

Tarea 2 IMA543 Redes Neuronales Avanzadas de Aprendizaje Profundo

Benjamin Higuera, Felipe Pooley

28-05-2025

Ítems

Pregunta 1

Tomando el Dataset "dataset_flores" en la carpeta share del curso.

1. Cargue los conjuntos de imagenes de entrenamiento y testeo que entregados para esta evaluación. Genere una versión a escala de grises de las imágenes de ambos conjuntos. En la entrega de esta evaluación incluya una versión comprimida de las imágenes generadas.
2. Defina una arquitectura de autoencoder que pueda ser entrenada en base al conjunto en entrenamiento para colorear las imagenes a escala gris en base a las imágenes en colores correspondientes. Muestre un summary del modelo definido. El modelo debe tener un mínimo de 40 millones de parámetros.
3. Compile el modelo definido utilizando una función de perdida personalizada que combine dos términos en una suma ponderada. El primer término debe calcular el MSE entre la imagen a color generada por el auto encoder (salida del modelo autoencoder) y su correspondiente imagen original. El segundo término debe calcular el MSE entre la versión llevada a escala de grises de la imagen original (entrada del modelo autoencoder) y la versión llevada a escala de grises de la imagen a color generada por el autoencoder (salida del modelo autoencoder). La figura 1 muestra un diagrama de los términos mencionados.
4. Entrene el modelo autoencoder propuesto con la función de pérdida personalizada en el servidor khipu. Utilice al menos 100 épocas y salve el modelo entrenado. Suba evidencias del entrenamiento, el archivo .py utilizado y el modelo entrenado. Identifique tres imágenes del conjunto de testeo que tengan el peor MSE entre la imagen a color original y la generada por el modelo. Identifique también otras tres imágenes del conjunto de testeo que tengan el mejor MSE. Indique claramente cuáles fueron las seis imágenes identificadas y para cada una de ellas muestre la imagen en gris, la original y la generada por el modelo.

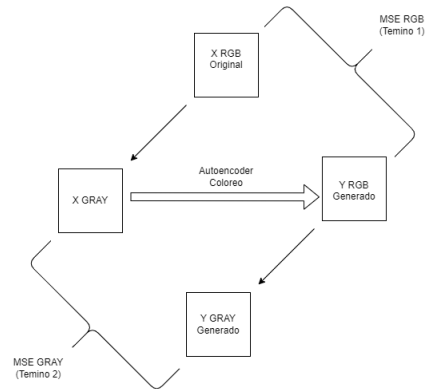


Figure 1: Diagrama función de pérdida

- Repita el experimento con otras 2 funciones de pérdida a su elección (MSE, PSNR, MAE, Chairbonner)

Pregunta 2

Tomando el conjunto de frases en "spa.txt" recree el código de la clase 14 (corregido) y :

- Entrene el modelo al menos 15 épocas (ojalá 30), seleccionando dos de los siguientes cambios e implementando las modificaciones correspondientes, y reporte rendimiento (accuracy en validación y BLEU en test). Explique brevemente los resultados.
 - Cambiar el número de cabezas de atención a 2 y 4 (de ser posible 16).
 - Reducir dense_dim de 2048 a 1024.
 - Incluir una capa adicional de encoder (es decir, stacking: pasar de 1 bloque de TransformerEncoder a 2 bloques secuenciales).
 - Aplicar dropout de 0.3 (en lugar de 0.5) solo en la salida del decoder antes de la capa densa.
- Prepare una función que recibe una oración en inglés y entrega la correspondiente traducción al español que genera el modelo entrenado. Tanto la entrada como la salida del modelo deben ser string. Utilice la función para generar la traducción de 10 frases aleatoriamente elegidos desde los datos de testeo.
- Cargue las oraciones en inglés del archivo eng-sentences.txt y utilice la función preparada en el inciso anterior para generar las traducciones al español que entrega su modelo entrenado. Discuta los resultados y, eventualmente, sugiera vías para mejorar los mismos.

Entregables

- **Acerca de la implementación :**
 - Códigos ocupados en Python (TensorFlow-Keras), de manera ordenada y legible.
 - Evidencias de ejecución en el servidor Khipu (pantallazos o logs de consola).
- **Informe en PDF** (máx. 9 páginas): resumen, métodos, resultados y discusión.

Criterios de Evaluación

Criterio	Descripción	Ponderación
Implementación	Correcta adaptación de modelos y técnicas	40%
Resultados	Claridad y profundidad del análisis	40%
Orden	Orden general y legibilidad del código	10%
Presentación	Organización, coherencia y visualización	10%

Notas Finales

- Se debe usar exclusivamente la versión del dataset disponible en la carpeta share del curso y todo el código debe ejecutarse correctamente en el servidor Khipu.
- El informe no tiene un formato específico pero lo que se espera es que no solo mencione métricas, sino también comente por qué un modelo funciona mejor o peor, un formato suficiente es:
 1. Descripción general de arquitecturas y modificaciones.
 2. Metodología de sus experimentos.
 3. Resultados
 4. Discusión y conclusiones.
- Use `screen` o `nohup` para ejecuciones largas. Se recomienda el uso de `verbose=2` al entrenar con `nohup`.
- No es obligatorio pero se recomienda usar distintos callbacks (checkpoints, early stopping con restore best weights, etc) para manejar sus experimentos y técnicas de aumentación si busca mejores resultados.

Plazo final de entrega: Viernes 6 de Junio 2025