

**UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ**

**Colegio de Ciencias e Ingenierías**

**Clasificación de volúmenes tridimensionales utilizando  
redes neuronales artificiales**

**Manolo Nicolás Noboa Lasso**

**Ingeniería en Ciencias de la Computación**

Trabajo de fin de carrera presentado como requisito  
para la obtención del título de  
Ingeniero en Ciencias de la Computación

Quito, 21 de diciembre de 2021

# **UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ**

**Colegio de Ciencias e Ingenierías**

**HOJA DE CALIFICACIÓN  
DE TRABAJO DE FIN DE CARRERA**

**Clasificación de volúmenes tridimensionales utilizando  
redes neuronales artificiales**

**Manolo Nicolás Noboa Lasso**

**Nombre del profesor, Título académico**

**Daniel Riofrio, PhD**

Quito, 21 de diciembre de 2021

## © DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en la Ley Orgánica de Educación Superior del Ecuador.

Nombres y apellidos: Manolo Nicolás Noboa Lasso

Código: 00132716

Cédula de identidad: 1722304340

Lugar y fecha: Quito, 22 de diciembre de 2021

## ACLARACIÓN PARA PUBLICACIÓN

**Nota:** El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en <http://bit.ly/COPETheses>.

## UNPUBLISHED DOCUMENT

**Note:** The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on <http://bit.ly/COPETheses>.

## RESUMEN

Presentamos diferentes arquitecturas de redes neuronales para la clasificación de volúmenes tridimensionales, incluyendo los detalles de los procedimientos de preprocesamiento de los datos. Los mencionados procedimientos y arquitecturas fueron aplicados para la clasificación de granos de suelo, y se evaluaron con diversas métricas que son presentadas en sus respectivas secciones. El origen de los volúmenes tridimensionales es una tomografía computacional sobre la cual se aplicó procesamiento de imágenes para obtener una representación matemática conocida como curvas de nivel o campos de distancia con signo. Se obtuvo un procedimiento intuitivo y visual que depende únicamente de los datos de la superficie y utiliza redes neuronales convolucionales para realizar la clasificación. Esto implica que dicho procedimiento puede ser aplicado para clasificar según los estándares requeridos sobre cualquier origen de datos del que se pueda obtener una superficie que describa su volumen. También se describen posibles nuevas arquitecturas que podrían ser implementadas para mejorar aún más los resultados obtenidos, con prometedores nuevos caminos por recorrer en el desarrollo de métodos para clasificación de volúmenes tridimensionales.

**Palabras clave:** Clasificación, Volúmenes, Redes Neuronales, Convolución, Superficie, Tridimensional, Procesamiento, Computacional

## ABSTRACT

We present different neural network architectures designed for classifying tridimensional volumes, including the details of the data preprocessing procedures. The mentioned procedures and architectures are applied in the classification of soil grains and were evaluated with various metrics that are stated in their respective sections. The origin of the tridimensional volumes is a computational tomography, on which we applied image processing to obtain a mathematical representation known as level sets or signed distance fields. A visual and intuitive procedure was obtained that depends only on the surface data and uses convolutional neural networks to obtain the classification. This procedure can be applied for classification according to the required standards over any data origin from which a surface describing the volumes can be obtained. Furthermore, additional architectural possibilities are described that could be implemented to improve the obtained results, with promising new paths to explore in the development of methods for classifying tridimensional volumes.

**Keywords:** Classification, Volumes, Neural Networks, Convolution, Surface, Tridimensional, Processing, Computational

## TABLA DE CONTENIDO

1. Introducción.....	10
2. Curvas de nivel (SDF) .....	12
2.1. Arquitectura de la red.....	13
2.2. Matriz completa.....	15
2.3. Matriz corregida .....	18
2.4. Matriz cercana a la superficie.....	20
3. Muestreo de Superficie.....	22
3.1. Coordenadas cartesianas .....	23
3.2. Coordenadas esféricas .....	23
4. Conclusión .....	25
REFERENCIAS BIBLIOGRÁFICAS.....	28
ANEXO A: Alteraciones en el procesamiento de imágenes .....	29
ANEXO B: Clases de errores .....	31
ANEXO C: Gráficas de entrenamiento y validación.....	33

## ÍNDICE DE TABLAS

Tabla 1: Disposición de las clases en la muestra.....	11
Tabla 2: Precisión (Accuracy), Área bajo la curva ROC (AUC) y Recall de la clasificación sobre la matriz completa.....	15
Tabla 3: Resultados de la clasificación sobre la matriz corregida .....	19
Tabla 4: Resultados del experimento 3 para $\mu = 0$ , $\alpha = 0.75$ y $\beta = 0.25$ .....	21
Tabla 5: Resultados para la clasificación usando coordenadas cartesianas .....	23
Tabla 6: Resultados de la clasificación para el caso de coordenadas esféricas.....	24

## ÍNDICE DE FIGURAS

Figura 1: Valores de la matriz SDF en un corte bidimensional con el contorno del grano .....	12
Figura 2: Contornos correctos (izquierda) e incorrectos (derecha) de un grano en un corte bidimensional .....	12
Figura 3: Proceso de convolución con un kernel de 3x3. ....	13
Figura 4: Arquitectura de la red neuronal convolucional.....	14
Figura 5: Curva de entrenamiento (azul) y validación (naranja) de la métrica recall en el caso de etiquetas múltiples para 100 épocas con sus respectivas desviaciones estándar (área sombreada).....	16
Figura 6: Curva de entrenamiento del segundo experimento .....	20
Figura 7: Curvas de nivel para un corte bidimensional completo (izquierda) y cercano a la superficie (derecha) .....	21
Figura 8: Curva de entrenamiento para el caso de coordenadas esféricas .....	24
Figura 9: Primeros pasos del método de Vlahinic et al aplicado sobre la muestra ..	29

# 1. Introducción

La representación de objetos como modelos tridimensionales presenta a las distintas áreas de la ciencia incontables nuevas posibilidades, así como una ventaja para permitirnos la visualización y cuantificación de distintos fenómenos físicos. En la medicina, se puede aplicar a la creación de imágenes musculoesqueléticas, permitiendo el fácil diagnóstico de fracturas y traumas, así como la identificación de masas de tejidos blandos (The Additional Diagnostic Value of the Three-dimensional Volume Rendering Imaging in Routine Radiology Practice, 2019). Los volúmenes tridimensionales también pueden ser caracterizados, permitiéndonos usarlos en simulaciones computarizadas que nos indiquen, por ejemplo, las características de un determinado suelo. Sin embargo, los métodos utilizados para obtener dichas representaciones no han conseguido que todas contengan fielmente las características de las muestras. Este documento analiza la posibilidad de clasificar eficientemente los volúmenes tridimensionales obtenidos para encontrar aquellos que modelan apropiadamente a las muestras.

Los datos utilizados para este documento provienen de la tomografía de una muestra de suelo de la región ecuatoriana de Jaramijó, sobre la cual se aplicó procesamiento de imágenes para obtener modelos matemáticos conocidos como curvas de nivel o campos de distancias con signo (SDF por sus siglas en inglés). El método aplicado para obtener estos modelos está basado en el propuesto por Vlahinic et al. en el artículo, "Towards a more accurate characterization of granular media: extracting quantitative descriptors from tomographic images" (2014) con ciertas alteraciones explicadas a detalle en el Anexo A.

Los resultados del procesamiento de imágenes presentaron diversas anomalías, las cuales fueron agrupadas en 6 clases observables: *Valles hondos*,

*Puntas pronunciadas, Caras planas, Dos Volúmenes Separados, Dos Volúmenes Unidos y Forma arbitraria.* La Tabla 1 muestra el total de muestras junto el porcentaje de aparición de cada clase de error, y en el anexo B se encuentran gráficos de ejemplo de cada una de las clases de error. Las clases observadas no son excluyentes entre sí, pudiendo un volumen presentar por ejemplo tanto caras planas como valles hondos.

<b>Clase de error</b>	<b>Porcentaje de aparición</b>
Formas arbitrarias	7.5%
Puntas pronunciadas	31.5%
Dos volúmenes unidos	12.25%
Dos volúmenes separados	1%
Valles hondos	26.25%
Caras planas	9.25%
Sin errores	41.5%
Numero de muestras	2950

*Tabla 1: Disposición de las clases en la muestra*

Los resultados clasificados son usados para el entrenamiento supervisado de cinco arquitecturas de redes neuronales convolucionales, basadas en dos tipos de datos de entrada. Las tres primeras arquitecturas toman como dato de entrada una matriz SDF, mientras que las dos últimas utilizan una técnica desarrollada de mapeo de superficie. Los detalles de cada arquitectura, incluyendo su preprocesamiento, entrenamiento, y sus resultados se muestran en las secciones a continuación.



superficies de nivel obtenidas para dos distintas distancias de la misma matriz SDF son equidistantes para todos sus puntos, sin embargo, dados los errores producidos por la naturaleza iterativa de la generación de la matriz SDF se observa que esta propiedad no siempre se cumple (véase Figura 2).

## 2.1. Arquitectura de la red

Los datos objetivo serán de dos tipos: el arreglo de clases de errores descrito anteriormente (clasificación de etiquetas múltiples), y una etiqueta binaria. La etiqueta binaria de cada volumen representa si este contiene o no alguna de las clases de errores. Para la clasificación se generan dos modelos, basados en las arquitecturas comunes de redes neuronales convolucionales. Estas están diseñadas para identificar jerarquías espaciales en las características de entrada mediante diversas operaciones (Yamashita, Nishio, Kinh Gian Do, & Togashi, 2018), entre ellas el proceso de convolución (véase Figura 3). Ambos modelos utilizan como función objetivo la entropía binaria cruzada junto con la activación sigmoide en la capa de salida.

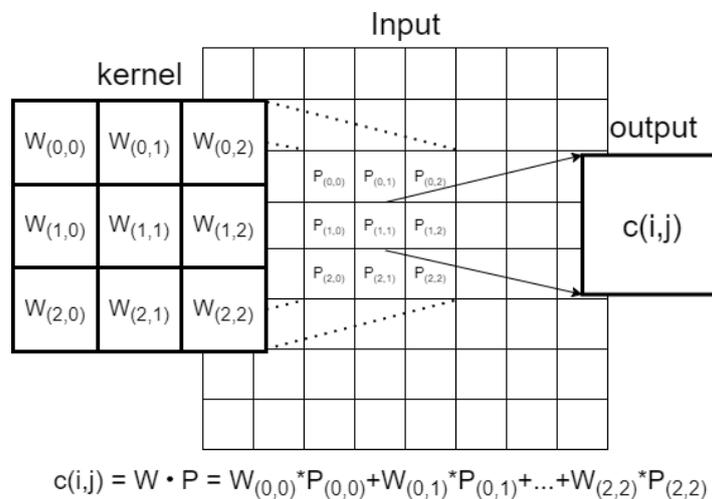


Figura 3: Proceso de convolución con un kernel de 3x3

Ambos modelos contienen tres capas ocultas convolucionales, y varían únicamente en la cantidad de neuronas de la capa de salida. Diversas arquitecturas fueron probadas, siendo la mostrada en la Figura 4 la que mejores resultados mostró respecto a los siguientes parámetros de comparación:

- Valores de promedio y desviación estándar de las métricas en un sistema de validación cruzada.
- Estabilidad en las gráficas de las métricas en los periodos de entrenamiento, tiempo de convergencia, y tendencia al sobreajuste.

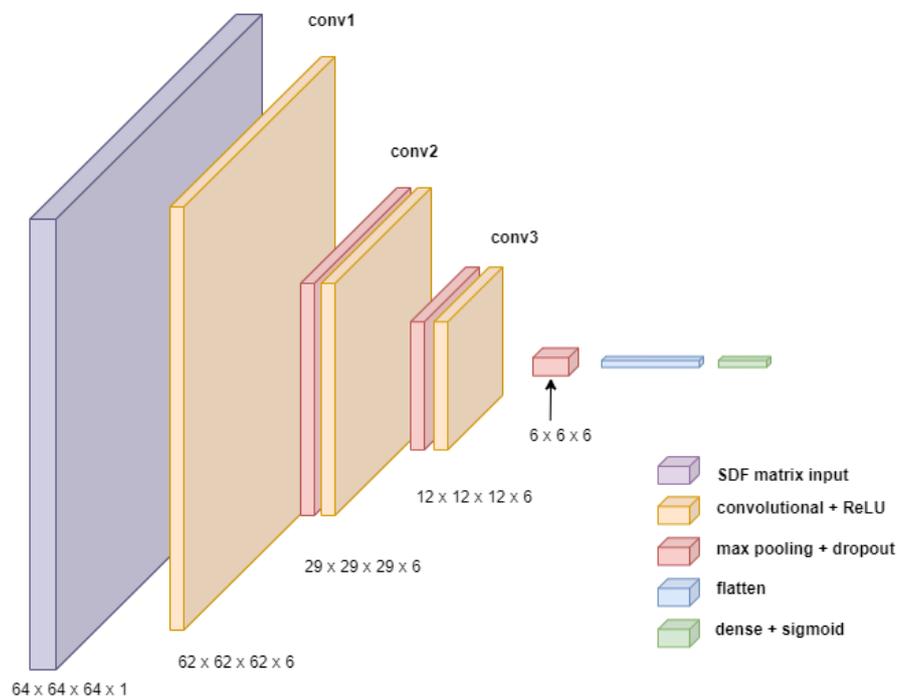


Figura 4: Arquitectura de la red neuronal convolucional

Las gráficas tanto de las métricas como de la función objetivo resultantes del entrenamiento y validación de los modelos que propone este documento se encuentran en el Anexo B.

## 2.2. Matriz completa

Para el primer experimento usamos directamente la matriz SDF con errores. Esto nos permitirá obtener un punto de partida con el cual comparar los resultados de los siguientes experimentos. El preprocesamiento de este experimento es sencillo, con únicamente dos pasos.

Primero, se escala la matriz SDF a un tamaño regular, para permitir la comparación uniforme con las demás matrices. Para esto utilizamos interpolación cúbica para convertir cada matriz, independientemente de su tamaño original, a una matriz cúbica con 64 voxels de cada lado. Segundo, normalizamos la matriz, alterando su escala de valores para que se encuentren en el rango de 0 a 1. Esta normalización se realiza independientemente para cada matriz, con el propósito de evitar que las diferencias de tamaño entre volúmenes generen tendencias equivocadas al momento de entrenar la red neuronal.

<i>Clasificación</i>	<i>N. salidas</i>	<i>Accuracy</i>	<i>AUC</i>	<i>Recall</i>
Binaria	1	0.70 ± 0.1	0.80 ± 0.069	0.811 ± 0.151
Etiquetas Múltiples	6	0.875 ± 0.021	0.76 ± 0.095	0.147 ± 0.110

*Tabla 2: Precisión (Accuracy), Área bajo la curva ROC (AUC) y Recall de la clasificación sobre la matriz completa.*

Para los cálculos de las métricas se calculan los valores de la matriz de confusión, una vez por cada uno de los 10 cortes usados para la validación cruzada. En el caso de la clasificación utilizando las clases de errores, cada uno se cuenta como un acierto o una falla, y con estos se calculan las demás métricas. Para la métrica *Recall*, se utiliza un único umbral para la binarización de las predicciones de la red, establecido en 0.5, para después calcular su valor dividiendo el número de

verdaderos positivos para la suma de los verdaderos positivos y falsos negativos. Para la métrica *AUC* (acrónimo en inglés para *Area Under the Curve*) se utilizan 200 umbrales para el cálculo de los puntos de la curva ROC (acrónimo en inglés para *Receiver Operating Characteristic*). Esto nos permite evaluar el desempeño de la red dado que compara los resultados obtenidos con los resultados reales. En la Tabla 2 se encuentran compilados los resultados de dichas métricas para el experimento propuesto en esta sección, y en el Anexo C se encuentran las gráficas de entrenamiento y validación para todos los experimentos analizados en este documento.

Los resultados demuestran la capacidad de la red neuronal de identificar las características deseadas, mas no a un nivel confiable. Adicionalmente, comparando la métrica *AUC* entre la clasificación binaria y la de etiquetas múltiples observamos que la red no pudo identificar correctamente a que clase pertenecían los errores encontrados en la matriz.

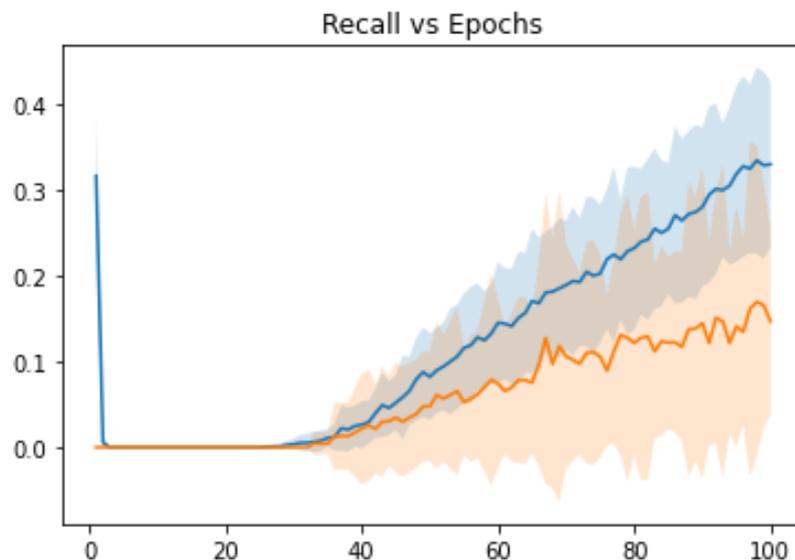


Figura 5: Curva de entrenamiento (azul) y validación (naranja) de la métrica recall en el caso de etiquetas múltiples para 100 épocas con sus respectivas desviaciones estándar (área sombreada)

Ahora analicemos la métrica *Recall*. Observando la curva de entrenamiento de esta métrica en la Figura 5 podemos ver que inicialmente empieza en 0, y solo empieza a incrementarse tras varias épocas, lo que nos indica que los valores de las predicciones de la red tienden a ser bastante bajos, siendo posteriormente categorizados como negativos. En la misma gráfica, vemos también que la desviación estándar incrementa con el paso de las épocas en lugar de disminuir, dando a entender que el modelo es inestable.

Para corregir los errores evidenciados, se generan algunas hipótesis acerca del origen de estos. La primera hipótesis asume que el origen son los datos preprocesados de las matrices SDF, y lo relaciona directamente con las fallas encontradas estas observadas en la Figura 2. Al existir zonas donde no existe una variación considerable en la matriz, la operación de convolución podría estar obteniendo mapas de características errados de estas secciones, ocasionando las predicciones con valores bajos. En el caso de etiquetas múltiples donde los valores de salida contienen mayor cantidad de ceros, el entrenamiento no consigue corregir los pesos del kernel de convolución y las curvas de entrenamiento empiezan a evidenciar sobreajuste. Esta hipótesis será comprobada en la sección 2.3.

La segunda hipótesis asume que el origen es la alta cantidad de datos por cada volumen, lo cual lleva a que la red disperse su entrenamiento en zonas no tan relevantes de la matriz SDF, impidiendo que se enfoque en las zonas donde las características a clasificar son más evidentes.

La tercera hipótesis asume que el origen es el desbalance de los datos (ver Tabla 1) en la clasificación de etiquetas múltiples, lo cual provoca que tienda a predecir todos los valores como negativos. Un experimento sencillo se ejecutó para probar esta hipótesis: entrenar la red únicamente con las clases con mayor

porcentaje de aparición (*puntas pronunciadas* y *valles hondos*). Los resultados encontrados no se diferenciaron de los presentados en el experimento anterior, descartando temporalmente esta hipótesis. Alternativamente se intentó aplicar diversas técnicas de aumento de datos, más las matrices generadas mostraban diferentes combinaciones de errores, algunas nuevas clases de errores (por ejemplo, superficies abiertas) y algunos no contenían ningún error. Experimentos para comprobar las hipótesis uno y dos son presentadas en las secciones a continuación.

### **2.3. Matriz corregida**

Para el segundo experimento utilizaremos una matriz SDF sin los errores observados en la Figura 2. Este proceso consta de tres pasos, el primero es mediante interpolación encontrar una superficie definida por triángulos en el nivel de distancia 0 (por lo tanto, la superficie real del grano). Se escoge el nivel 0 dado que el procesamiento de imágenes converge cuando los valores cercanos al nivel 0 dejan de variar, y siendo este nivel la guía para la clasificación.

Con esta superficie tridimensional, se podría recrear la matriz SDF calculando la mínima distancia de cada punto en la matriz hacia la superficie, y dándole el signo que tenga la magnitud de la diferencia entre el punto y la normal más cercana. En lugar de esto, se utilizó el método propuesto en el artículo “DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation” en el cual se propone no utilizar toda la superficie, si no una nube de puntos tomada alrededor de esta y empleando una red construir la matriz SDF (Park, Florence, Straub, Newcombe, & Lovegrove, 2019) con el fin de optimizar el tiempo requerido de preprocesamiento. Este proceso también incluye la alineación de la superficie con

sus ejes principales y su escalado a un espacio de rango  $[-1, 1]$ . Finalmente se aplican los mismos pasos de la sección 2.2, y se entrenan las mismas redes definidas anteriormente.

<i>Clasificación</i>	<i>N. salidas</i>	<i>Acc</i>	<i>AUC</i>	<i>Recall</i>
Binaria	1	$0.651 \pm 0.053$	$0.755 \pm 0.077$	$0.413 \pm 0.188$
Etiquetas Múltiples	6	$0.85 \pm 0.03$	$0.761 \pm 0.044$	$0.276 \pm 0.075$

*Tabla 3: Resultados de la clasificación sobre la matriz corregida*

Analizando los resultados compilados en la Tabla 3, podemos observar que efectivamente la clasificación de etiquetas múltiples ha mejorado considerablemente en la métrica *Recall*. Al contrario, se observa una desmejora en los resultados de la clasificación en las otras métricas, en especial en la clasificación binaria. Podemos concluir que el modelo estaba aprendiendo no de los errores presentados en la superficie de los volúmenes, sino más bien en los errores encontrados en la generación de las matrices SDF. Este no es un comportamiento deseable, pues la red neuronal solo podría clasificar errores en volúmenes encontrados con el método de Vlahinic et al., y no sobre los datos de la superficie, por lo cual es preferido el preprocesamiento de esta sección, siendo aplicable sobre cualquier superficie tridimensional.

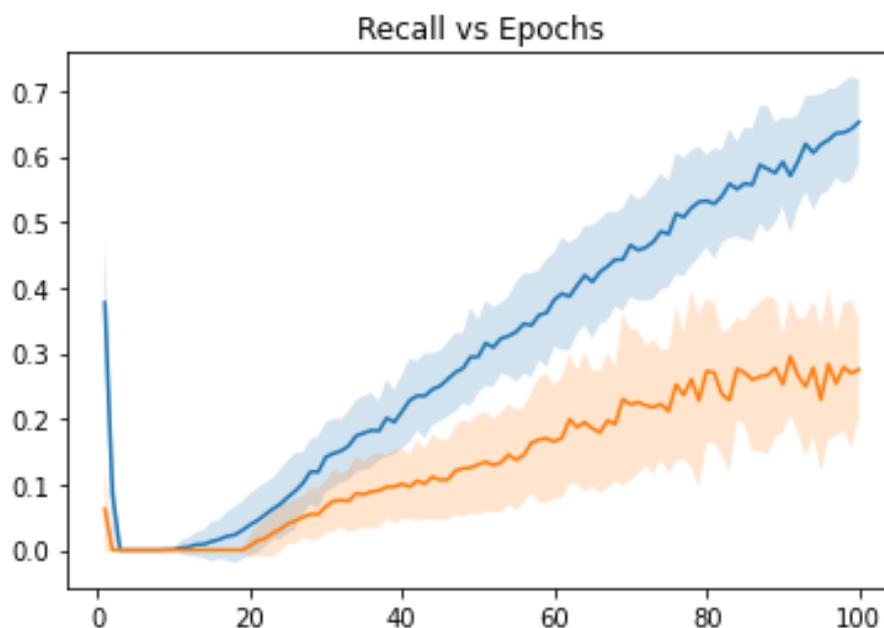


Figura 6: Curva de entrenamiento del segundo experimento

Observando en la Figura 6 la curva de entrenamiento para la métrica *Recall*, notamos que se recupera mucho más rápido del problema observado en la Figura 5 sin presentar el incremento en la desviación estándar, más aún existe sobreajuste. Adicionalmente también existe diferencia entre la clasificación binaria y la de etiquetas múltiples, pero es mucho menos marcada que en el experimento anterior. Con esto consideramos comprobada la hipótesis uno, y las siguientes secciones se enfocan en experimentar basados en la hipótesis dos.

## 2.4. Matriz cercana a la superficie

Para el tercer experimento utilizamos el preprocesamiento de la sección anterior, agregándole un paso adicional: pasamos cada valor de la matriz por una función por partes, de acuerdo con la cual los valores mayores al nivel  $\alpha$  son reemplazados por un valor de relleno  $\mu$ , así como los menores al nivel  $-\beta$ , dejando intactos los valores entre ambos niveles. En la Figura 7 se grafica un ejemplo de dicho procedimiento

aplicado sobre un corte bidimensional de un grano con *puntas pronunciadas*, para los valores de  $\alpha = \beta = 0.1$ .

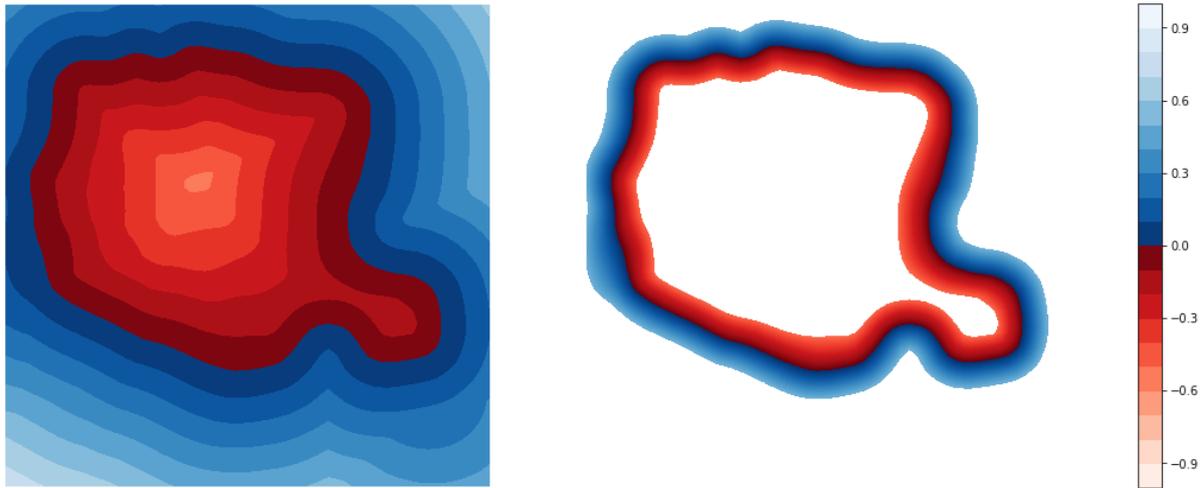


Figura 7: Curvas de nivel para un corte bidimensional completo (izquierda) y cercano a la superficie (derecha)

Así se espera que la red neuronal se entrene únicamente sobre los niveles cercanos a la superficie, evitando que los niveles que no contienen las características deseadas tengan influencia en los resultados de la clasificación. Por ejemplo, observemos el nivel -0.3 en las curvas de nivel izquierdas de la Figura 7. Este nivel y los inferiores ya no contienen la *punta pronunciada* que se puede observar en el nivel de la superficie.

Clasificación	N. salidas	Acc	AUC	Recall
Binaria	1	$0.631 \pm 0.137$	$0.705 \pm 0.128$	$0.440 \pm 0.209$
Etiquetas Múltiples	6	$0.850 \pm 0.020$	$0.718 \pm 0.059$	$0.227 \pm 0.112$

Tabla 4: Resultados del experimento 3 para  $\mu = 0$ ,  $\alpha = 0.75$  y  $\beta = 0.25$

Los resultados no demuestran una mejora confiable, aún probando con distintos valores de  $\mu$  (0, 10, -10),  $\alpha$  (0.1, 0.5, 0.75) y  $\beta$  (0.1, 0.25, 0.5), siendo los mejores los compilados en la Tabla 4. Analizando también las propiedades de la nueva entrada, se observa que las matrices tienen distintas cantidades de puntos a ser tomados en cuenta dependiendo de la forma de la superficie que estos representen. Por estos motivos, y considerando que esta transformación es irreversible, se considera mejor no aplicarla. Dado que no se ha diseñado aún una manera de reducir los datos de otra manera, las siguientes secciones exploran otras transformaciones para experimentar sobre la hipótesis dos.

### **3. Muestreo de Superficie**

Para el cuarto experimento reducimos el espacio de entrada a únicamente los puntos que se encuentren en la superficie de nivel 0. Estas son las mismas superficies usadas en los primeros pasos del preprocesamiento de la sección 2.3. Esto nos permite reducir la dimensionalidad de la entrada a  $N$  vectores de longitud 3, donde  $N$  es el número de puntos que se desea extraer, caracterizado por sus coordenadas. Teniendo la representación mínima de la superficie, vamos aumentando la densidad de los puntos que la componen hasta alcanzar los  $N$  puntos deseados. Para finalizar el preprocesamiento, ordenamos los puntos obtenidos en orden ascendente de su ángulo de elevación y azimut. Así obtenemos una especie de espiral que recorre los puntos de la superficie. Probamos este preprocesamiento utilizando coordenadas cartesianas y esféricas, usando mil puntos por grano.

Posteriormente, usamos modelos similares a los presentados en la sección 2.1, cambiando las capas de convolución tridimensionales por convoluciones unidimensionales. Cada componente del vector consta de su kernel individual, teniendo cada capa convolucional 3 kernels.

### 3.1. Coordenadas cartesianas

<i>Clasificación</i>	<i>N. salidas</i>	<i>Acc</i>	<i>AUC</i>	<i>Recall</i>
Binaria	1	0.714 ± 0.066	0.745 ± 0.081	0.61 ± 0.121
Etiquetas Múltiples	6	0.868 ± 0.015	0.604 ± 0.045	0.193 ± 0.082

*Tabla 5: Resultados para la clasificación usando coordenadas cartesianas*

Para el caso de las coordenadas cartesianas, se normalizan los vectores alterando su escala a  $[0, 1]$ , donde el mayor valor se convierte en 1 y el menor en 0. Observando los resultados compilados en la Tabla 5, descartamos el uso de coordenadas cartesianas como datos de entrada, pues tanto el *AUC* y el *Recall* se encuentran por debajo de los resultados de secciones anteriores.

### 3.2. Coordenadas esféricas

Para el caso de coordenadas esféricas, se normaliza cada componente de los vectores alterando su escala a  $[0, 1]$ . Los valores máximos y mínimos se toman en cuenta únicamente para el componente del radio, dado que para los componentes de elevación y azimut ya se conoce su escala original ( $[-\pi/2, \pi/2]$  y  $[-\pi, \pi]$  respectivamente).

<i>Clasificación</i>	<i>N. salidas</i>	<i>Acc</i>	<i>AUC</i>	<i>Recall</i>
Binaria	1	$0.746 \pm 0.077$	$0.805 \pm 0.076$	$0.654 \pm 0.140$
Etiquetas Múltiples	6	$0.888 \pm 0.014$	$0.799 \pm 0.054$	$0.315 \pm 0.095$

Tabla 6: Resultados de la clasificación para el caso de coordenadas esféricas

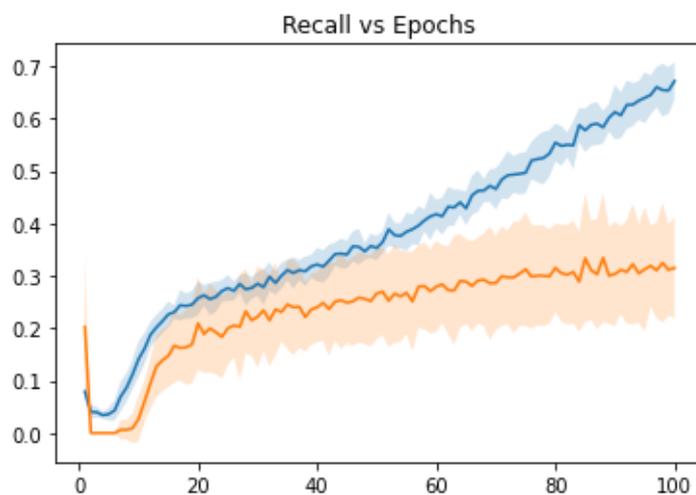


Figura 8: Curva de entrenamiento para el caso de coordenadas esféricas

Los resultados de este experimento se encuentran compilados en la Tabla 6. Analizando los resultados para la clasificación de etiquetas múltiples, si bien es cierto aún no alcanzan rangos aceptables, estos son mejores que los obtenidos en cualquiera de los otros experimentos, en especial comparado con el de la sección 3.1. Adicional, en la curva de entrenamiento es visible el punto en el que empieza a ocurrir el sobreajuste, manteniendo las curvas de entrenamiento y validación una similitud sobre una mayor cantidad de épocas que las observadas en otras curvas (véase Figura 8).

## 4. Conclusión

Diversos métodos para la clasificación de volúmenes tridimensionales fueron probados y documentados, basados en dos diferentes espacios de características. En el espacio de características definido por matrices SDF, los resultados no alcanzan niveles confiables, y las variaciones entre experimentos son poco considerables. La causa puede ser atribuida precisamente al método en el que se basa la construcción de dicho espacio.

Los valores en una matriz SDF están relacionados únicamente a su distancia a la superficie, siendo esta la que define el espacio, y sus características no son identificables al ser iteradas a través de un espacio euclidiano. Por lo tanto, las operaciones convolucionales tradicionales no consiguen construir un mapa de características apropiado para luego ser utilizado como la base de la clasificación.

Podemos también aplicar esta lógica para comprender la causa de la mejora en los resultados observados en el muestreo de las superficies, en específico el caso de coordenadas esféricas. Sobre el espacio que construyen estas coordenadas, resulta más sencillo interpretar las características de las superficies y están más relacionadas a las características identificables. Por ejemplo, en el caso de *puntas pronunciadas* y *valles hondos*, estas características pueden ser identificadas con la siguiente proposición lógica: en un conjunto de puntos con ángulo de elevación y azimut cercanos, pero con radios muy distintos, será *punta pronunciada* si existen más radios con valor bajo que alto, de lo contrario, será *valle hondo*.

Sin embargo, los resultados de la clasificación sobre coordenadas esféricas tampoco alcanzan niveles confiables. Podemos generar dos hipótesis acerca del motivo: el método de extracción de puntos no alcanza a abarcar todas las

características observables en la superficie; o el orden de los puntos extraídos reduce la posibilidad de identificar una jerarquía espacial completa. A partir de estas hipótesis podemos diseñar nuevos experimentos para comprobarlas o refutarlas.

Uno de estos experimentos para comprobar la primera hipótesis podría basarse en la técnica de muestreo propuesta por Park et al., que propone encontrar el valor de distancia de Q puntos mediante *exploraciones laser virtuales* y vectores normales para conseguir una nube de puntos (DeepSDF: Learning Continuous Signed Distance Functions, 2019). A cada punto en esta nube se le podría agregar las coordenadas esféricas que le corresponden, y usar los vectores  $(Q, 4)$  para alimentar una red neuronal densamente conectada. Alternativamente, si se desea mantener cierta relación espacial entre los puntos obtenidos, se puede usar el ordenamiento propuesto en la sección 3, agregándole adicionalmente una segunda entrada: los puntos ordenados también sobre el eje previamente ignorado. Con esta adición, se puede usar redes convolucionales como la propuesta, o incluso redes recurrentes.

Otro experimento podría ser uno basado en el proceso de generación de superficies de nivel conocido como *Marching Cubes*. Este método utiliza una iteración tridimensional de celdas cúbicas en las cuales cada esquina representa un valor de la matriz SDF. Usa estos valores para generar un mapa que indique que tipo de polígonos requiere esta sección, para luego interpolar y encontrar los vértices de estos (Lopes & Brodlie, 2003). En lugar de interpolar, se podría utilizar el tipo de polígono como representación para identificar las características de cada sección, agregándole pesos entrenables a cada esquina, siendo esta la base para la clasificación.

También podemos relacionar el error al formato de salida. Dada la complejidad de los datos de entrada, la red no consigue asociar apropiadamente las características encontradas mediante las capas convolucionales con sus etiquetas de clasificación. Para superar este problema, se podría alterar tanto la arquitectura como los datos de salida para utilizar un *Autoencoder*. Esto aumenta la complejidad del método al generar el conjunto de datos, mas disminuye la complejidad del proceso de entrenamiento, y los resultados de la predicción permitirían determinar las zonas donde se encuentran las características deseadas.

## REFERENCIAS BIBLIOGRÁFICAS

- Bradley, D., & Roth, G. (2007). Adapting Thresholding Using the Integral Image. *Journal of Graphics Tools*, 13-21.
- Duran, A. H., Duran, M. N., Masood, I., Maciolek, L. M., & Hussain, H. (2019). The Additional Diagnostic Value of the Three-dimensional Volume Rendering Imaging in Routine Radiology Practice. *Cureus* .
- Lopes, A., & Brodlie, K. (2003). Improving the robustness and accuracy of the marching cubes algorithm for isosurfacing,. *IEEE Transactions on Visualization and Computer Graphics*, 16-29.
- Park, J. J., Florence, P., Straub, J., Newcombe, R., & Lovegrove, S. (Enero de 2019). DeepSDF: Learning Continuous Signed Distance Functions. Estados Unidos.
- Vlahinic, I., Andò, E., Viggiani, G., & Andrade, J. E. (2014). Towards a more accurate characterization of granular media: extracting quantitative descriptors from tomographic images. *Granular Matter*, 9-21.
- Yamashita, R., Nishio, M., Kinh Gian Do, R., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, 611-629.

## ANEXO A: Alteraciones en el procesamiento de imágenes

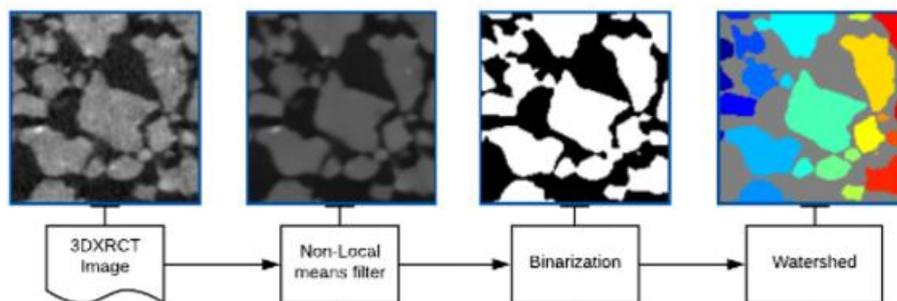


Figura 9: Primeros pasos del método basado en el propuesto por Vlahinic et al. aplicado sobre la muestra

El método propuesto por Vlahinic et al. consta de una serie de procedimientos a ser realizados en secuencia, uno de ellos siendo la binarización. Para este paso, el método originalmente propone encontrar un umbral mediante la observación del histograma de intensidades encontradas en la tomografía. (Towards a more accurate characterization of granular media: extracting quantitative descriptors from tomographic images, 2014).

Para la muestra utilizada en este documento, se observó que dicho histograma de intensidades no contenía un umbral reconocible, posiblemente debido a la presencia de elementos con distintas densidades dentro de la muestra o la baja resolución de la tomografía. En su lugar, utilizamos el procedimiento de binarización con umbral adaptable propuesto por Bradley & Rod (Adapting Thresholding Using the Integral Image, 2007).

Adicionalmente, para el procedimiento de separación de los granos, Vlahinic et al. proponen una transformada de distancia euclidiana para la generación de “picos y valles”, posteriormente aplicando la técnica de *Watershed*. En la muestra utilizada, sin embargo, este procedimiento no consiguió separar apropiadamente los granos más grandes de sus vecinos más pequeños, fusionándolos en un único grano. Para superar este inconveniente, utilizamos una transformada de distancia

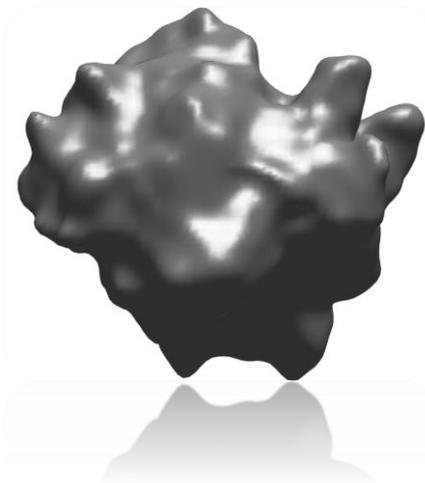
manhattan. Un ejemplo de la aplicación de estos pasos se encuentra ilustrado en la Figura 9.

Finalmente, para el procedimiento de generación de las matrices SDF, Vlahinic et al. proponen utilizar una matriz inicial obtenida con la función paso definida como -2 para los puntos dentro del grano, y 2 en su exterior, así como el uso de distintas fuerzas convergentes parametrizadas para ajustar iterativamente dicha matriz. Dicha parametrización representó una alta dificultad al momento de intentar ajustarla para ser usada en los granos con distintas formas y tamaños presentes en la muestra, por lo cual se realizó dos alteraciones. La primera, usar como matriz inicial la transformada de distancia euclidiana, y la segunda usar parámetros bajos (mucho menores a 1) con un número de iteraciones limitadas a 100.

## ANEXO B: Clases de errores

El conjunto de granos utilizado para los experimentos de clasificación fue inspeccionado manualmente, graficando la superficie (nivel 0) de cada uno individualmente. Se identificaron de esta manera seis distintas características indeseables que pueden ser observadas en sus superficies. Para el caso de *Forma arbitraria* se considera si la superficie contiene varias irregularidades impropias de la naturaleza de un grano. En el caso de *Dos granos unidos* se considera si el origen de la superficie antinatural se puede asociar a una falla en el paso de separación de dos granos que estaban en cercanía o contacto. Se consideran *Dos granos separados* si se evidencian dos superficies separadas en una misma matriz SDF. Las clases *Puntas pronunciadas*, *Valles hondos* y *Caras planas* se asocian a propiedades fácilmente observables en la superficie, como sus nombres indican. Un ejemplo de cada clase se presenta a continuación.

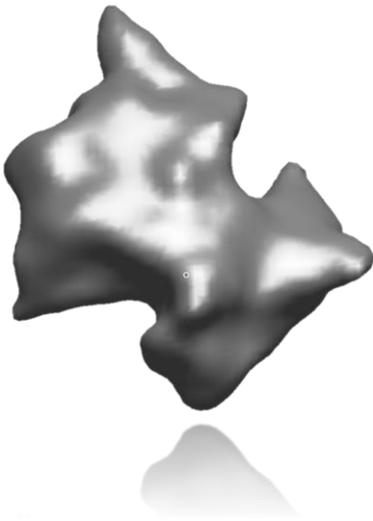
*Forma arbitraria*



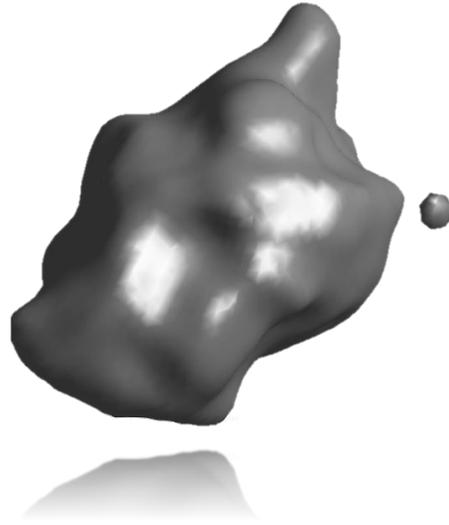
*Puntas pronunciadas*



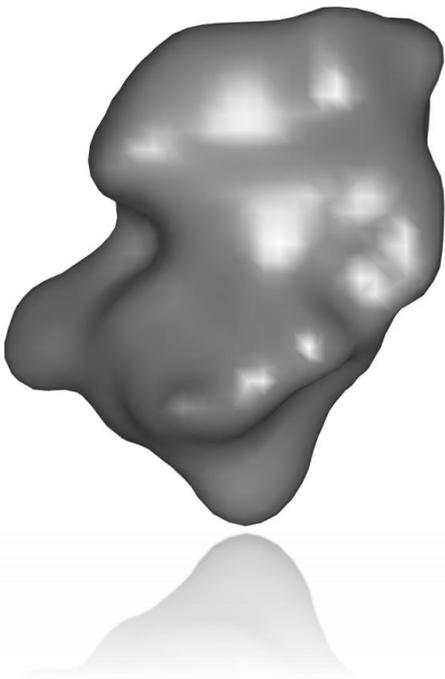
Dos granos unidos



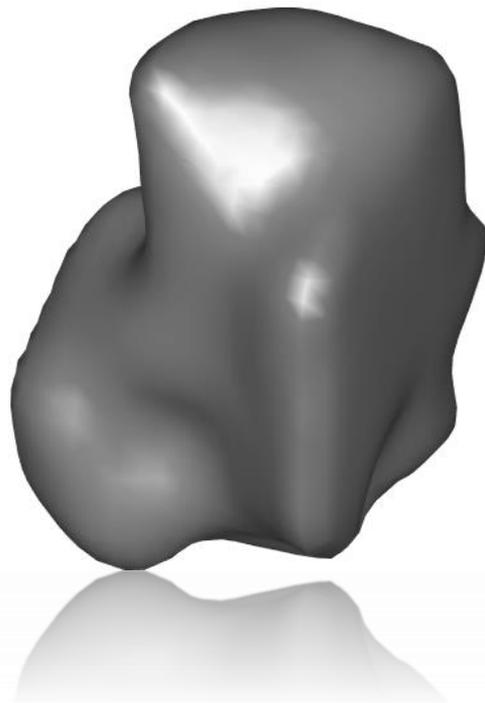
Dos granos separados



Valles hondos

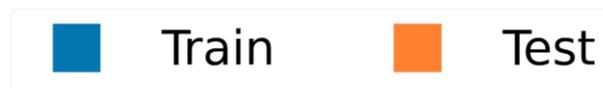


Caras planas

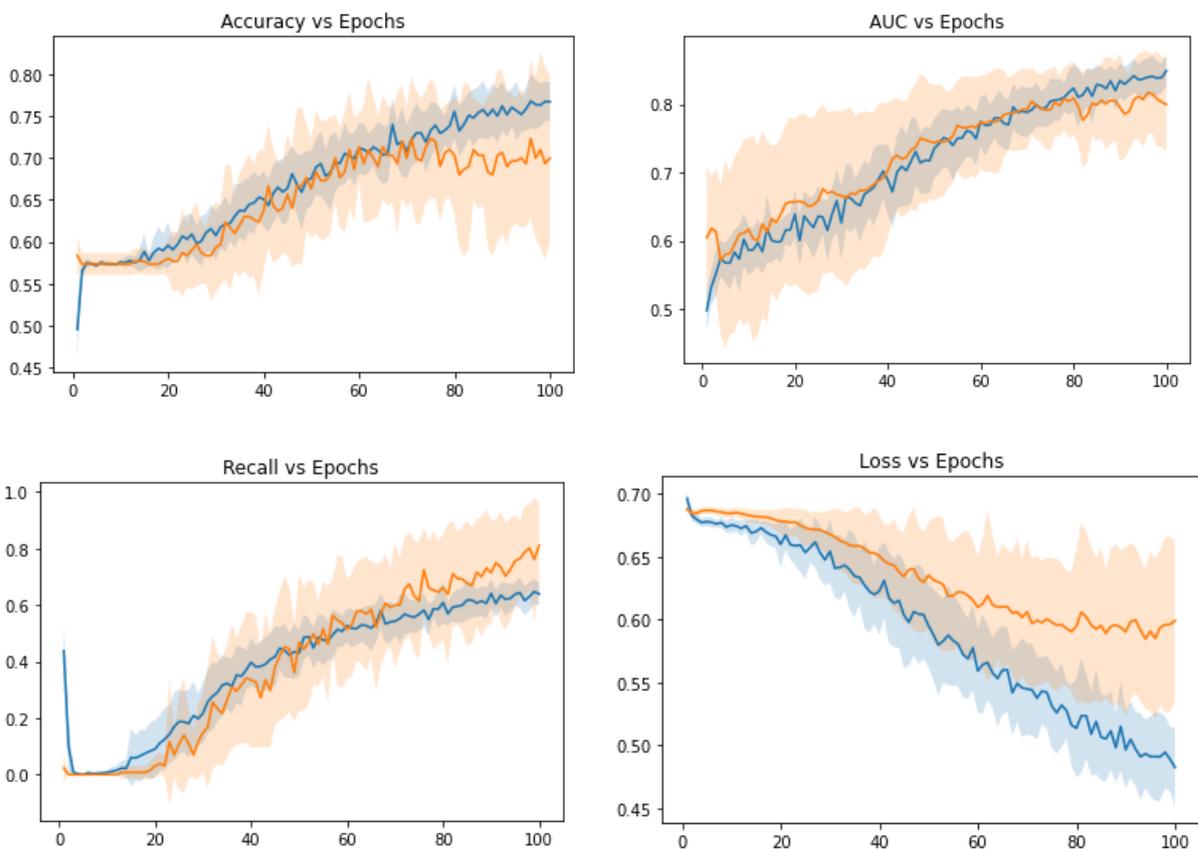


## ANEXO C: Gráficas de entrenamiento y validación

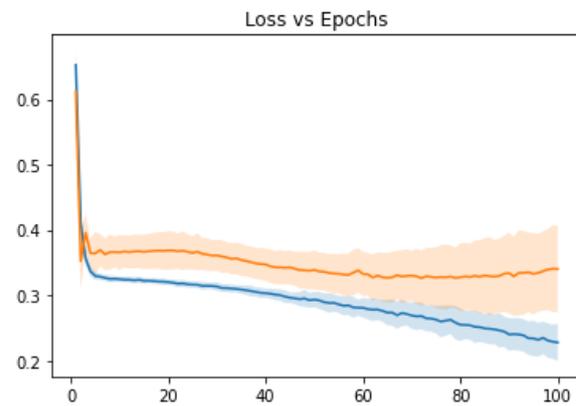
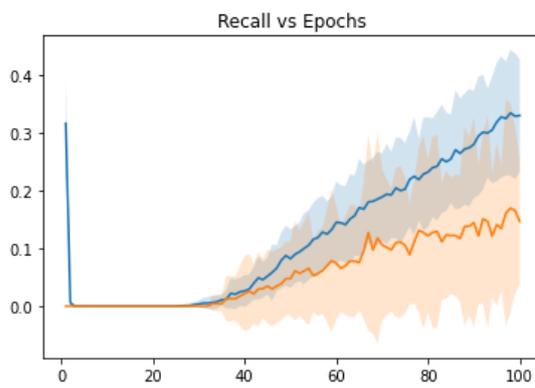
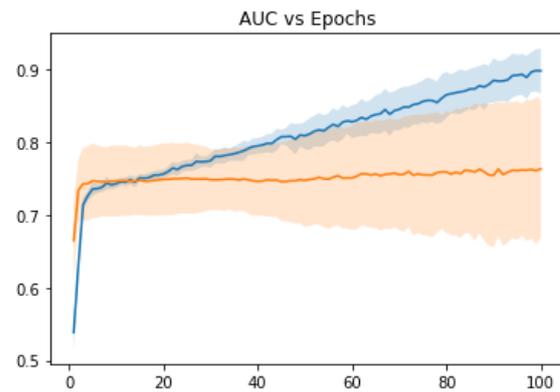
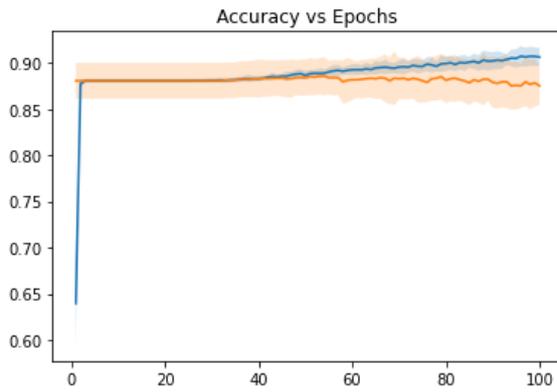
Aquí se presentan los gráficos de entrenamiento y validación tanto para las métricas analizadas en el documento como la función objetivo *binary crossentropy* conocida como *Loss*. Todos los gráficos presentados siguen el mismo código de color, con sus respectivas desviaciones estándar en sombreado.



- Sección 2.2
  - Binario

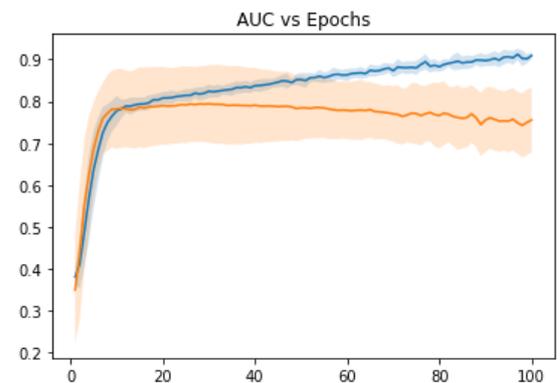
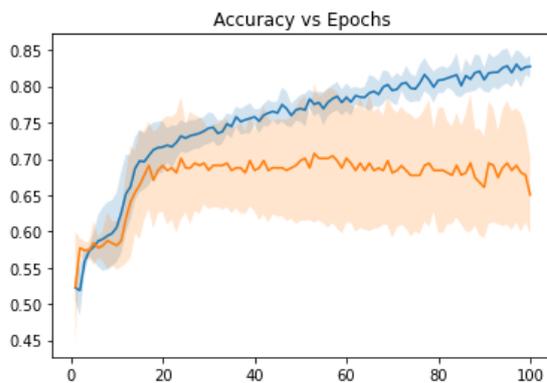


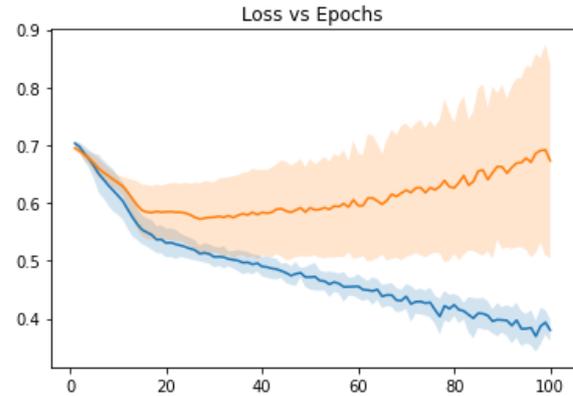
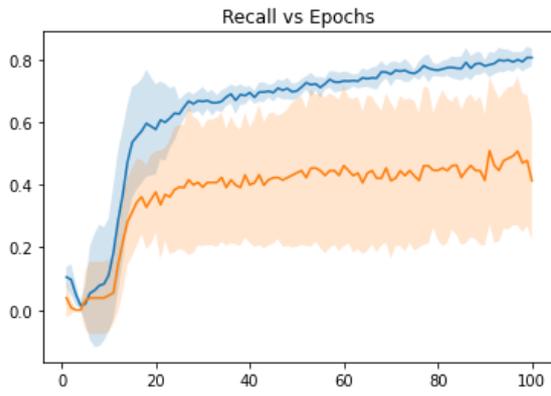
- Etiqueta múltiple



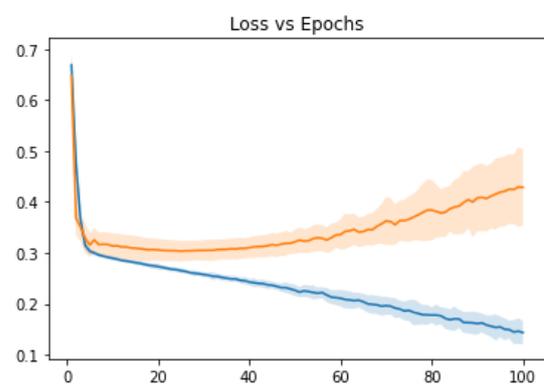
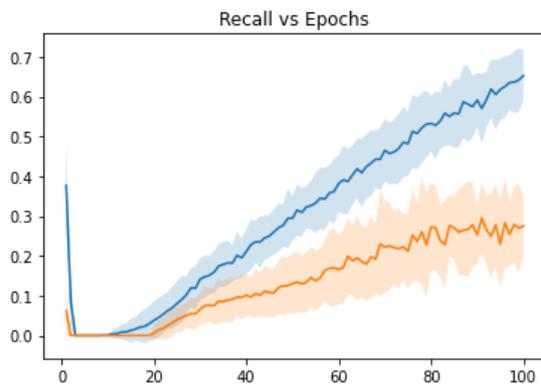
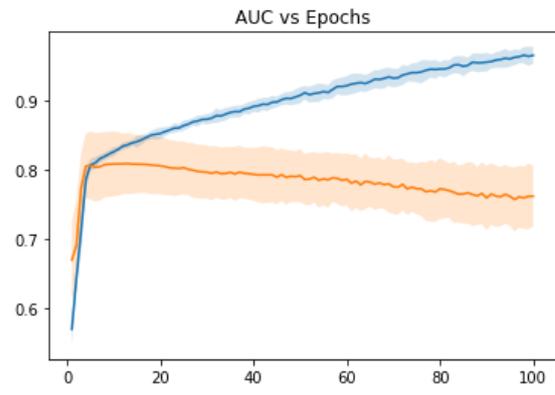
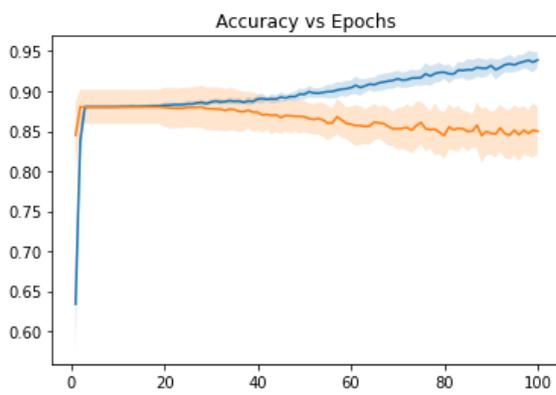
- Sección 2.3

- Binario



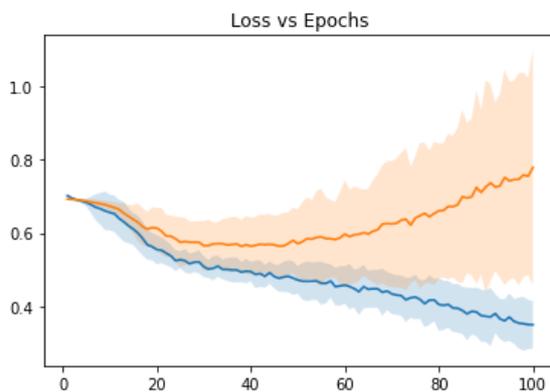
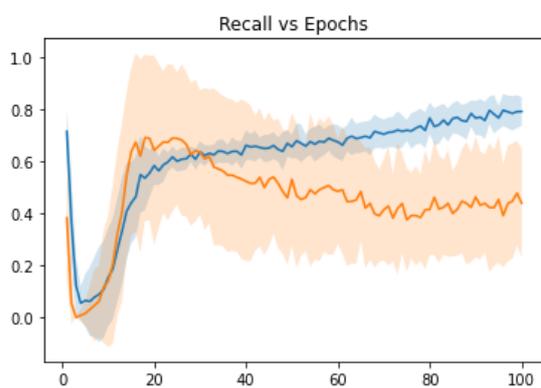
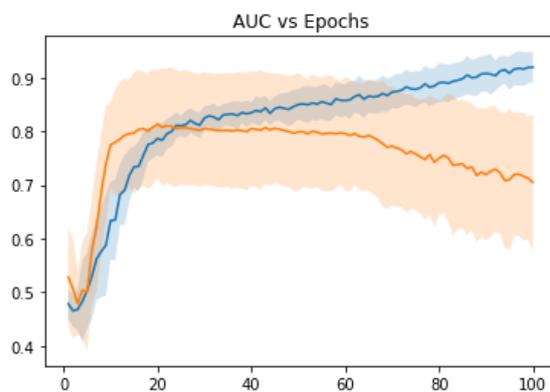
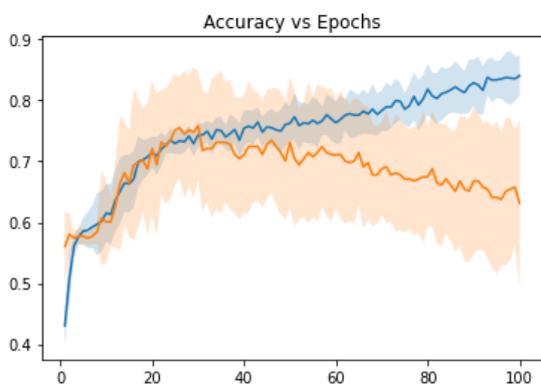


○ Etiqueta múltiple

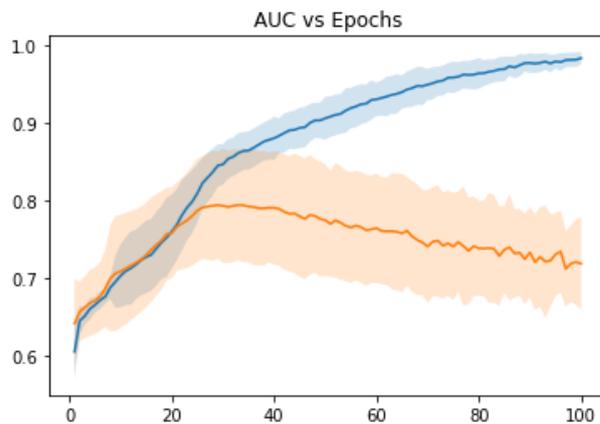
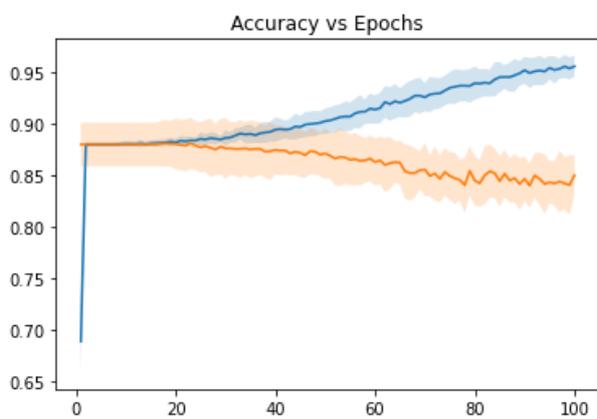


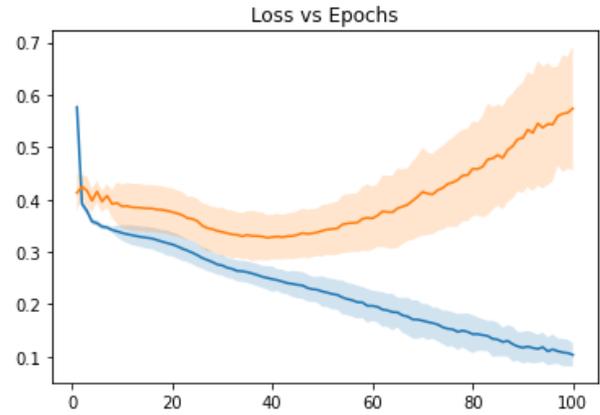
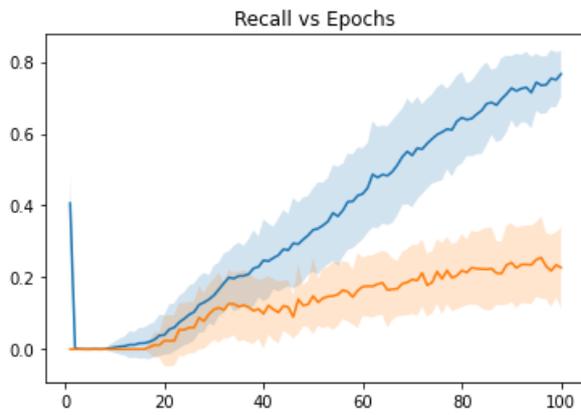
- Sección 2.4

- Binario



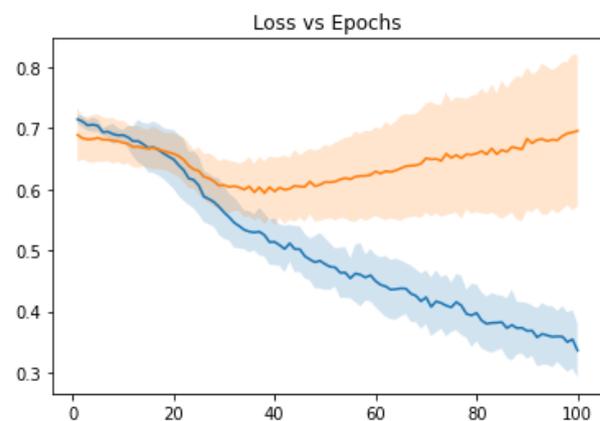
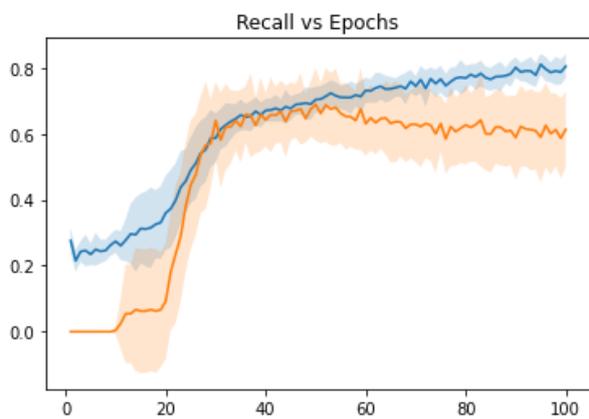
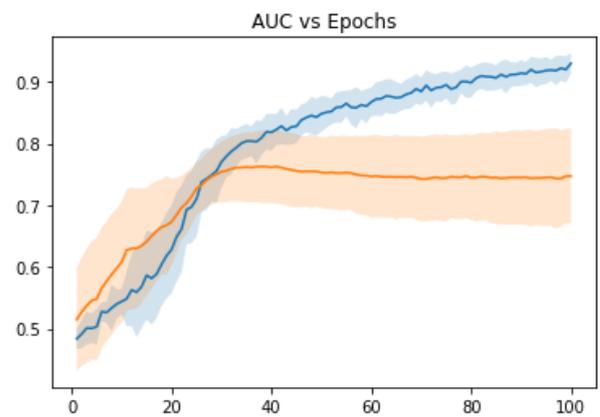
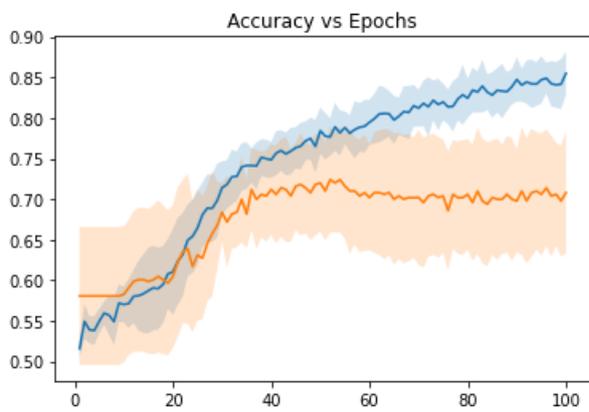
- Etiqueta múltiple



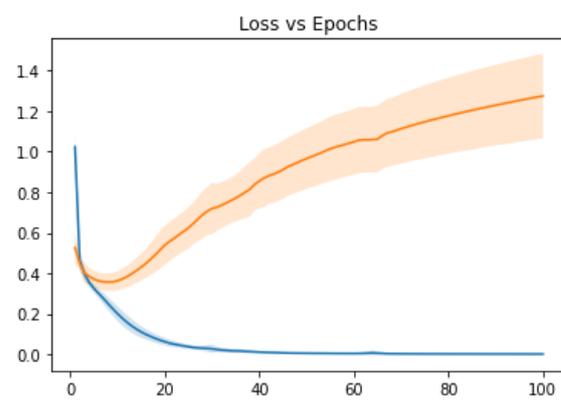
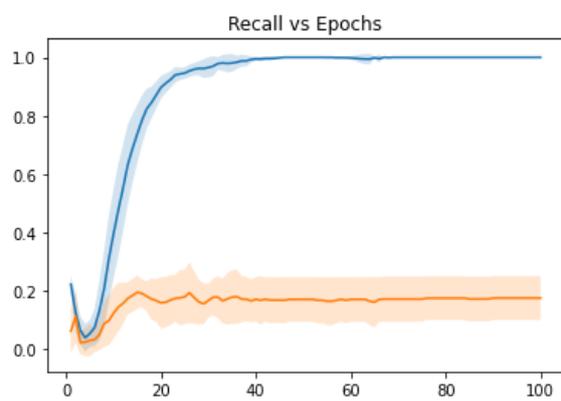
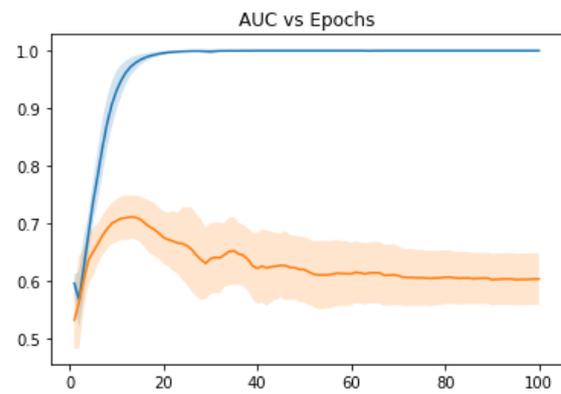
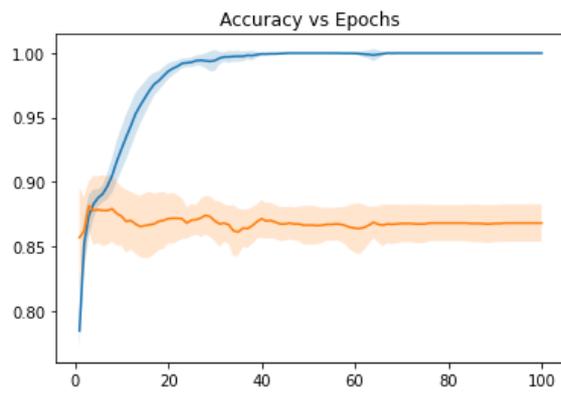


- Sección 3.1

- Binario

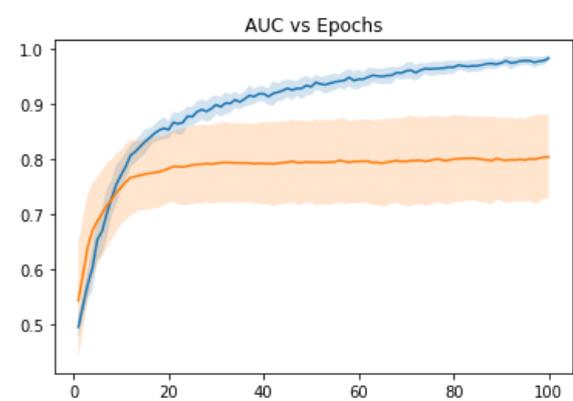
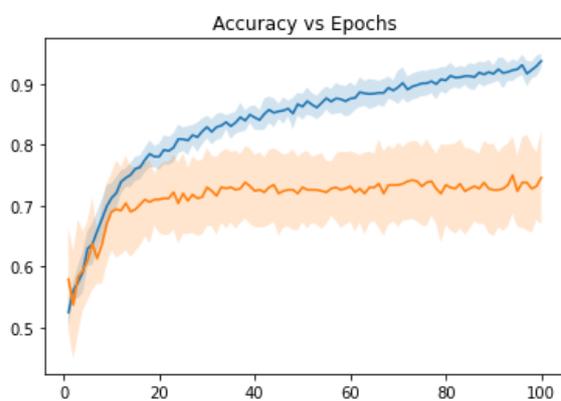


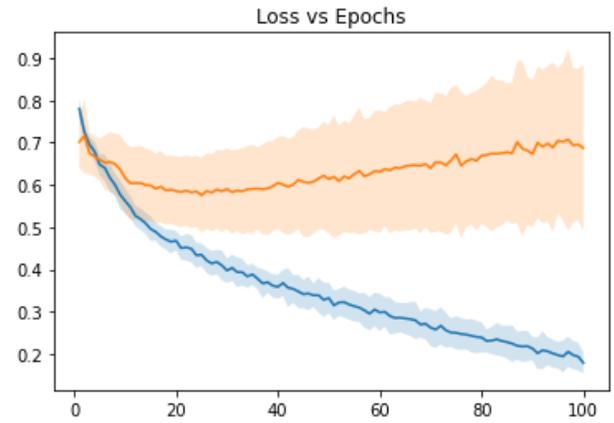
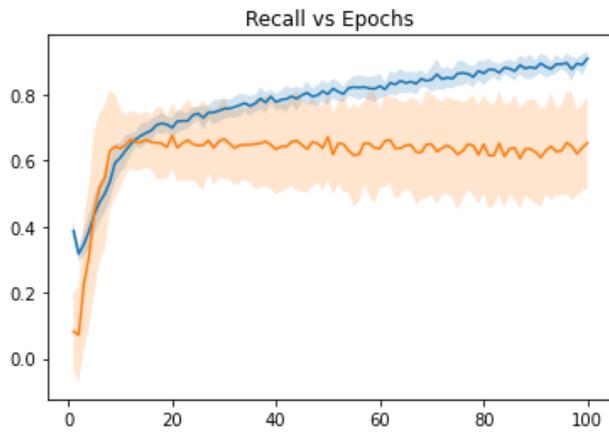
- Etiqueta múltiple



- Sección 3.2

- Binario





○ Etiqueta múltiple

