Class Test on Convolutional Neural Networks (CNN)

Total Marks: 20
Time: 40 Minutes

Answer the following questions in brief and simple form.

1. **What is a Convolutional Neural Network (CNN)? Why is it mainly used?** (1)
   A CNN is a deep learning model mainly used for image-related tasks. It automatically learns spatial features (like edges and patterns) from images. For example, CNNs are used in face recognition, object detection, and handwriting recognition systems.

2. **Why is the convolution process necessary in CNN?** (1)
   Convolution helps extract useful features by applying filters across the image. It detects edges, colors, and shapes that help the network understand visual patterns. For example, an early convolution layer might detect edges, while deeper layers detect objects like eyes or wheels.

3. **What is the main difference between CNN and a fully connected neural network?** (1)
   CNN uses local connections and shared weights to process images efficiently, while a fully connected network connects every neuron to all others, leading to more parameters. CNNs are more suitable for image data due to their spatial awareness and lower computational cost.

4. **What is the purpose of using filters (kernels) in CNN?** (1)
   Filters (small matrices like 3×3 or 5×5) slide over the image to detect features such as lines, textures, or corners. Each filter extracts a specific type of pattern from the image.

5. **Define "feature extraction" in CNN.** (1)
   Feature extraction is the process of identifying key visual patterns like edges, colors, and textures from an image. Convolution and pooling layers perform this automatically. For instance, a CNN can extract features that help recognize whether an image contains a cat or a dog.

6. **What is padding? Why is it used in CNN?** (1)
   Padding means adding extra pixels (usually zeros) around the input image. It helps preserve the image size after convolution and prevents loss of information at the edges. For example, using "same padding" ensures the output size equals the input size.

7. **Explain stride and its effect on the output size.** (1)
   Stride is the number of pixels the filter moves at each step. A larger stride reduces the output size and speeds up computation, while a smaller stride keeps more detail. For example, stride = 1 keeps full detail, while stride = 2 halves the output size.

8. **What is pooling? Differentiate between Max Pooling and Average Pooling.** (2)
Pooling reduces the spatial size of feature maps and makes the model more robust.
Max Pooling: Takes the largest value from a region (keeps strong features).
Average Pooling: Takes the average of values in a region (smooths features).
Example: In a 2×2 region [1, 3; 2, 4], max pooling gives 4, average pooling gives 2.5.

9. **What is the purpose of the activation function (ReLU) in CNN?** (1)
ReLU (Rectified Linear Unit) adds non-linearity by converting negative values to zero. This helps CNNs learn complex patterns. For example, $f(x) = \max(0, x)$ ensures faster learning and prevents the vanishing gradient problem.

10. **Name the main layers used in CNN architecture.** (1)
The main layers are: Convolution Layer, Pooling Layer, Activation Layer (ReLU), Flatten Layer, Fully Connected Layer, and Output Layer.

11. **What is flattening and why is it required before the fully connected layer?** (1)
Flattening converts a 2D matrix (from convolutional layers) into a 1D vector. This allows the fully connected layer to process the extracted features and make final predictions. For example, a $4 \times 4 \times 3$ feature map becomes a single 48-element vector.

12. **What is the function of the fully connected layer?** (1)
The fully connected (FC) layer combines all extracted features to classify the image into categories. For instance, after flattening, the FC layer may output probabilities like [Cat: 0.9, Dog: 0.1].

13. **Explain the concept of weight sharing in CNN.** (2)
Weight sharing means the same filter (set of weights) is used across all positions of the input. This reduces the total number of parameters and ensures the same feature is detected everywhere. For example, a "vertical edge detector" filter detects vertical lines in all image regions.

14. **How does CNN handle large image inputs efficiently?** (1)
CNN uses convolution and pooling layers to reduce dimensions gradually while keeping key features. This reduces computation and memory usage without losing important visual details.

15. **Mention two real-life applications of CNN.** (1)
1. Face recognition in security systems.
2. Medical image analysis (e.g., detecting tumors in MRI scans).

16. **How does CNN achieve translation invariance?** (1)
Through convolution and pooling, CNN recognizes an object even if it moves slightly within the image. For example, a dog can be detected whether it's in the left or right corner of the photo.

17. **What is overfitting and how can it be reduced?** (1)
Overfitting happens when the model memorizes training data but performs poorly on new data. It can be reduced by techniques like dropout, data augmentation,

and regularization (L2). Example: Randomly flipping images helps avoid overfitting.

18. **What are pre-trained CNN models? Give one example.** (1)
Pre-trained models are CNNs trained on large datasets (like ImageNet) and reused for other tasks. Example: VGG16, ResNet50, and InceptionV3 are common pre-trained models.

19. **Why do we use dropout layers in CNN?** (1)
Dropout randomly turns off some neurons during training to prevent over-reliance on specific paths. This improves model generalization and reduces overfitting.

20. **How can data augmentation improve CNN performance?** (1)
Data augmentation increases dataset variety by applying transformations like rotation, flipping, zooming, or brightness changes. This helps CNN generalize better to unseen images.

21. **Why do we need to freeze some neurons while convolution operations?** (1)
Freezing neurons (or layers) means keeping their weights unchanged during training. This is done when using pre-trained models to retain learned features like basic edges or colors, while only training new layers for a specific task.

Total Marks: 20

Best of luck!