Q3

Due 5 Feb at 5:30 **Points** 100 **Questions** 14 **Available** 29 Jan at 9:00 - 5 Feb at 5:30 7 days **Time limit** None

Submission details:

Time:	20 minutes
Current score:	70 out of 100
Kept score:	70 out of 100

Instructions

Instructions:

1. You have 30 minutes to attempt the quiz
2. Once you start the quiz, you cannot go back and re-attempt it
3. You will not find answers online, so please make sure you are ready for the quiz
4. For Multiple Answer Questions, ALL the answers must be correct to score any point

Sometimes you might see multiple empty options. Please do not consider those empty options, that's some rendering issue, the options you see are the only options available for that question.

Attempt history

	Attempt	Time	Score
LATEST	Attempt 1	20 minutes	70 out of 100

Score for this quiz: **70** out of 100
Submitted 5 Feb at 1:48
This attempt took 20 minutes.

Question 1 — 0 / 5 pts

(Mostly) whenever we see kernel visualizations online (or some other reference) we are actually seeing:

- Correct! ☑ What kernels extract
- You Answered ☑ Feature Maps
- ☐ How Kernels Look

Question 2 — 3.33 / 5 pts

What all do we need to consider when we decide the number of kernels in our 11x11 receptive field layer?

- Correct! ☑ Expressiveness required
- Correct answer ☐ Inter and intra class variations
- Correct! ☑ Hardware capacity
- ☐ Total number of images in the dataset

Question 3 — 6.67 / 10 pts

Select the ones which are true

- Correct answer ☐ We use strides sometime on resource constraint hardware
- Correct! ☑ We tend not to use strides as they do not read spatial data evenly, causing checkboard issue
- Correct! ☑ When using strides, the channels created after convolutions are blurry (not consistent)
- ☐

Question 4 — 10 / 10 pts

What are the benefits of 1x1 Convolution?

- Correct! ☑ Lesser computation requirement for reducing the number of channels
- Correct! ☑ Use of existing channels to create complex channels (instead of re-convolution)
- Correct! ☑ Less number of parameters
- Correct! ☑ Reduces the burden of channel selection on 3x3. (select this answer "as well" even though you may not know this, we will discuss this in the class)

Question 5 — 5 / 5 pts

Why do we not use 1x1 to increase the number of channels?

- Correct! ☑ That's not true. We can use 1x1 to increase the number of channels, just that we need to have a purpose
- ☐ Because 1x1 is not an ideal method increase the number of channels

Question 6 — 5 / 5 pts

Why do we need an activation function?

- Correct! ☑ To provide decision making power to the neurons/DNN
- Correct! ☑ To provide non-linearity

Question 7 — 10 / 10 pts

Why do we need non-linearity in our neural networks?

- Correct! ☑ Not everything can be expressed using linear functions
- Correct! ☑ Non-linearity allows DNN to act like a Universal Activation Function

Question 8 — 0 / 5 pts

Why sigmoid activation functions are not used?

- Correct! ☑ They cause vanishing gradient issue
- You Answered ☑ They cause gradient explosion

Question 9 — 5 / 5 pts

Select which activation function from the ones below you'll use in CNN at TSAI. Please note if you do not select ReLU, you will get 0 as the marks for this question.

- Correct! ☑ ReLU
- ☐ ELU
- ☐ Sigmoid
- ☐ SELU
- ☐ TanH
- ☐ SReLU
- ☐ LeakyReLU

Question 10 — 5 / 5 pts

Promise me you'll try and never use Fully Connected Layers at TSAI!

- Correct! ☑ I Promise
- ☐ No, I know they will add too many parameters and there are better alternatives, but I will still use FC layers.

Question 11 — 0 / 10 pts

Why do we generally not prefer to add stride of more than 1?

- Correct! ☑ It causes checkerboard issue, as we are not reading all pixels equal number of time (ignoring the corner pixels)
- ☐ It increases the number of parameters
- ☐ It increases the channel size
- You Answered ☑ It reduces the channel size

Question 12 — 10 / 10 pts

What all features does ReLU provide us?

- Correct! ☑ Easy way to communicate with BackProp to use negative values if that information needs to be filtered out
- Correct! ☑ Easy way to communicate with BackProp to use positive values if some information needs to be not filtered out
- Correct! ☑ Very low computation requirements

Question 13 — 10 / 10 pts

ReLU is defined as:

0 when x is less than or equal to zero
x when x is more than zero

Any activation function must be differentiable if we were to use it in our DNNs (else backprop would not work). Knowing that we indeed use ReLU, what do you think is the derivative of ReLU?

- Correct! ☑ 0 when x is less than or equal to zero, 1 when x is positive
- ☐ 0 when x is less than zero, not defined when x is equal to zero, and 1 when x is positive
- ☐ 1 when x is less than or equal to zero, 0 when x is positive
- ☐ 0 when x is less than or equal to zero and x when x is more than zero

Question 14 — 0 / 5 pts

We know that when we use a kernel of size 3x3 and a stride of 1, the receptive field increases by 2.

If we use MaxPooling with kernel size 3x3 and with a stride of 1, will the receptive field increase by 2?

- Correct answer ☐ True
- You Answered ☉ False

Quiz score: **70** out of 100