

Implementación de un modelo de Segmentación de imágenes para detección de defectos en producción de acero y tumores cerebrales.

Pablo Jerez

pjerez2016@udec.cl

Gabriel Briceño

gbriceno2016@udec.cl

1 Introducción

La segmentación de imágenes es un proceso de gran utilidad que puede ser implementado bajo algoritmos de machine Learning y Deep learning. Este proceso es importante ya que se entrenan inteligencias artificiales con el objetivo de obtener la información más relevante de una imagen, eliminando lo que no sea de interés y así resolver un problema concreto. Esta segmentación generalmente es representada con una nueva imagen donde solo se puedan ver las "máscaras" (siluetas) de los objetos de interés.

Este proceso de segmentación se lleva a cabo en distintas áreas ya sean comerciales, productivas y médica. He aquí su importancia para buscar y generar soluciones para realizar el proceso de forma automática e ir beneficiando departamentos tales como los de mantenimiento, control, detección, producción y diagnóstico dentro de las organizaciones.

Es así como en el presente artículo se abordará el problema de automatización de detección y localización de defectos en la producción de acero además de tratar el problema segmentación de tumores cerebrales para poder visualizar su localización dentro del cerebro.

Ambos problemas pueden ser abordados de forma análoga bajo una misma arquitectura, ya que hay que implementar un algoritmo capaz de clasificar si tiene o no la anomalía, para posteriormente identificar el área de interés (segmentación). Para resolver estos problemas se utilizará una arquitectura que posee básicamente dos modelos. Esta consta de un clasificador ResNet para luego segmentar utilizando Unet y ResUnet. Estos dos últimos algoritmos serán implementados para comparar su rendimiento y lograr obtener el mejor modelo que se acomode al problema a resolver.

2 Materiales y metodología

2.1 Exploración y análisis de datos

El dataset que contiene imágenes de productos de acero corresponden a elementos manufacturados a una fabrica ubicada en San Diego, Estados Unidos. El conjunto de datos contiene 5.748 instancias de productos defectuosos y no defectuosos, con su respectiva máscara.

Para el problema de detección de tumores, el dataset contiene imágenes de resonancia magnética del cerebro junto con máscaras hechas de manera manual que indican la sección con las anomalías de tipo FLAIR. El conjunto de datos se obtuvo de The Cancer Imaging Archive (TCIA), y corresponden a 110 pacientes que formaron parte del proyecto The Cancer Genome Atlas (TCGA)

En ambos grupos las imágenes tienen una dimensión de 256x256.

Luego de preprocesar ambos datasets, se crean el grupo de entrenamiento y el de prueba (85% y 15% respectivamente) para entrenar los modelos.

2.2 Definición de Arquitectura

Para abordar ambos problemas se utilizará básicamente dos modelos: Un modelo clasificador capaz de discriminar las instancias con y sin el defecto, y un modelo de segmentación capaz de generar una máscara que muestre el área que comprenda el defecto sobre la imagen.

El modelo de clasificación a utilizar será una ResNet la cual será implementada con la librería de ResNet50 (aprendizaje por transferencia) de Keras y reentrenada en su nueva estructura en la sección de fully connected (TOP).

Posteriormente, los modelos de segmentación a comparar son el Unet y ResUnet. Unet es una estructura basada básicamente de capas convolucionales utilizando un codificador que reduce gradualmente la dimensión espacial de la capa de agrupación, para que luego su decodificador

restaure gradualmente los detalles y las dimensiones espaciales de la imagen. En cambio ResUnet utiliza la estructura de Unet, pero utilizando bloques residuales para superar los problemas acarreados del desvanecimiento de los gradientes.

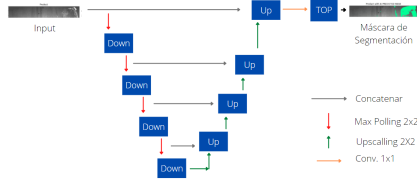


Figure 1: Estructura Unet
Fuente: Elaboración propia

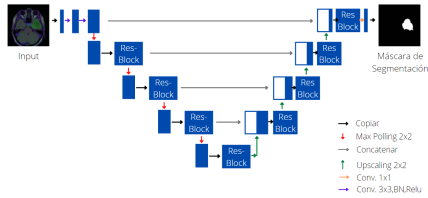


Figure 2: Estructura ResUnet
Fuente: Elaboración propia

La configuración de hiperparámetros generales a implementar en cada modelo son los siguientes:

Hiperparámetro	ResNet50	Unet	ResUnet
Learning rate	0.0001	0.05	
Dropout	0.3	0.1	0
Optimizador	Nadam	Adam	
Loss	Binary Cros.	Focal T.	
Batch size	16		

Table 1: Comparación de métricas de segmentación para Productos de acero.

2.3 Métricas de evaluación y comparación

Para evaluar y comparar el rendimiento de los modelos de segmentación se utilizan las métricas de Tversky y Focal Tversky(Loss). Ambas métricas se utilizan para evaluar modelos de segmentación tipo Unet y permiten manejar datasets desbalanceados y obtener resultados con mejor precisión y recall.

3 Experimentación y Resultados

3.1 Métricas de clasificación

Los resultados obtenidos se ven en la siguiente tabla:

Métricas/Dataset	Productos A.	Tumores C.
F1-Score	0.7583	0.7741
Recall Score	0.9970	0.9896
Precision Score	0.6118	0.6357

Table 2: Resultados métricas de clasificación para ambos datasets.

Las matrices de confusión respectivas a cada uno de los datasets se aprecian en las siguientes imágenes:

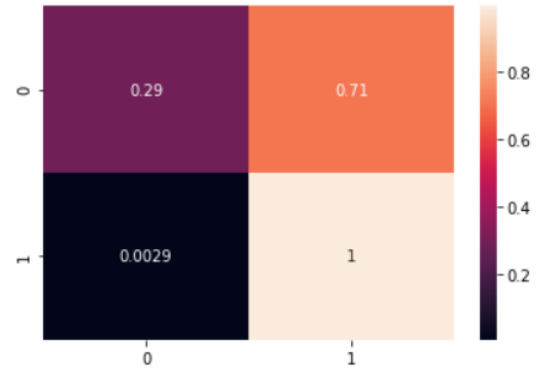


Figure 3: Matriz de confusión Productos de acero
Fuente: Elaboración propia

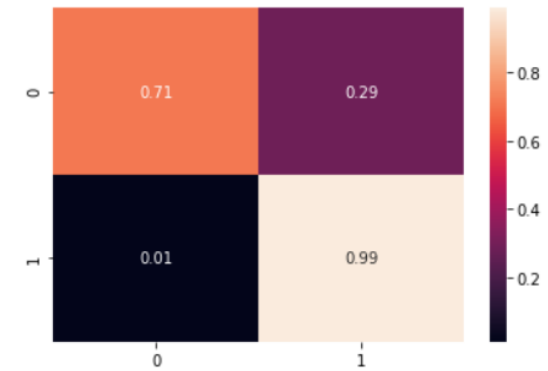


Figure 4: Matriz de confusión Tumores cerebrales
Fuente: Elaboración propia

3.2 Métricas de segmentación

Los resultados obtenidos del modelo luego de la implementación de modelos Unet y ResUnet se aprecian en las siguientes tablas:

Métricas/Modelo	Unet	ResUnet
Loss	0.5922	0.5340
Val loss	0.6192	0.5228
Tversky	0.4990	0.5645
Val tversky	0.4690	0.5769

Table 3: Comparación de métricas de segmentación para Productos de acero.

Métricas/Modelo	Unet	ResUnet
Loss	0.2922	0.2474
Val loss	0.3353	0.2817
Tversky	0.8047	0.8427
Val tversky	0.7659	0.8149

Table 4: Comparación de métricas de segmentación para Tumores cerebrales.

El comportamiento de las curvas de las métricas para cada epoch se visualizan en los siguientes gráficos:

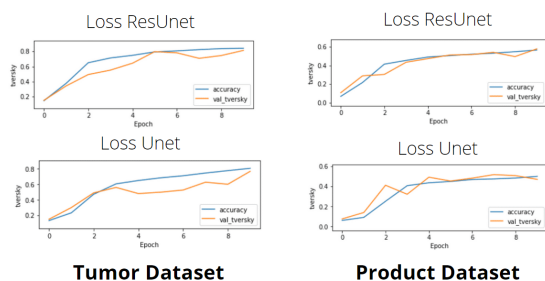


Figure 5: Curvas Accuracy de modelos Unet y ResUnet

Fuente: Elaboración propia

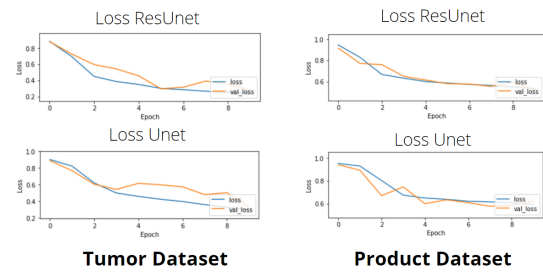


Figure 6: Curvas Loss de modelos Unet y ResUnet
Fuente: Elaboración propia

De los resultados obtenidos se logra evidenciar un mejor desempeño para el modelo de ResUnet para resolver estos problemas de segmentación.

Algunas máscaras obtenidas para cada escenario se aprecian en las siguientes imágenes:

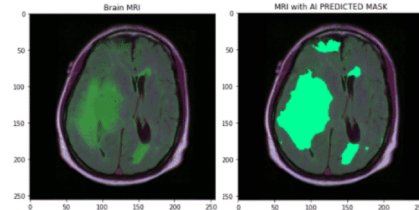


Figure 7: Máscara ResUnet tumor cerebral
Fuente: Elaboración propia

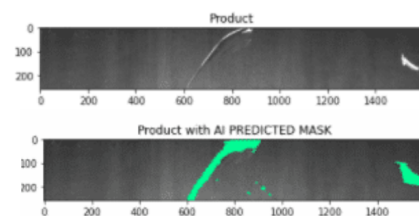


Figure 8: Máscara ResUnet producto de acero
Fuente: Elaboración propia

4 Discusión

La clasificación del Resnet50 tuvo un buen rendimiento al ajustar el threshold a 0.01. Es preferible tener falsos positivos que falsos negativos, para así identificar mejor los candidatos que podrían tener un defecto, descartando aquellos que notoriamente no lo tengan.

El modelo ResUnet presentó un mejor rendimiento que el Unet para el problema de segmentación de defectos de la producción de acero y para la segmentación de tumores cerebrales, mostrando mejores métricas de Tversky y Loss.

Sin embargo, en ambos escenarios siguen habiendo un número no menor de máscaras predecidas de manera imprecisa. Para el caso de los productos de acero, esto puede significar un inconveniente ya que puede dececharse más o menos material del debido, según lo que estime la IA. Para el caso de la segmentación de cerebros, esto puede generar ciertos problemas al indicar el grado de avance o la magnitud del tumor que padece la persona.

5 Conclusión

Para abordar problemas de segmentación en el ámbito médico de tumores y manufacturero de la producción de acero, el modelo ResUnet presentó un mejor rendimiento para segmentar imágenes con defectos. Las métricas Tversy y Focal Tversky(Loss) mostraron ser superiores en el ResUnet a diferencia del Unet.

Si bien la arquitectura final propuesta presenta algunas instancias con máscaras imprecisas que puedan derivar ciertos inconvenientes para los profesionales de cada área, no deja de ser un modelo con un rendimiento aceptable.

Al ser un modelo perfectible, sería interesante ver en futuros trabajos variaciones de esta arquitectura, ya sea probando con otros parámetros, optimizadores u otros modelos de segmentación y comparar sus resultados. También se pueden probar otras arquitecturas capaces de realizar tareas de segmentación similares, ya sea utilizando autoencoders o estructuras Gans que codifiquen y decodifiquen máscaras más exactas. Por otro lado, se pueden poner a prueba la arquitectura implementada en el estudio y los otros modelos anteriormente mencionados, a otros problemas similares tales como situaciones asociadas a la crisis sanitaria, segmentando mascarillas en rostros humanos.

6 Referencias

1. <https://www.kaggle.com/mateuszbuda/lgg-mri-segmentation>
2. <https://www.kaggle.com/abhiheetthakare5/tumor-sign-detection-segmentation-resunet>
3. <https://www.kaggle.com/sumantindurkhya/brain-mri-segmentation-unet>
4. <https://github.com/nabsabraham/focal-tversky-unet/blob/master/losses.py>
5. https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28
6. <https://arxiv.org/abs/1810.07842v1>
7. <https://www.udemy.com/courses/search/?q=machine%20learning>