Image Segmentation Questions and Solutions

ETH Zurich

Note: For all multiple choice questions, exactly one option is correct.

Preparation Questions

1. Analyzing K-means Algorithm Complexity

Question: What is the algorithmic complexity of the K-means algorithm per iteration, where K is the number of clusters and N is the number of data points?

a) $O(N^2)$

b) $O(K \cdot N)$

c) $O(K^2 \cdot N)$

d) $O(N \cdot \log N)$

Correct Answer: b) $O(K \cdot N)$

Explanation: The K-means algorithm consists of two main steps in each iteration:

1. Assignment step: Each of the N data points must be compared to all K centroids to find the nearest centroid, requiring $O(K \cdot N)$ operations.
2. Update step: Computing new centroids requires averaging the points in each cluster, taking $O(N)$ operations.

Therefore, the dominant term is $O(K \cdot N)$, making this the overall complexity per iteration.

2. Role of Bandwidth in Kernel Density Estimation (KDE)

Question: In Kernel Density Estimation (KDE), what is the primary role of the bandwidth parameter h?

a) The smoothness of the estimated density function

b) The number of clusters in the resulting distribution

c) The dimensionality of the feature space

d) The mean of the underlying data distribution

Correct Answer: a) The smoothness of the estimated density function
Explanation: The bandwidth parameter h determines:

- The spread of each kernel function around its data point
- Larger h values result in smoother density estimates where each data point influences a wider area
- Smaller h values produce more detailed, peaked estimates that closely follow individual data points
- Serves as a key parameter balancing between over-smoothing and under-smoothing of the density estimate

3. Calculating mIoU and mPA from Ground Truth and Predictions

Question: Given the ground truth matrix Y and prediction matrix P:

$$Y = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Calculate the mean Intersection over Union (mIoU) and mean Pixel Accuracy (mPA).

Solution:

1. **Intersection over Union (IoU):**
   - Class 1:
   $$\text{IoU}_1 = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} = \frac{1}{1+1+1} = \frac{1}{3}$$
   - Class 0:
   $$\text{IoU}_0 = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} = \frac{1}{1+1+1} = \frac{1}{3}$$
   - mIoU:
   $$\text{mIoU} = \frac{\text{IoU}_0 + \text{IoU}_1}{2} = \frac{1/3 + 1/3}{2} = \frac{1}{3} \approx 0.333$$
2. **Pixel Accuracy (PA):**
   - Class 1:
   $$\text{PA}_1 = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{1}{2} = 0.5$$
   - Class 0:
   $$\text{PA}_0 = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{1}{2} = 0.5$$
   - mPA:
   $$\text{mPA} = \frac{\text{PA}_0 + \text{PA}_1}{2} = \frac{0.5 + 0.5}{2} = 0.5$$

4. Output Size Calculation for Transposed Convolution

Question: Calculate the output dimensions of a transposed convolution with the following parameters:

- Input size: 2×2
- Kernel size (K): 2×2
- Stride (S): 1
- Padding (P): 0

Solution: For transposed convolution, the output size (O) is given by:

$$O = S \cdot (I - 1) + K - 2P$$

where I is the input size.

Substituting the values:

$$O = 1 \cdot (2 - 1) + 2 - 2 \cdot 0 = 1 + 2 = 3$$

Therefore, the output feature map will be 3×3.