Break & Quiz

Q 1.1: Which of the following is **not** true?

- A. Adding more layers can improve the performance of a neural network.
- B. Residual connections help deal with vanishing gradients.
- C. CNN architectures use no more than ~20 layers to avoid problems such as vanishing gradients.
- D. It is usually easier to learn a zero mapping than the identity mapping.

Break & Quiz

Q 1.1: Which of the following is **not** true?

- A. Adding more layers can improve the performance of a neural network.
- B. Residual connections help deal with vanishing gradients.
- **C. CNN architectures use no more than ~20 layers to avoid problems such as vanishing gradients.**
- D. It is usually easier to learn a zero mapping than the identity mapping.

Break & Quiz

Q 1.1: Which of the following is **not** true?

- A. Adding more layers can improve the performance of a neural network. (Yes, as long as we're careful, e.g., ResNets.)
- B. Residual connections help deal with vanishing gradients. (Yes, this is an explicit consideration for residual connections.)
- **C. CNN architectures use no more than ~20 layers to avoid problems such as vanishing gradients.** (No, much deeper networks.)
- D. It is usually easier to learn a zero mapping than the identity mapping. (Yes: simple way to learn zero is to make weights zero)

Break & Quiz

Q 2.1: If we apply data augmentation blindly, we might

(i) Change the label of the data point

(ii) Produce a useless training point

- A. (i) but not (ii)
- B. (ii) but not (i)
- C. Neither
- D. Both

Break & Quiz

Q 2.1: If we apply data augmentation blindly, we might

(i) Change the label of the data point

(ii) Produce a useless training point

- A. (i) but not (ii)
- B. (ii) but not (i)
- C. Neither
- **D. Both**

Break & Quiz

Q 2.1: If we apply data augmentation blindly, we might

(i) Change the label of the data point

(ii) Produce a useless training point

- A. (i) but not (ii) (Can do (ii): imagine turning up the contrast till the image is completely black and is unusable).
- B. (ii) but not (i) (Can change label: rotate a 6 into a 9).
- C. Neither (Can do either).
- **D. Both**

Break & Quiz

Q 2.2: What are some consequences of data augmentation?

(i) We have to store a much bigger dataset in memory

(ii) For a fixed batch size, there will be more batches per epoch

- A. (i) but not (ii)
- B. (ii) but not (i)
- C. Neither
- D. Both

Break & Quiz

Q 2.2: What are some consequences of data augmentation?

(i) We have to store a much bigger dataset in memory

(ii) For a fixed batch size, there will be more batches per epoch

- A. (i) but not (ii)
- **B. (ii) but not (i)**
- C. Neither
- D. Both

Break & Quiz

Q 2.2: What are some consequences of data augmentation?

(i) We have to store a much bigger dataset in memory

(ii) For a fixed batch size, there will be more batches per epoch

- A. (i) but not (ii) ((i) is false: can store original points only, and then transform them on-the-fly).
- **B. (ii) but not (i)**
- C. Neither ((ii) is true: more points, same batch size means more batches).
- D. Both