AI CAMP

school of ai Algiers

Image Classification ✨

By : Kamel Brouthen & Inas Bachiri

Workshop Plan

01

Datasets & Evaluation

02

CNN Architectures

03

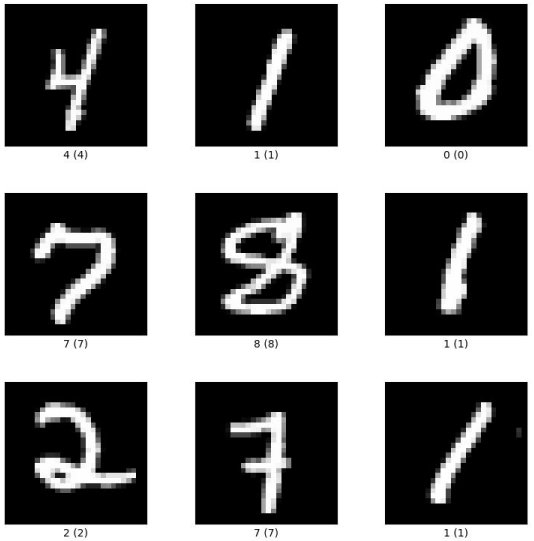
Transfer Learning

01

Datasets & Evaluation

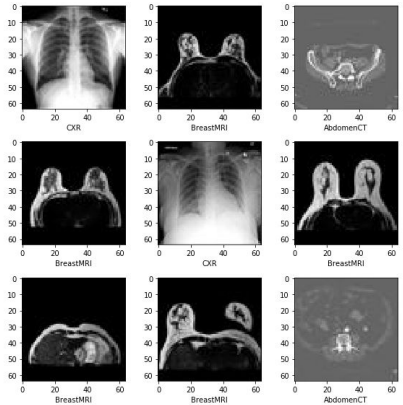
MNIST & Variants

- The "Hello World" of image classification.
- Consists of 28x28 grayscale images of handwritten digits (0-9)
- Cntains 60,000 training images and 10,000 testing images.



MNIST & Variants

Medical MNIST

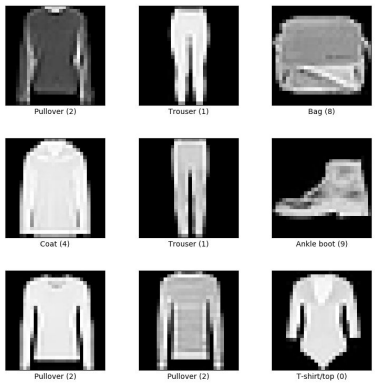


Learn more

Sign Language MNIST

Learn more

Fashion MNIST



Learn more

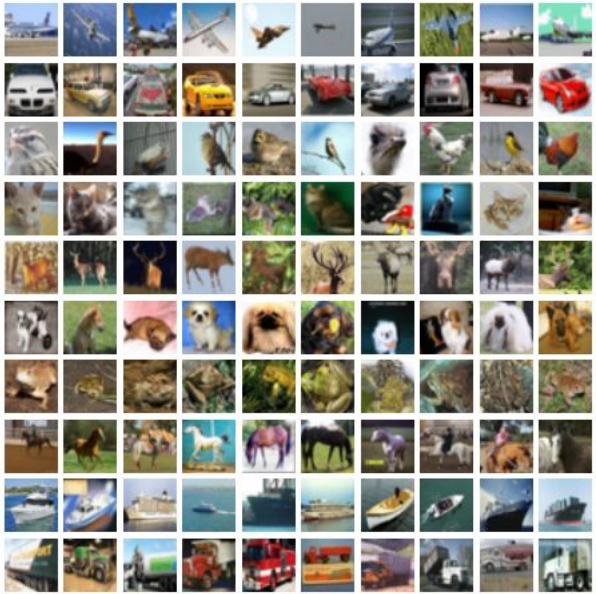
AI CAMP

school of ai Algiers

CIFAR

CIFAR-10

- Consists of 32x32 colored images in 10 classes, with 6000 images per class.
- The classes include common objects such as airplanes, cars, birds, cats, and more.



CIFAR-100

- Similar to CIFAR-10 but contains 100 classes, with 600 images per class.

ImageNet

- One of the most widely used and influential datasets in computer vision.
- Consists of over 14 million images.
- Has +20,000 classes.
- Each image is annotated with bounding boxes and hierarchical labels.

Evaluation

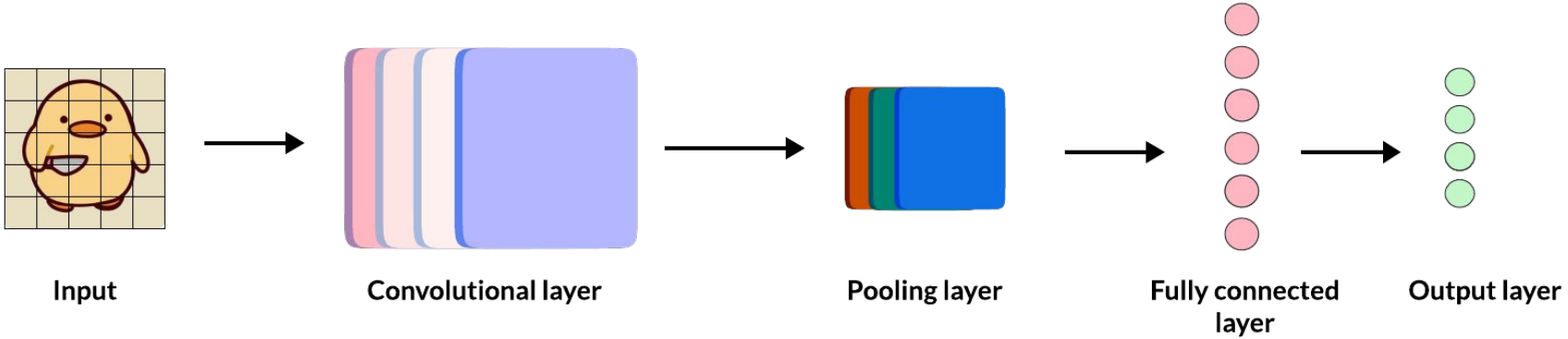
- As any other classification task, image classification models can be evaluated using the classic metrics :
  - Accuracy
  - Precision
  - Recall
  - F1 Score
  - Cross-entropy Loss
  - ....
- However, you might hear about some confusing terms. I got your back !
  - Top-1 Accuracy : The percentage of images for which the correct class is predicted as the top prediction, a.k.a ol' plain accuracy :)
  - Top-k Accuracy : The percentage of images for which the correct class is predicted within the top k predictions.

02

CNN Architectures

CNN Architectures

In the previous workshop, you learned what a simple **convolutional neural network architecture** is !

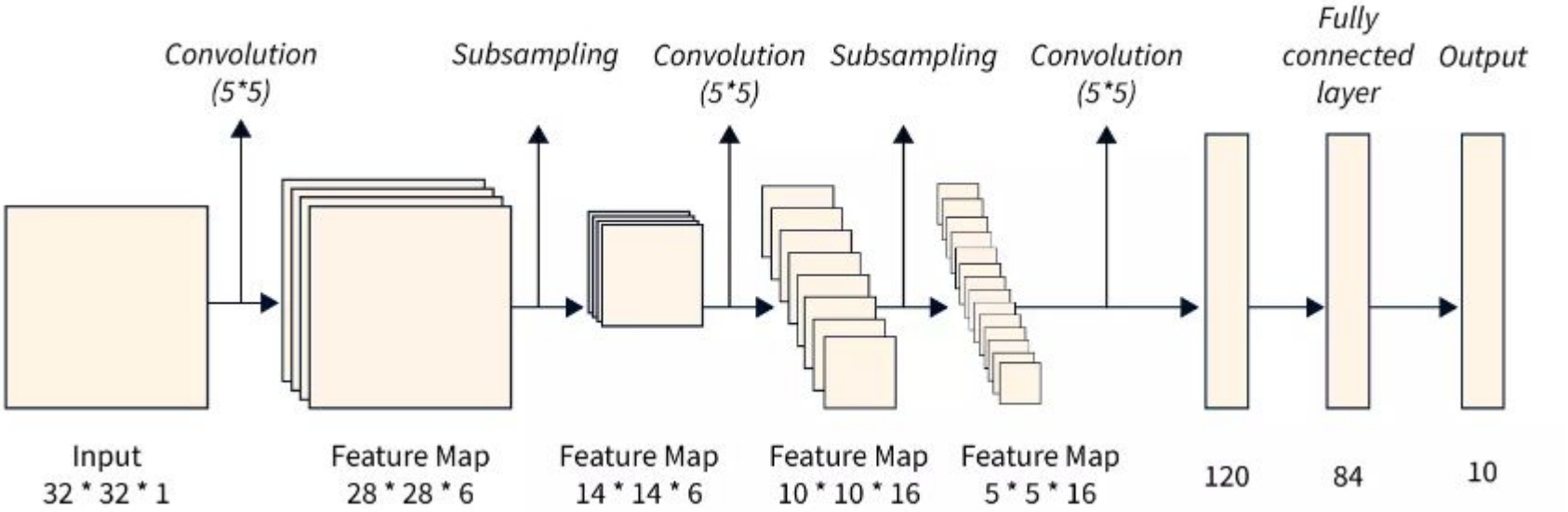


CNN Architectures

- The **number**, **arrangement**, **parameters** and **connections** between these components allow us to define various **CNN architectures** that work well according to the **dataset**, **task** and **hardware**.

- Let's explore some famous **CNN architectures** designed to tackle different complexities of **image classification** tasks

LeNet

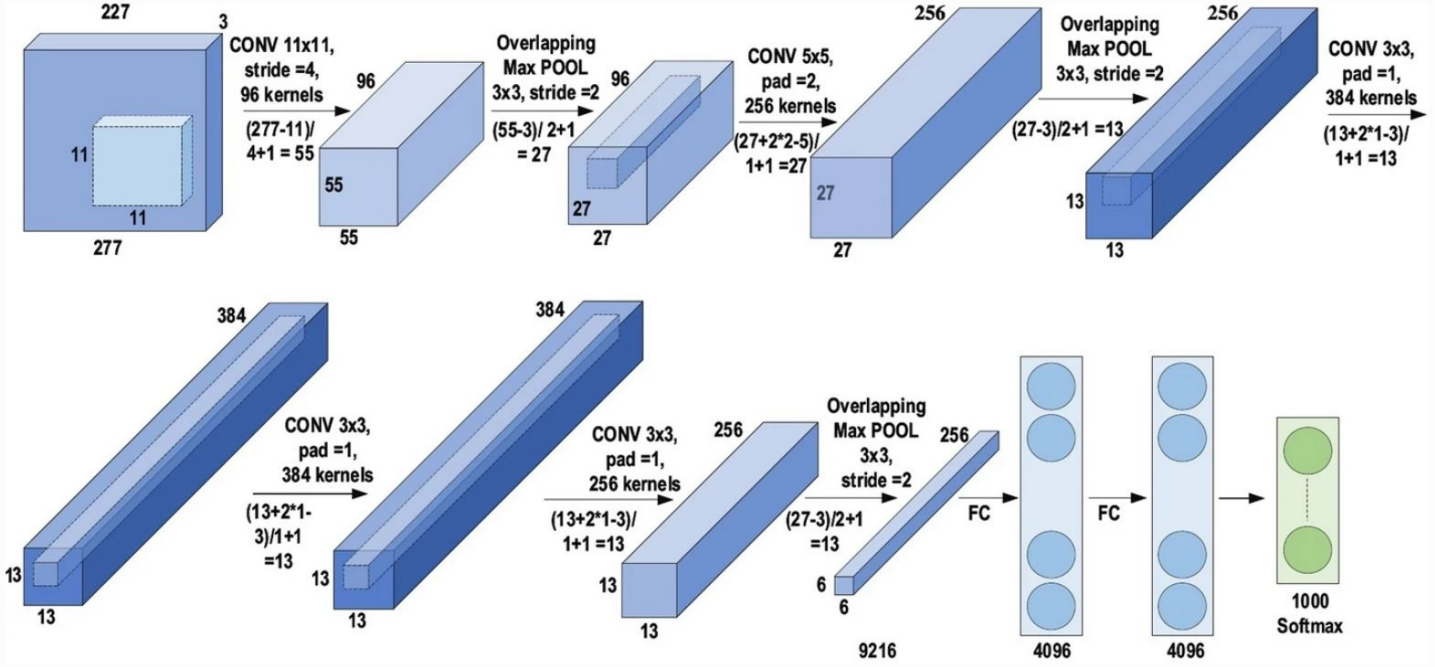

Convolution
(5*5)
Subsampling
Convolution
(5*5)
Subsampling
Convolution
(5*5)
Fully
connected
layer
Output
Input
32 * 32 * 1
Feature Map
28 * 28 * 6
Feature Map
14 * 14 * 6
Feature Map
10 * 10 * 16
Feature Map
5 * 5 * 16
120
84
10

LeNet

- Consists of convolutional layers followed by pooling layers and fully connected layers. Yeah, the grandma of CNNs :-)

- Initially used for handwritten digit recognition tasks.
- Pioneered the concept of trainable CNNs.

AlexNet



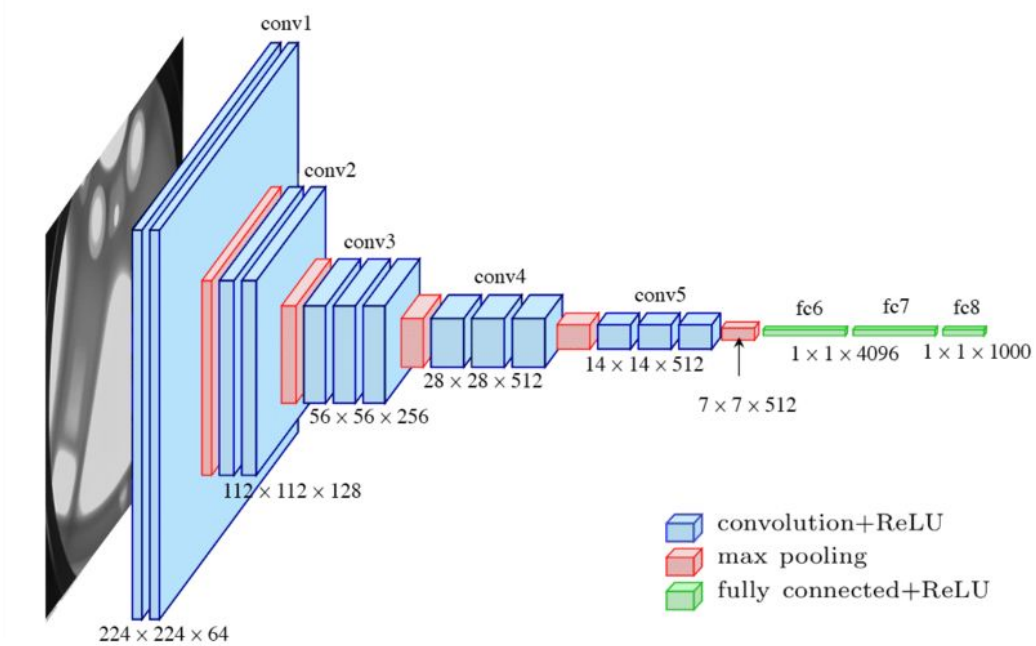
AI CAMP

school of ai Algiers

AlexNet

- Significantly deeper than previous networks at the time.
- Uses ReLU activation functions, dropout for regularization, and data augmentation techniques. She got it all 😎
- Won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012, sparking the deep learning revolution ✨

VGGNet



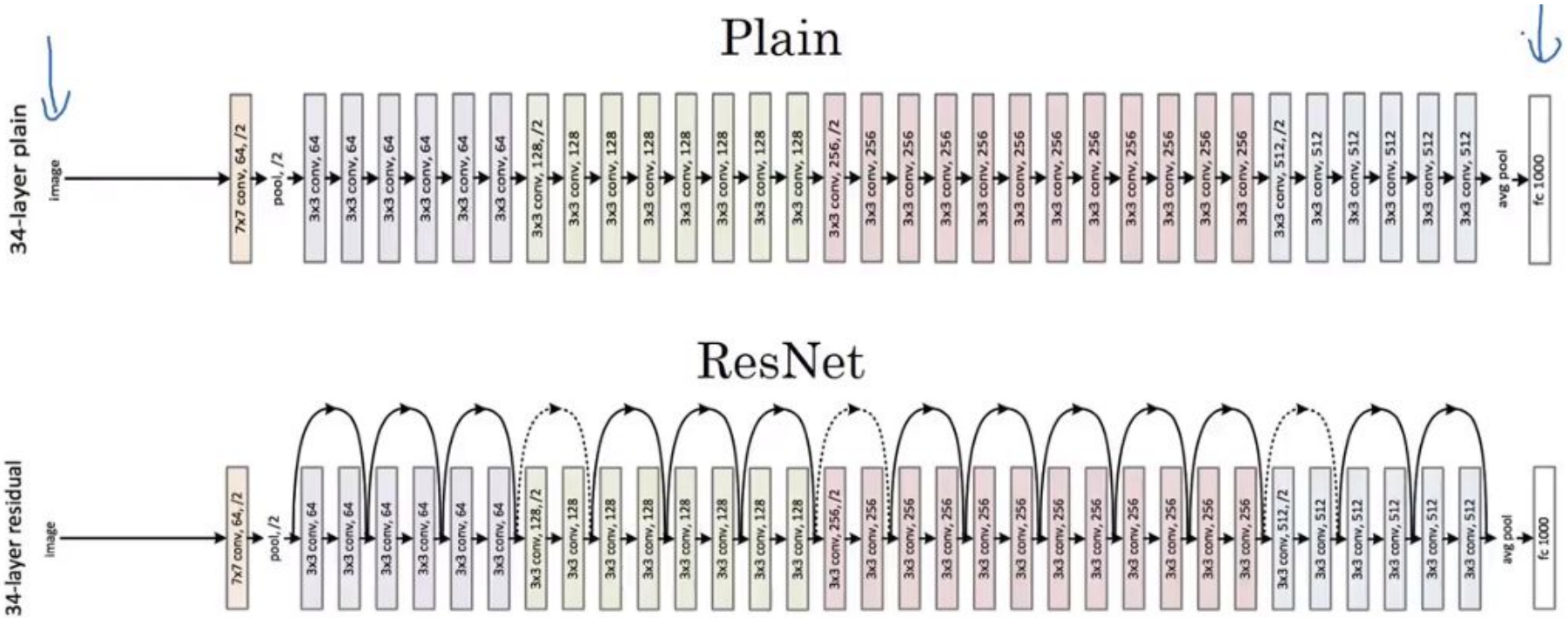
AI CAMP

school of ai Algiers

VGGNet

- Known for its simplicity and uniform architecture, consisting of multiple 3x3 convolutional layers.
- Available in different variants with varying depths (e.g., VGG16, VGG19).
- Achieved competitive performance on ImageNet classification.

ResNet



AI CAMP

school of ai Algiers

ResNet

- Addressed the vanishing gradient problem by utilizing skip connections (or residual connections) to enable training of very deep networks.
- Available in various depths, such as ResNet-18, ResNet-50, ResNet-101, and ResNet-152.
- Achieved state-of-the-art results on various image classification benchmarks.

There's more to explore !

ZFNet

GoogLeNet

EfficientNet

SENet

Xception

ResNeXt

DenseNet

AI CAMP

school of ai Algiers

As you can see, there isn't a limit to how many options we can explore. Crafting a good architecture is thus a complicated task, that requires a lot of time and expertise, and a non-ending loop of trials and errors !

So instead of starting from scratch with every new task, we can use the existing AWESOME models that AI scientists have came up with and tweak them according to our needs 

03

Transfer Learning

AI CAMP

school of ai Algiers

Transfer Learning

We call the existing awesome models : pretrained models.

- A pretrained model is a saved network that was previously trained on a large dataset, typically on a large-scale image classification task.

- Intuition : if a model is trained on a large and general enough dataset, this model will effectively serve as a generic model of the visual world.

  So we can then take advantage of them without having to start from scratch by training a large model on a large dataset.

Transfer Learning

We can use transfer learning in two ways :

1. **Using a pretrained model for feature extraction:** this means borrowing a pretrained model's knowledge to extract relevant features from new data, enhancing the performance of tasks without modifying the original model's parameters.

2. **Fine-tuning a pretrained model:** this involves adjusting the parameters of a pretrained model to better suit a new task or dataset. By fine-tuning, we allow the model to learn specific patterns of the new data while retaining the valuable knowledge it gained during pretraining.

Install linux, windows is for children

AI CAMP

school of ai Algiers

Thank you for attending

any questions ?

AI CAMP

school of ai Algiers

Thank you for attending

any questions ?

Don’t be shy, ask