

Apellidos: _____

Nombres: _____

Fecha: 11/02/2025

Nota:

Indicaciones:

La Duración es de **40 minutos**.

La evaluación consta de **10 preguntas**.

Las preguntas de opción múltiple valen 1.5 punto.

1. ¿Cuál de la ventaja principal de DeepLabV3+ sobre PSPNet en la fusión de características?

- a) DeepLabV3+ emplea un módulo de decoder que únicamente integra high-level features para reducir la redundancia, lo que minimiza el costo computacional, pero a costa de una menor precisión en los bordes.
 - b) DeepLabV3+ combina el uso de ASPP para extraer características multiescala con un módulo de decoder que integra explícitamente low-level features para refinar las segmentaciones, superando las limitaciones de PSPNet en la precisión de los bordes.
 - c) DeepLabV3+ reemplaza el uso de atrous spatial pyramid pooling (ASPP) por una serie de capas fully-connected, permitiendo capturar dependencias globales sin fusionar explícitamente información local.
 - d) DeepLabV3+ se basa en el uso exclusivo de atrous convolutions sin incorporar mecanismos de fusión, lo que facilita el procesamiento de imágenes de alta resolución, pero limita la integración de escalas.
 - e) DeepLabV3+ introduce un módulo de atención que descarta low-level features, enfocándose únicamente en la información semántica profunda para mejorar la segmentación en regiones homogéneas.
2. ¿Cuál es la diferencia de utilizar InfoNCE en comparación con Triplet Loss?
- a) InfoNCE se diferencia de Triplet Loss al emplear una normalización softmax sobre un amplio conjunto de muestras negativas, lo que permite calcular la información mutua exacta entre las representaciones, en contraste con el uso de un margen fijo en Triplet Loss.
 - b) InfoNCE y Triplet Loss son esencialmente equivalentes, ya que ambos optimizan la separación entre muestras positivas y negativas sin incorporar mecanismos de normalización, haciendo que sus rendimientos sean comparables.

c) A diferencia de Triplet Loss, InfoNCE descarta el uso de muestras negativas, concentrándose únicamente en maximizar la similitud entre el anchor y el positivo, lo que simplifica el entrenamiento, pero reduce la diversidad en las representaciones.

d) InfoNCE incorpora una penalización adicional para evitar el sobreajuste en las muestras negativas, reduciendo la sensibilidad al tamaño del batch en comparación con Triplet Loss.

e) InfoNCE maximiza la probabilidad de identificar correctamente el par positivo entre múltiples distractores, operando en un espacio probabilístico que refleja la información mutua, a diferencia de Triplet Loss que se centra en empujar a las muestras negativas a distancias mayores sin considerar la distribución global.

3. Dada la función Softmax definida como

$$\text{Softmax}_i(z) = \frac{\exp(z_i/\tau)}{\sum_j \exp(z_j/\tau)}$$

donde τ es el parámetro temperature, ¿cuál es el efecto de aumentar el valor de τ en la distribución de probabilidades resultante?

- a) Aumentar τ intensifica las diferencias entre las probabilidades, elevando la probabilidad de la clase con mayor logit y haciendo la distribución más puntiaguda.
- b) Disminuir τ conduce a una distribución más uniforme, ya que los logits se normalizan de manera equitativa entre todas las clases.
- c) Aumentar τ suaviza la distribución, reduciendo las diferencias entre las probabilidades asignadas a cada clase y haciendo que la distribución sea más uniforme.
- d) Reducir τ suaviza la distribución, lo que genera mayor incertidumbre al asignar probabilidades similares a todas las clases.
- e) Modificar τ solo escala los logits sin alterar la forma relativa de la distribución, por lo que su efecto es meramente numérico.

4. Analice las implicaciones del diseño de la función loss en contrastive learning, discutiendo cómo elementos como el parámetro temperature y la definición de la distancia en el feature space afectan la capacidad del modelo para capturar similitudes y diferencias relevantes. [5 pts]

5. Se define Edge-aware smooth loss como:

$$L_s = |\partial_x d_t^*| \cdot e^{-|\partial_x I_t|} + |\partial_y d_t^*| \cdot e^{-|\partial_y I_t|}$$

donde:

- d_t^* es la profundidad predicha en la muestra t
- I_t es la imagen correspondiente en la muestra t
- ∂_x y ∂_y denotan las derivadas direccionales en los ejes horizontal y vertical, respectivamente.

Interprete correctamente cada componente de la función. Justifique.

[5 pts]

6. Se define Dynamic Margin Depth Loss (DMDL) para estimación de profundidad:

$$\mathcal{L}_{\mathcal{D}\mathcal{D}}(d, \hat{d}) = \frac{1}{N} \sum_{i=1}^N \max(0, |d_i - \hat{d}_i| - \tau_i)$$

con

$$\tau_i = \gamma \cdot \exp(-\beta \|\nabla I_i\|)$$

donde:

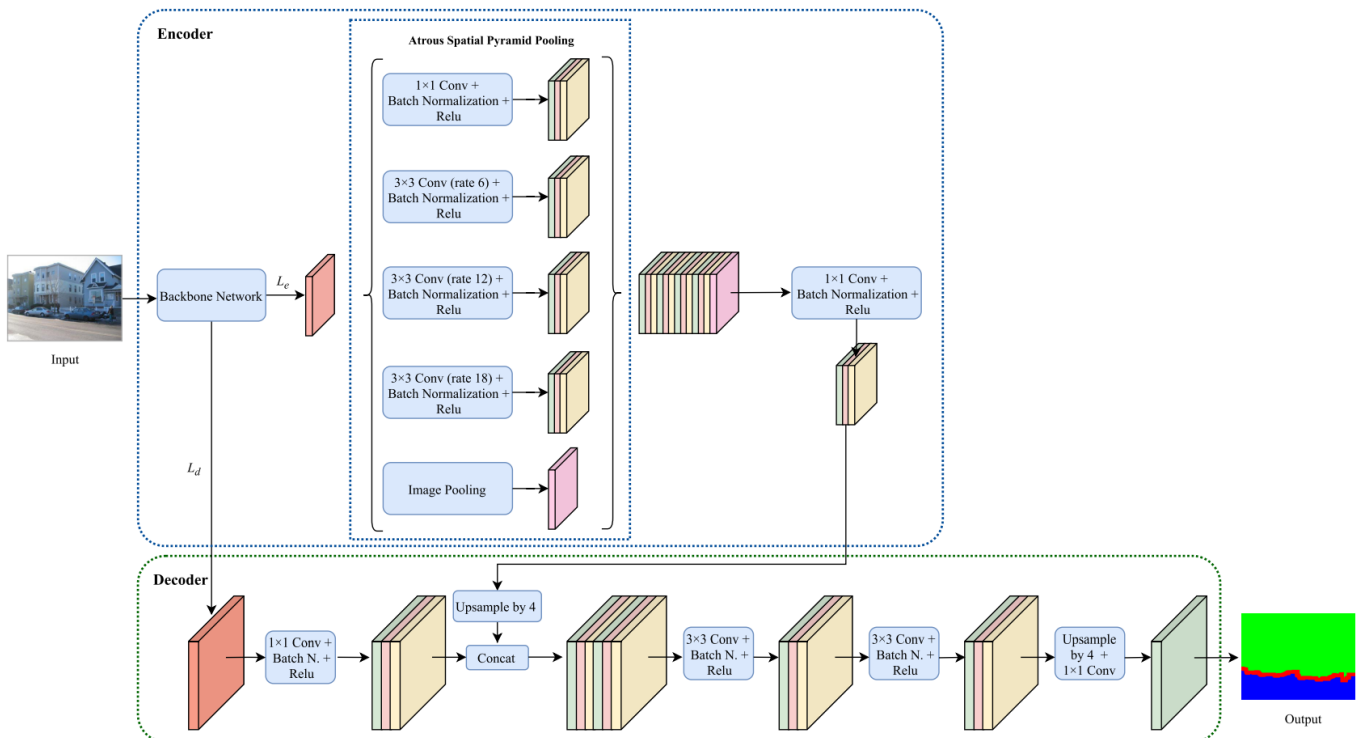
- d_i y \hat{d}_i son, respectivamente, la profundidad real y la predicha en el píxel i ,
- I es la imagen de entrada,
- ∇I_i representa el gradiente (local) de I en el píxel i ,
- γ y β son constantes.

Interprete correctamente cada componente de la función. Justifique.

[5 pts]

7. Implemente DeepLabV3+.

[5 pts]



Asuma que puede importar el backbone de su elección.