Convolutional Neural Networks for Image Classification

Pablo Mesejo

pmesejo@go.ugr.es

Universidad de Granada

Departamento de Ciencias de la Computación e Inteligencia Artificial

UNIVERSIDAD DE GRANADA

DaSCI

Instituto Andaluz de Investigación en **Data Science** and **Computational Intelligence**

Readings

- Goodfellow, Bengio and Courville (2016), Chapter 9.
- Szeliski (2022), Chapter 5.3, 5.4, and 6.2.
- Zhang, Lipton, Li and Smola (2023), *Dive into Deep Learning*, Chapter 7 and 8.
- Stanford University CS231n: Deep Learning for Computer Vision. Lectures 5 and 6.

Reviewing ConvNets

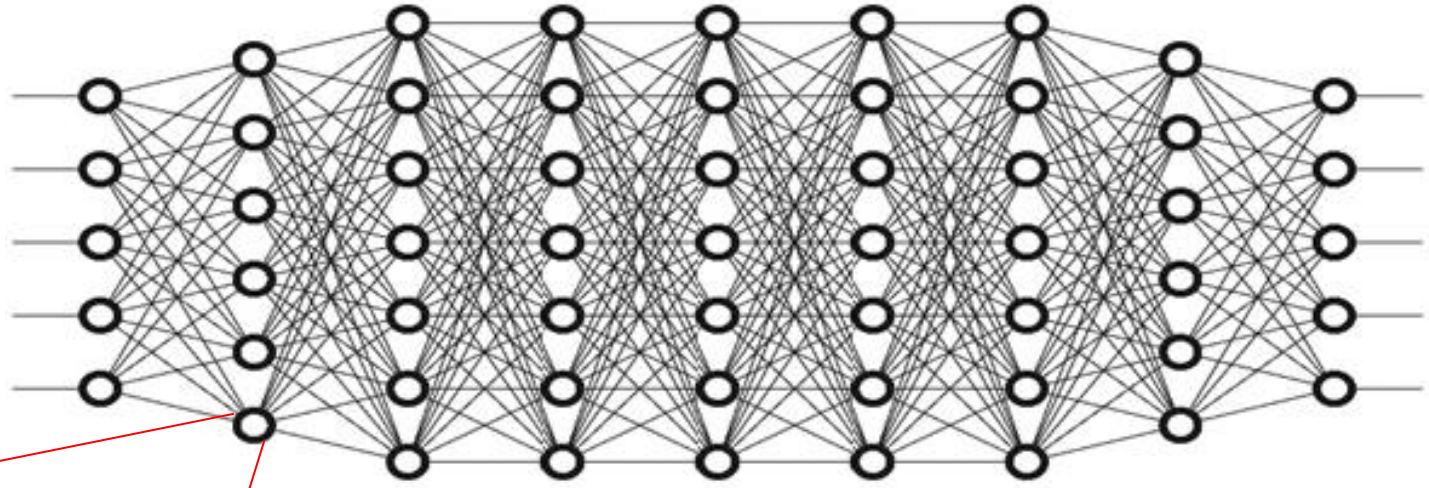
Reviewing ConvNets

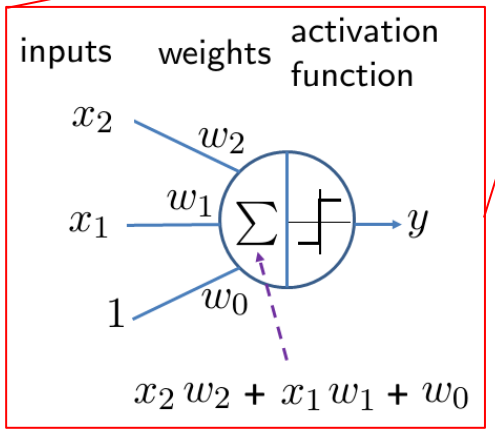
- We have devoted 3-4 weeks of practical sessions to introduce the main concepts related with Convolutional Neural Networks (ConvNets) and Deep Learning (DL).

- Now, we'll review some of those concepts using quizzes!

Question 1

- Could this architecture be considered a ConvNet?





Every unit in the network above corresponds to a perceptron like the one displayed on the left. Furthermore, the network also incorporates batch normalization and dropout.

Question 2

- Which layer is the main responsible for downsampling and reducing the spatial dimensions in a ConvNet?

A. Convolutional layer
B. Pooling layer
C. Fully connected layer
D. Batch normalization layer

Question 3

- In a ConvNet, what is the purpose of the activation function typically used after the convolutional and fully connected layers?

A. Faster training
B. Introduce non-linearity
C. Improve regularization
D. Normalize input data

Question 4

- What is the purpose of the kernel (filter) in a convolutional layer?

A. Reducing dimensionality
B. Feature extraction
C. Model regularization
D. Gradient computation

Question 5

- Which of the following statements about transfer learning in ConvNets is true?

A. It requires retraining the entire model from scratch.
B. It needs to manually set some of the filters.
C. It involves using pre-trained models on similar tasks.
D. It can only be applied to shallow networks.

Question 6

- What is the purpose of the "stride" parameter in a convolutional layer?

A. Control the learning rate
B. Define the size of the kernel
C. Specify the number of filters
D. Set the step size for sliding the kernel

Question 7

- What is the role of data augmentation in training deep neural networks?

A. Increase model complexity
B. Improve generalization
C. Reduce the number of parameters
D. Speed up training time

Question 8

- What does the term "padding" refer to in a convolutional layer?

A. Adding extra layers to the model
B. Filling missing values in the data
C. Adding values (zeros, for instance) around the input volume
D. Normalizing the input data

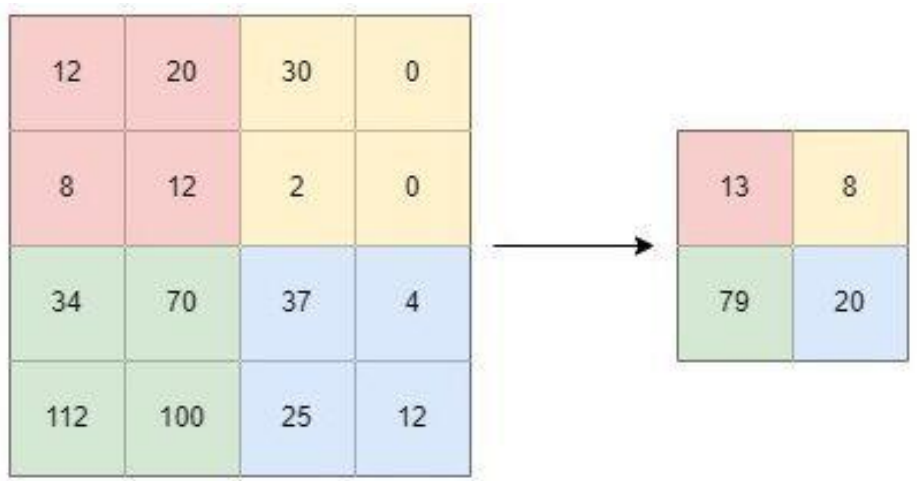
Question 9

- In the context of ConvNets, what is the purpose of the softmax activation function in the output layer?

A. Introduce non-linearity
B. Normalize the output into probabilities in a regression problem
C. Speed up training convergence
D. None of the above

Question 10

- In the context of ConvNets, what kind of operation do we visualize below?

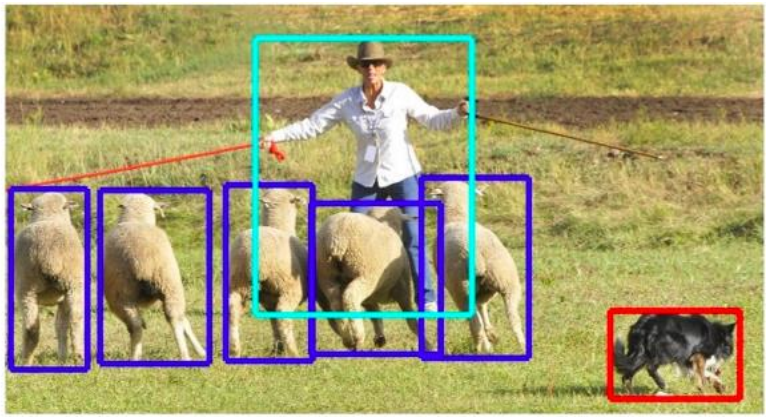


- How many extra parameters (weights) does it add to the network training?

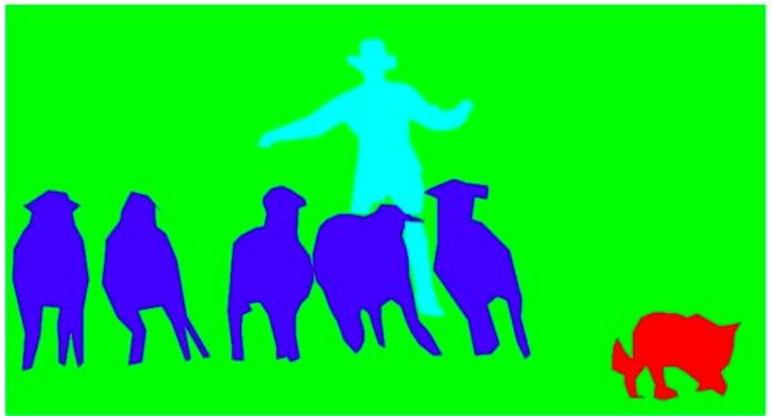
Question 11

- Could you associate images and computer vision tasks?

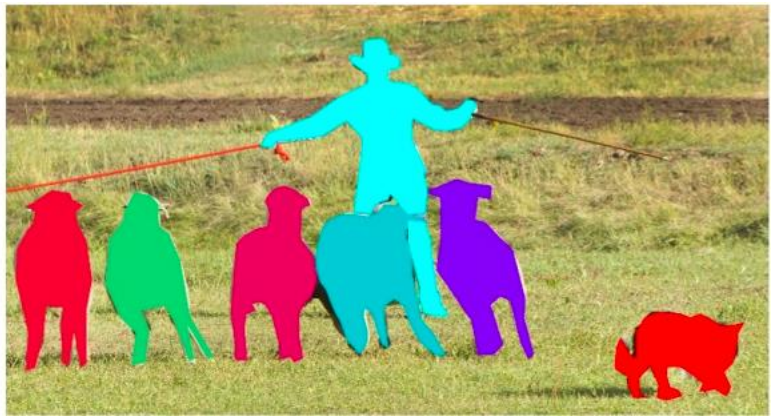
A



B



C



D

1) instance segmentation
2) object detection
3) image classification
4) semantic segmentation

Question 12

- Which of the following statements related with Dropout is incorrect?

A. Dropout is mainly a regularization technique

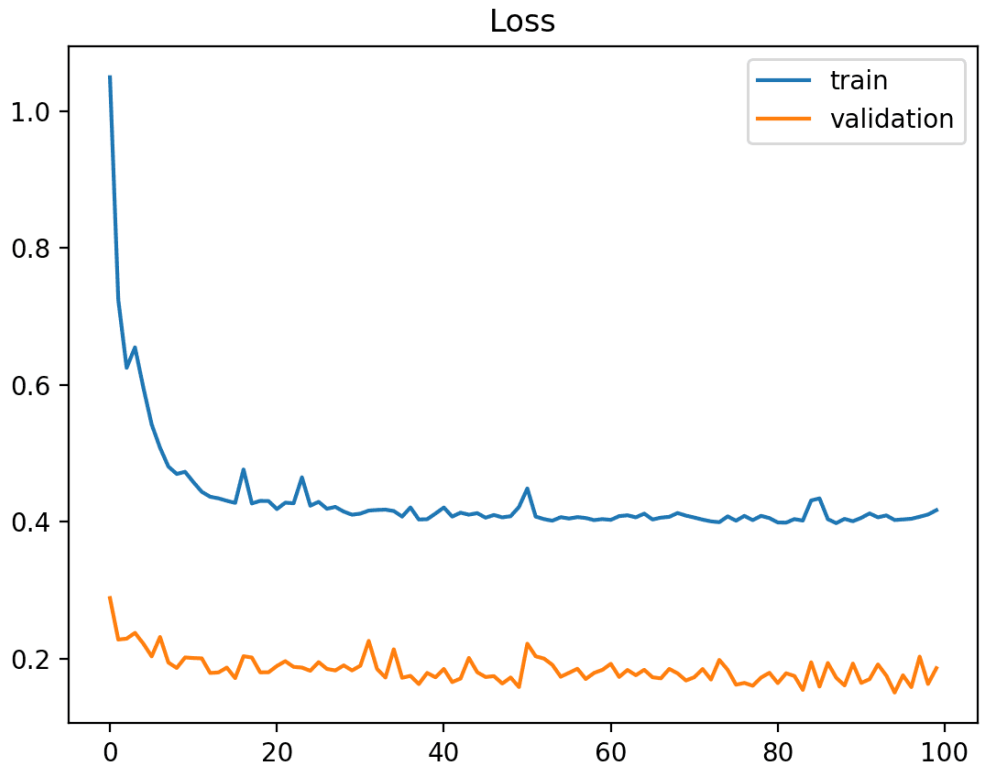
B. It does not introduce any extra parameter (weight)

C. It does not make much sense to use $p = 1$

D. It can only be used with fully connected layers

Question 13

- What would be your diagnosis/hypothesis if you get the following learning curves?



A. We are overfitting the training set. Therefore, we need regularization.

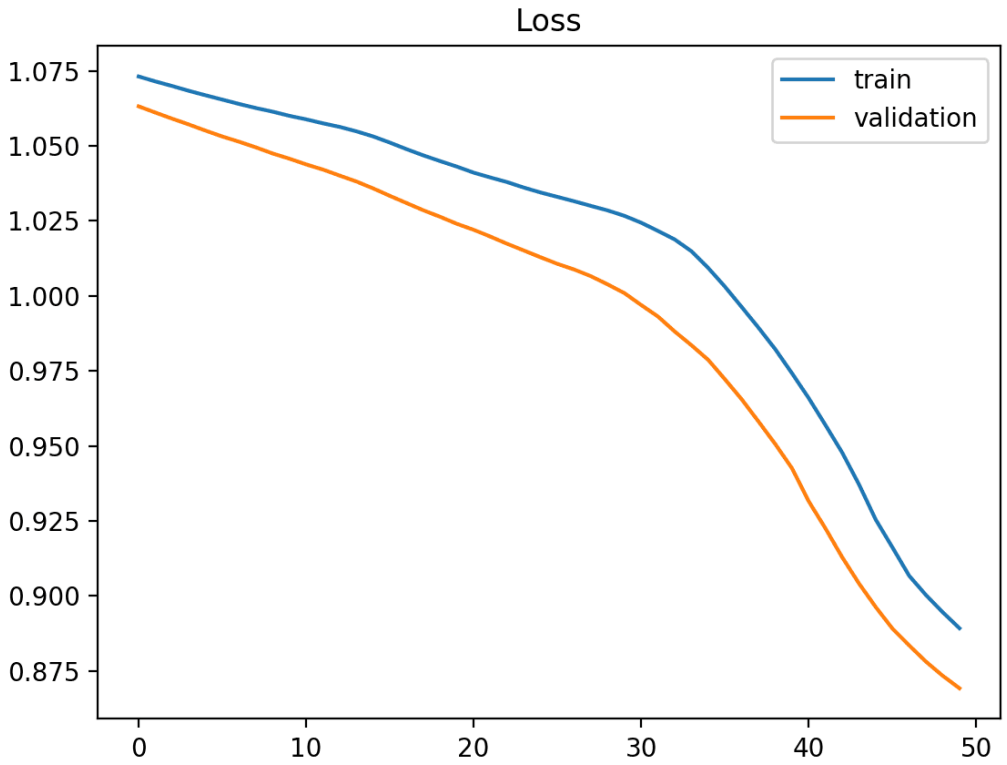
B. Slow start: improper initialization of weights.

C. Underfitting scenario (our model is maybe too simple, and it cannot learn the training set).

D. Unrepresentative validation set (it could be easier to predict than the training set)

Question 14

- What would be your diagnosis/hypothesis if you get the following learning curves?



A. Overfit: model too large/dataset too small.

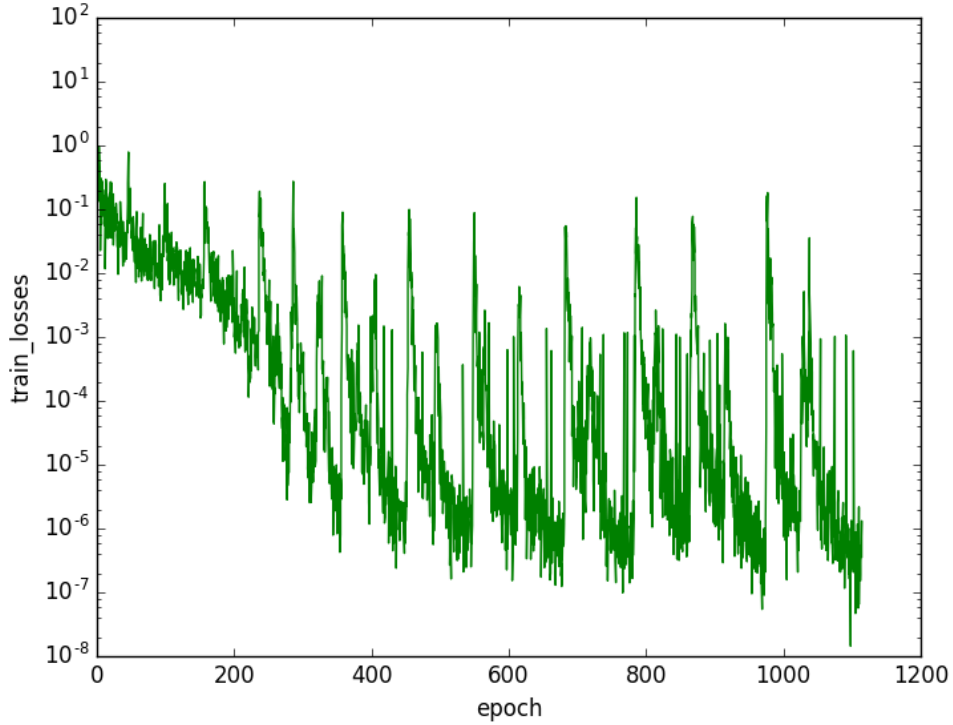
B. Underfitting scenario (our model may require further training).

C. Learning rate maybe too large.

D. These learning curves show directly a good fit. No more training is required.

Question 15

- What would be your diagnosis/hypothesis if you get the following training curve?



A. Gradients not applied to weights.

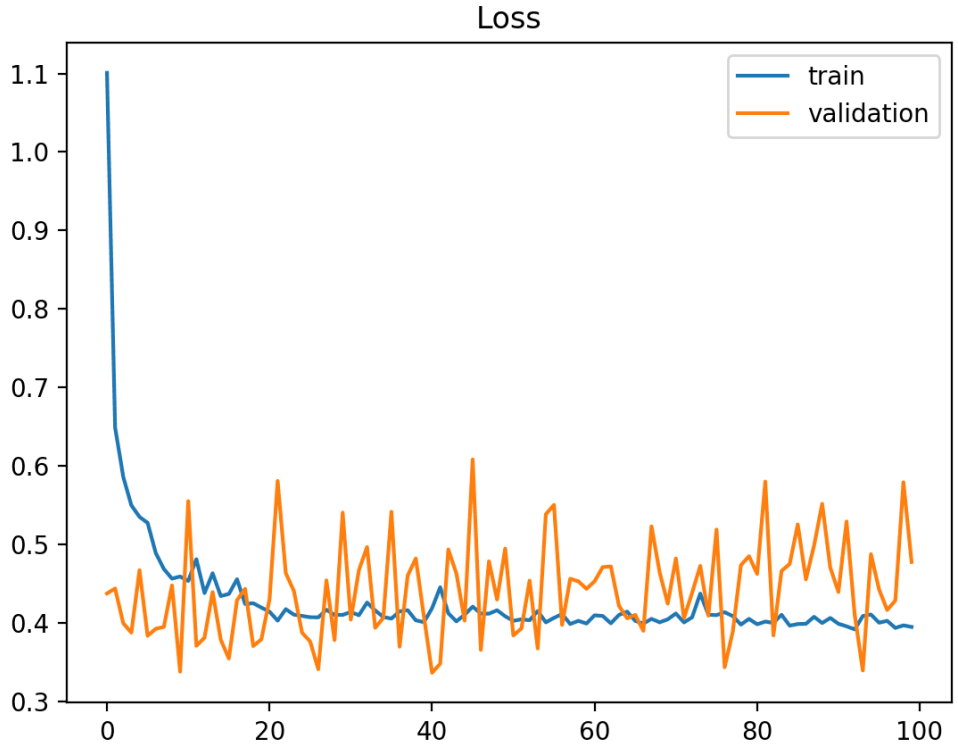
B. Applied the negative of gradients.

C. Learning rate maybe too small.

D. Not shuffling data after every epoch, periodical patterns in loss curve.

Question 16

- What would be your diagnosis/hypothesis if you get the following learning curves?



A. Overfitting scenario: there is a large generalization gap between training and validation.

B. Unrepresentative validation set (validation dataset has too few examples).

C. Underfitting scenario (our model is maybe too simple, and it cannot learn the training set).

D. Learning rate maybe too large.

Question 17

- If we have a dataset with 10.000 images and we decide to use minibatches of 8 images, how many iterations will take to complete an epoch?

Question 18

- Let’s assume the following input volume and the filter used. What is the resulting output volume?
  - Input volume: 224x224x3
  - Filters used: 68 filters of 3x3 with stride 2 and no padding

Question 19

- What is the number of parameters (weights) for this convolutional architecture?

Layer Type	Kernel Size (for convolutional layers)	Input | Output dimension	Input | Output channels (for convolutional layers)
Conv	3x3	32x32 | 30x30	3 | 15
BN	-	30x30 | 30x30	-
ReLU	-	30x30 | 30x30	-
MaxPooling	2x2	30x30 | 15x15	-
Conv	3x3	15x15 | 13x13	15 | 10
BN	-	13x13 | 13x13	-
ReLU	-	13x13 | 13x13	-
Dropout	-	13x13 | 13x13	-
FC	-	1690 | 50	-
ReLU	-	50 | 50	-
FC	-	50 | 25	-

Question 20

- How does Dropout (with probability *p*) operate at inference time (validation and test)?

A. The same as training, removing some units.

B. At test time, the unit is always present and the weights are multiplied by *p*.

C. At test time, the unit is always present but you don’t do anything in particular.

D. At test time, the unit is always present and the weights are divided by *p*.

Question 21

- If you take a ConvNet and remove all non linearities, can you improve performance if you…?

A. Increase the number of filters/kernels per convolutional block

B. Increase the number of layers in your network

C. Increase filters/kernels sizes

D. Train for more epochs

E. None of the above

Question 22

- What is a characteristic issue associated with saturating sigmoid activation functions?

A. Exploding gradients
B. Dead neurons
C. Vanishing gradients
D. Dead ReLU units

Question 23

- What is the primary effect and disadvantage of dead ReLU units in a neural network?

A. They lead to exploding gradients.
B. They contribute to overfitting.
C. They prevent the network from learning.
D. They cause vanishing gradients.

Question 24

- Evaluate whether the following statements are true or false.

A. You should not use a linear activation function in an output unit for a multiclass classification problem. True/False

B. You should not use cross-entropy loss in a regression problem. True/False

C. A convolutional layer is equivariant to translations. True/False

D. Pooling helps to make the representation approximately invariant to small translations of the input. True/False

Question 25

- Let’s say that you found an awesome network on some related work, but the input is not the same as the one of your problem… ☹

A. Bad luck. Forget about it, you just cannot do anything.

B. Nice! I can apply it directly to my problem!

C. Wait! I just need to do some finetuning!

D. Need to deal with larger/smaller input. E.g., you can stack one more conv+pool block, or change kernel size/stride

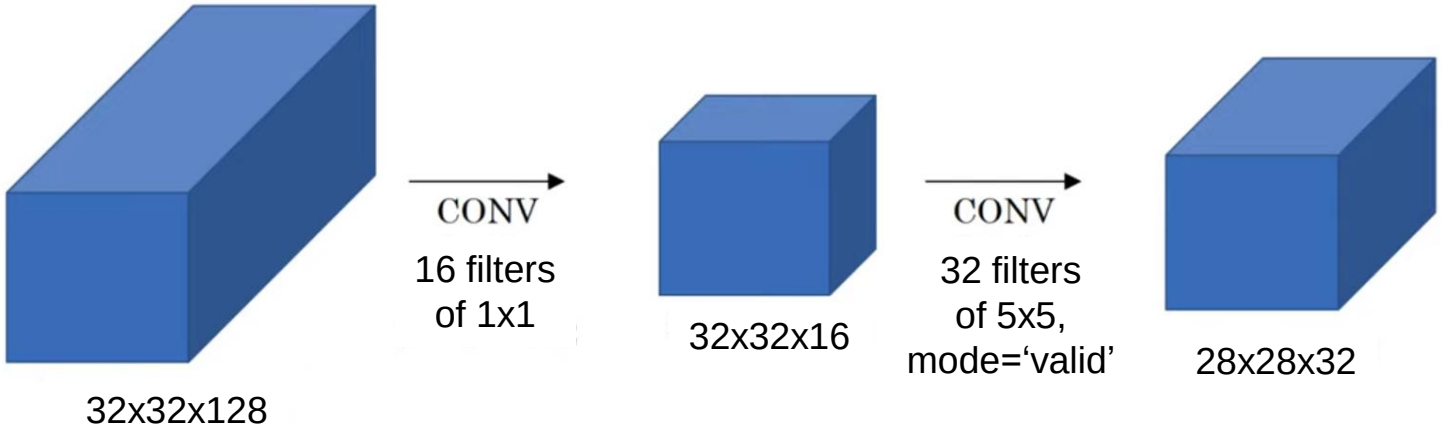
Question 26

- Let’s imagine that I decide to incorporate a new type of layer into my neural model. Which of the following is true?

A. Great! Let’s include it straight away in our model!
B. Wait… is it differentiable? Because otherwise we cannot include it… (unless we use non gradient-based approaches to train)
C. Wait… are we sure that it does not incorporate more weights to train?
D. It doesn’t make sense to introduce more types of layers. We know from theory that convolutional layers, pooling layers, and fully connected layers are all we need to approximate any function.

Question 27

- How many multiplications are performed in the figure below (where we move from 32x32x128 to 28x28x32, using 16 1x1 filters and 32 5x5 filters)?



Question 28

- If you are working with a ConvNet whose training and test accuracy are lower than expected, what would you think it would be a good strategy to follow?

A. Reduce the number of filters per convolutional block.

B. Increase regularization.

C. Reduce network depth.

D. None of the above.

Question 29

- If we want to reduce dimensionality in a ConvNet, what can we do?

A. Use a larger stride.

B. Use larger filters/kernels (no padding).

C. Use pooling.

D. All are correct.

Question 30

- Evaluate whether the following statements are true or false.

A. Gradient descent uses backpropagation. True/False

B. Backpropagation uses gradient descent. True/False

C. The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. True/False

D. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. True/False

Convolutional Neural Networks for Image Classification

Pablo Mesejo

pmesejo@go.ugr.es

Universidad de Granada

Departamento de Ciencias de la Computación e Inteligencia Artificial

UNIVERSIDAD DE GRANADA

DaSCI

Instituto Andaluz de Investigación en **Data Science** and **Computational Intelligence**