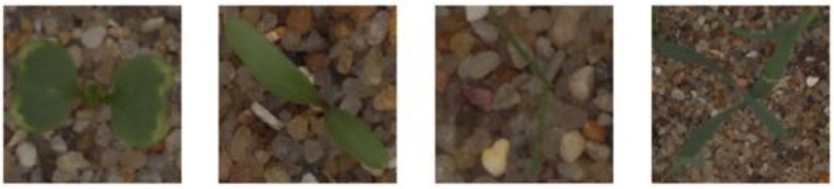
28 classes:

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

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

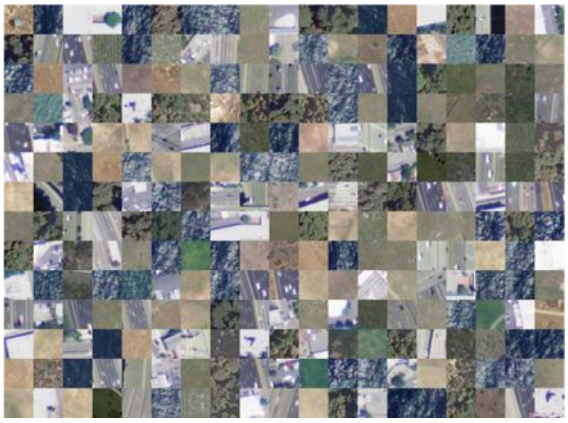
PlantDisease classification

38 classes

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

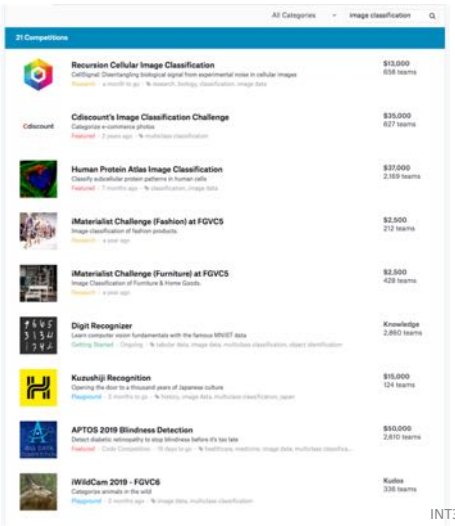
Image classification: satellite images

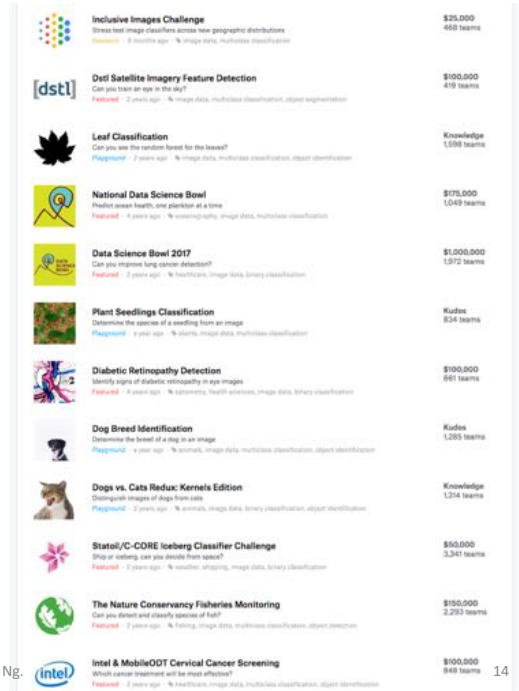
DeepSat SAT-4, four classes: barren land, trees, grassland, all land → red, green, blue and near infrared bands



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

And many others...





Problem statement

Definition of Image classification problem

- Image classification is a discriminant function of mapping images to symbols
- Single-output classification:
  - An image I is mapped to one symbol $c_i \in S = \{c_1, c_2, \dots, c_L\}$ and $|S| \geq 2$
  - If the target symbol set has two elements ($|S| = 2$), we call, "binary classification"
  - If $|S| > 2$, we call "multi-class classification"
- Multi-output classification:
  - An image I is mapped to multiple symbols, a vector of symbol $v = \{c_{s1}, c_{s2}, \dots, c_{sm}\}$ $where$ $c_{si} \in S_i = \{c_1^i, \dots, c_{L_i}^i\}$

Approaches

A (very) brief history

Classical approaches: pipeline

- Color histogram
- Shape, texture feature vector
- Hashing
- Scale-invariant features (Haar-like)
- Deep learning features
- ...

- Supervised:
  - SVM
  - K-Nearest neighbour
  - Maximum likelihood
  - Adaboost
  - XGBoost
- Unsupervised:
  - K-means
  - LDA
  - ...

Modern approaches: end-to-end

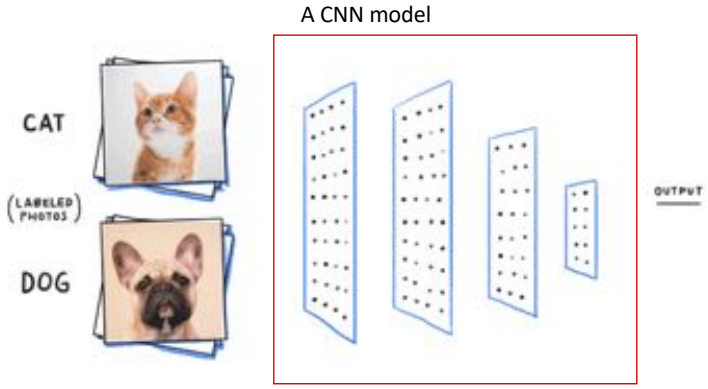


Image credit: https://becominghuman.ai/building-an-image-classifier-using-deep-learning-in-python-totally-from-a-beginners-perspective-be8dbaf22dd8

Arise of CNNs

"Transformers: Generation 1" (1984~)

- Kunihiko Fukushima [1980]: Neocognitron
  - a self-organizing artificial network of simple and complex cells that could recognize patterns and was unaffected by position shifts
- Yann Lecun [1989~]: LeNet-5
  - A backprop style learning algorithm to Fukushima's CNN
  - MNIST dataset (1998)
- Pascal VOC project [2006]: annual competition from 2006-2012
  - 20 object categories
- ImageNet benchmark [2009]: Large scale visual recognition competition (ILSVRC)
  - Over a million images, 1000 of object classes
  - AlexNet [2012] -> breakthrough moments for CNNs
    - 2010-2011: ILSVRC error rate hovered around 26%
    - AlexNet, LeNet-5: 16.4%
    - Current state-of-the-art: 13.6%

ImageNet leading board

Rank	Method	Top 1 Accuracy	Top 5 Accuracy	Number of params	Extra Training Data	Paper Title	Year	Paper	Code
1	FixResNeXt-101 32x48d	86.4%	98.0%	829M	✓	Fixing the train-test resolution discrepancy	2019		
2	ResNeXt-101 32x48d	85.4%	97.6%	829M	✓	Exploring the Limits of Weakly Supervised Pretraining	2018		
3	ResNeXt-101 32x32d	85.1%	97.5%	466M	✓	Exploring the Limits of Weakly Supervised Pretraining	2018		
4	EfficientNet-B7	84.4%	97.1%	66M	×	EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks	2019		
5	GPIPE	84.3%	97%	557M	×	GPipe: Efficient Training of Giant Neural Networks using Pipeline Parallelism	2018		
6	ResNeXt-101 32x16d	84.2%	97.2%	194M	✓	Exploring the Limits of Weakly Supervised Pretraining	2018		
7	EfficientNet-B6	84.0%	96.9%	43M	×	EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks	2019		
8	AmoebaNet-A	83.9%	96.6%	469M	×	Regularized Evolution for Image Classifier Architecture Search	2018		
9	FixPNASNet-5	83.7%	96.8%	86.1M	×	Fixing the train-test resolution discrepancy	2019		
10	MultiGrain PNASNet @ 500px	83.6%	96.7%	86M	×	MultiGrain: a unified image embedding for classes and instances	2019		
11	EfficientNet-B5	83.3%	96.7%	30M	×	EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks	2019		
12	MultiGrain SENet154 @ 400px	83.0%	96.5%	86M	×	MultiGrain: a unified image embedding for classes and instances	2019		
13	Oct-ResNet-152 + SE	82.9%	96.3%	67M	×	Drop an Octave: Reducing Spatial Redundancy in Convolutional Neural Networks with Octave Convolution	2019		
14	PNASNet-5	82.9%	96.2%	86.1M	×	Progressive Neural Architecture Search	2017		
15	NASNET-A(6)	82.7%	96.2%	89M	×	Learning Transferable Architectures for Scalable Image Recognition	2017		
16	SENet-154	82.7%	96.2%	146M	×	Squeeze-and-Excitation Networks	2017		
17	EfficientNet-B4	82.6%	96.3%	19M	×	EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks	2019		
18	FixResNet-50 Billion	82.5%	96.6%	25.6M	✓	Fixing the train-test resolution discrepancy	2019		
19	DPN-131	81.45%	95.84	80M	×	Dual Path Networks	2017		
20	PolyNet	81.3%	95.8%	92M	×	PolyNet: A Pursuit of Structural Diversity in Very Deep Networks	2016		
21	DPN-98	81.1%	95.6%		×	Dual Path Networks	2017		
22	EfficientNet-B3	81.1%	95.5%	12M	×	EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks	2019		
23	ResNeXt-101 64x4	80.9%	95.6%	83.6M	×	Aggregated Residual Transformations for Deep Neural Networks	2016		
24	DPN-92	80.7%	95.3%		×	Dual Path Networks	2017		
25	Inception ResNet V2	80.1%	95.1%	55.8M	×	Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning	2016		
26	RandWire-WS	80.1%	94.8%		×	Exploring Randomly Wired Neural Networks for Image Recognition	2019		
27	ScaleNet-152	79.94%	94.82%		×	Data-Driven Neuron Allocation for Scale Aggregation Networks	2019		
28	ResNet-200	79.9%	95.2%		×	Identity Mappings in Deep Residual Networks	2016		
29	Modified Aligned Xception	79.81%	94.83%		×	Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation	2018		
30	SKNet-101	79.81%			×	Selective Kernel Networks	2019		

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

In this class

Objectives

- To introduce modern approaches to important topics in Computer Science
- To help students build a ready-to-deploy image classification applications

Credits & activities

- Credits: 3
- Activity hours:
  - Class – theory: 21
  - Class – practice: 0
  - Self-study: 24
- Prerequisites:
  - INT2203 Data structures and Algorithms
  - INT1101 Statistics

Schedule

Week	Content	Class hour	Self-study hour
1	Introduction Image classification problem and its applications A toy problem with CIFAR10	2	1
2	CNN model architectures and visualization	2	1
3	Training and tuning parameters Automatic parameter learning	2	1
4	Data augmentation Data generator	2	2-6
5	Transfer learning	2	2-6
6	Multi-output image classification	2	2-6
7	Building a training dataset How to write a report	1	2-6
8, 9, 10, 11	Seminar: Bag of tricks with CNN (as mid-term tests)	1	2-6
12, 13, 14	Final project presentations	1-3	2-6
15	Class summarization	1	open

Mid-term tests

- Check state-of-the-art methods for each dataset
- Either:
  - Survey what to do to boost the performance
  - Apply SOTA in one dataset to another dataset and analyze the performance

Final projects

- Build a new dataset
- Create a model
- Train
- Find a nice set of hyper-parameters
- Report & present
- Extra point: implement UI applications

Assessment

- Attendance: 10%
- Seminars: 40%
  - Report + presentation
- Final projects: 50%
  - Report + presentation
- Plagiarism and Cheating: immediately 0 mark!

Libraries used in class

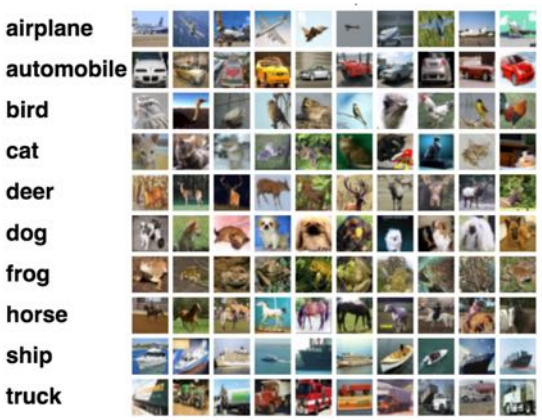
- Python + Keras
  - For image processing: OpenCV, imgaug
- However, for final projects or mid-term tests, any is okay
  - PyTorch
  - Torch
  - Tensorflow
- Make sure to report what you actually contributed

References

- Google scholar
- Google
- Kaggle
- http://rodrigob.github.io/are_we_there_yet/build/classification_datasets_results.html
- https://paperswithcode.com/task/image-classification

Toy problem CIFAR10

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



Environments

- Anaconda
- Tensorflow
- Keras
- Python3
- OpenCV 3
- Jupyter Notebook

Note: Run with GPU of Google Colab