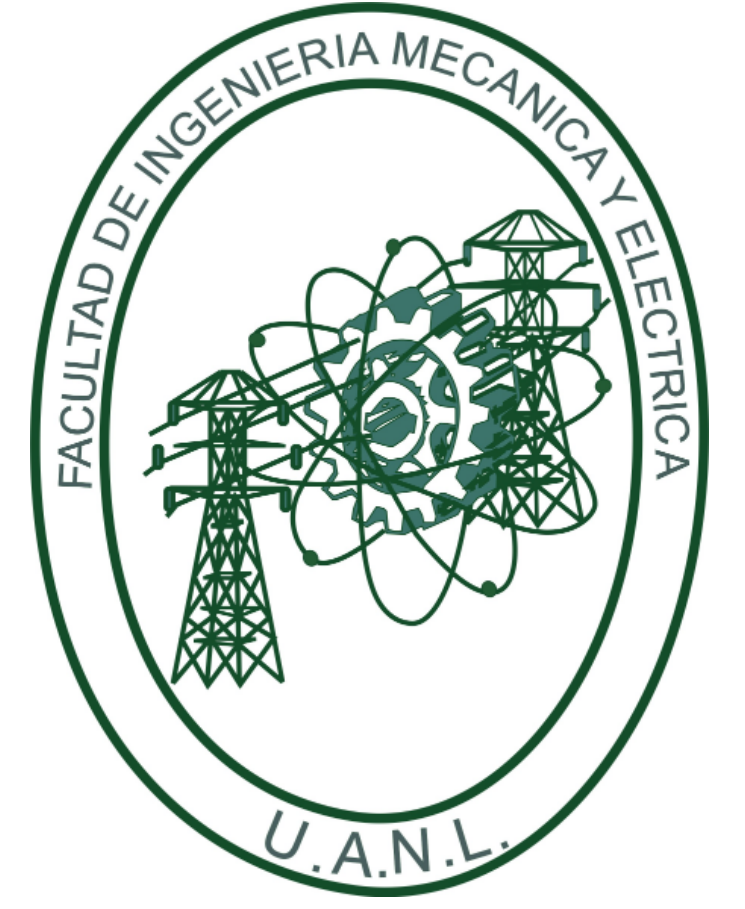


Detección de melanoma de piel mediante segmentación semántica

Redes Neuronales Convolucionales



Mario Alberto Flores Hernández
UANL - FIME

mario.floreshe@uanl.edu.mx

Introducción

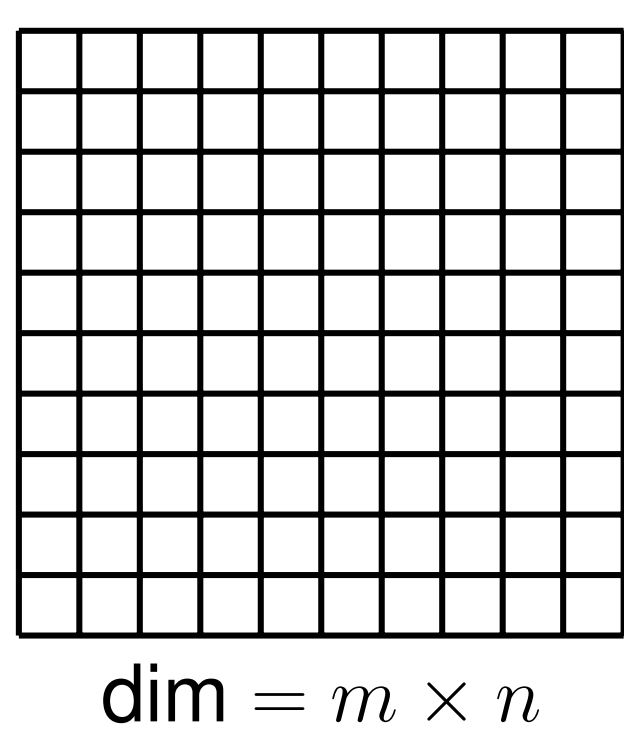
El **melanoma de piel** es un tipo de cáncer maligno que se propaga rápidamente a otras partes del cuerpo mediante el fenómeno de metástasis incrementando rápidamente su probabilidad de mortalidad. Una forma de amortiguar los riesgos de esta enfermedad es mediante la detección temprana de su presencia.

La inteligencia artificial brinda soluciones para la resolución de problemas que involucran visión computacional, una tecnología reciente que ha demostrado ser muy versátil es la red neuronal convolucional.

La **segmentación semántica** es una técnica de clasificación en la que se obtiene un mapa probabilístico con las regiones categorizadas de una imagen entrante, como resultado de una secuencia de filtros y transformaciones.

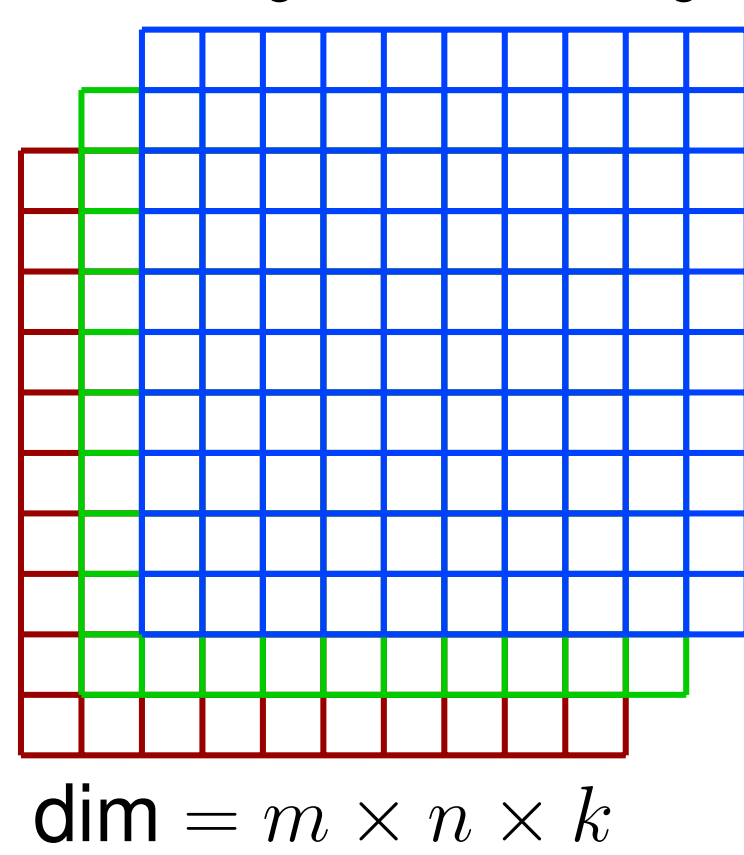
Características Dimensionales

Las imágenes a escala de grises y a color no comparten las mismas características dimensionales. Las imágenes a escala de grises se consideran un mapa bidimensional de intensidad de píxeles, mientras que las imágenes a color se consideran tridimensionales, siendo la tercera dimensión la capa de color.



dim = $m \times n$

Figura 1: Imagen a escala de grises.



dim = $m \times n \times k$

Figura 2: Imagen a color.

Imágenes Dermatológicas

Para entrenar un modelo de red neuronal convolucional de segmentación semántica es necesaria una base de datos que contenga dos subconjuntos: (a) imagen entrante, (b) salida deseada. La base de datos utilizada en esta implementación fue obtenida de la recopilación de imágenes dermatológicas internacional para la detección del melanoma [1].



(a) Imagen Entrante (b) Máscara Computada

Figura 3: Comparación de la imagen dermatológica entrante y su respectiva máscara segmentada.

Implementación

El tipo de arquitectura utilizada para la implementación se denomina FPN o red piramidal de características, como encodificador se utilizó el arreglo de capas de convolución RESNEXT.

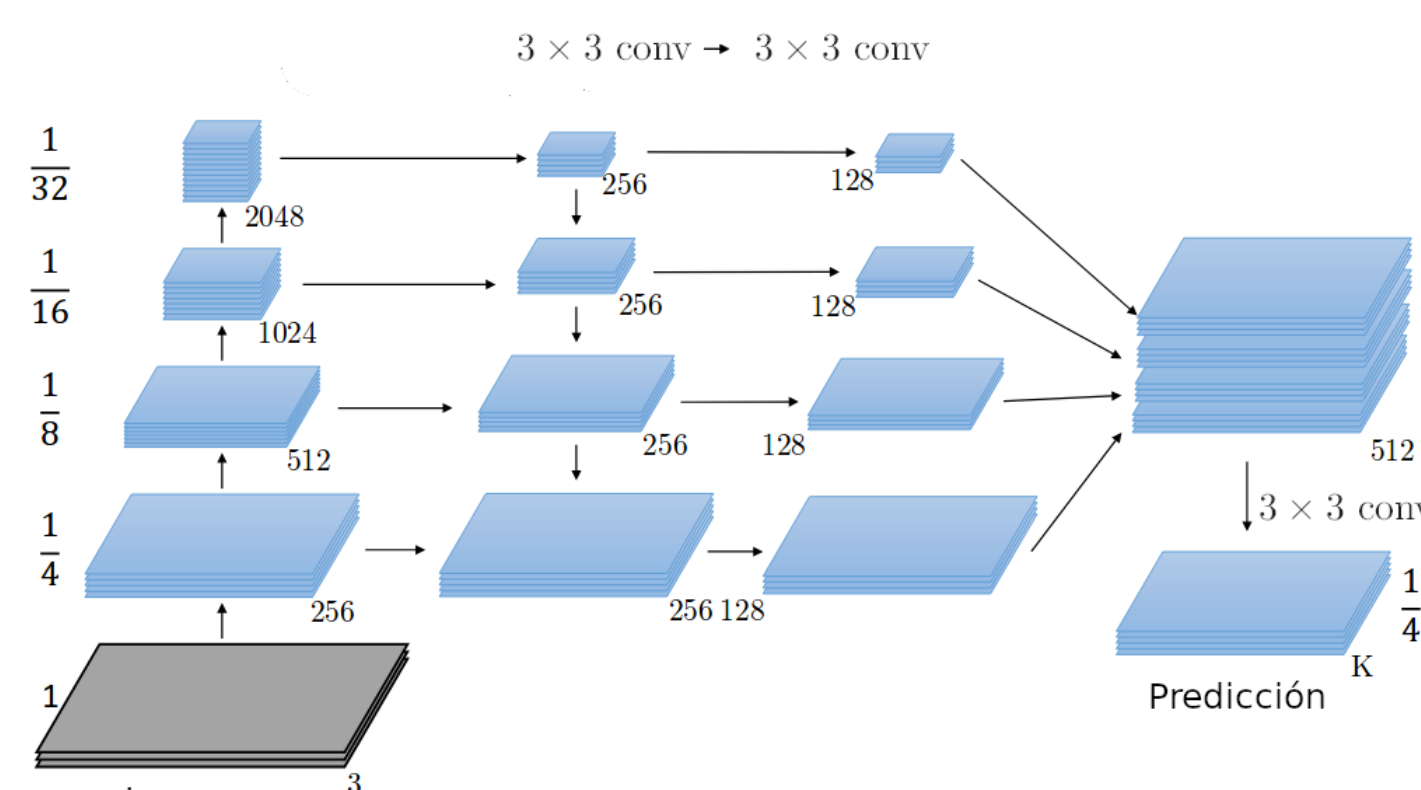


Figura 4: Representación de la arquitectura FPN [3].

El lenguaje utilizado para la implementación de la arquitectura de la red neuronal convolucional fue Python, debido a que es un lenguaje dinámico y que contiene la librería de Torch la cual es requerida para la manipulación de los subconjuntos de datos y el entrenamiento de modelos.

```
# train_model.py
max_score = 0
for i in range(0,40):
    print('\n Epoch: {}'.format(i))
    train_logs = train_epoch.run(train_loader)
    valid_logs = valid_epoch.run(valid_loader)
    if max_score < valid_logs['iou_score']:
        max_score = valid_logs['iou_score']
        torch.save(model, ('Model/'+encoder+'.pth'))
    print('Highest Score Model Saved: {}'.format(max_score))
```

Código 1: Bucle de entrenamiento.

Intensidad de Píxeles

El **filtro de umbral** es un tipo de procesamiento de imágenes que discretiza los valores de intensidad dependiendo de un **coeficiente límite**. Si el valor del píxel se encuentra por debajo del coeficiente límite, se reduce a la intensidad mínima; por el contrario si el valor rebaza el coeficiente límite, se amplifica a la intensidad máxima.

Coeficiente de Datos

El **coeficiente de datos** es una ecuación que determina el error entre dos matrices mediante la intersección de los píxeles coincidentes entre el total de píxeles. Este criterio es aplicado durante el proceso de entrenamiento.

$$CD = \frac{2|A \cap B|}{|A| + |B|} \quad (1)$$

Criterio de Jaccard

El **criterio de Jaccard** es una ecuación muy similar al **coeficiente de datos**, la diferencia radica en la penalización del error. El criterio de Jaccard penaliza aún más la diferencia entre las matrices y se utiliza en el proceso de validación.

$$CJ = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (2)$$

Resultados

Cuadro 1: Resumen de resultados por época.

Época	CD	CJ
1	0.2968	0.5616
2	0.2972	0.5546
3	0.2690	0.5879
4	0.2960	0.6389
5	0.2371	0.8415
6	0.0929	0.8626
7	0.0792	0.8626
8	0.0692	0.8798
9	0.0556	0.9017
10	0.0542	0.9039

Conclusión

Mediante los resultados obtenidos de los criterios de evaluación se puede observar un déficit en el error reportado por el coeficiente dados (CD) y un incremento en la similitud de acuerdo al criterio de Jaccard (CJ).

Trabajo a Futuro

- Implementar redes neuronales con segmentación semántica multi-categorica.

Repositorio de GitHub

Mediante el siguiente **código QR** puedes acceder a la implementación del proyecto y los datos generados de los experimentos.



<https://github.com/Btrox148/semantic-segmentation-skin>

Referencias

- [1] CODELLA, N. C. F., V. ROTEMBERG, P. TSCHANDL, M. E. CELEBI, S. W. DUSZA, D. GUTMAN, B. HELBA, A. KALLOO, K. LIOPYRIS, M. A. MARCHETTI, H. KITTLER y A. HALPERN (2019), «Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC)», *CoRR*, **abs/1902.03368**, 1902.03368 [cs.CV].
- [2] KINGMA, D. P. y J. BA (2014), «Adam: A method for stochastic optimization», *arXiv preprint arXiv:1412.6980[cs.LG]*.
- [3] LIN, T., P. DOLLÁR, R. B. GIRSHICK, K. HE, B. HARIHARAN y S. J. BELONGIE (2016), «Feature Pyramid Networks for Object Detection», *CoRR*, **abs/1612.03144**, 1612.03144.
- [4] YAKUBOVSKIY, P. (2020), «Segmentation Models Pytorch», https://github.com/qubvel/segmentation_models_pytorch.