

RED CONVOLUCIONAL DE CLASIFICACIÓN DE RESIDUOS

FRANCISCO PERIS CAROT
BOOTCAMP FEB23 THEBRIDGE



INTRODUCCIÓN

Durante años la gestión de residuos ha sido un problema en nuestra sociedad y aún a día de hoy lo está siendo en aquellos países o zonas en las cuales no hay una concienciación de este problema.

Cambios normativos han provocado que este sector se industrialice pasando de un modelo de selección manual a un modelo mecánico y automatizado. Pero realmente, ¿estamos recuperando todo lo posible?

El objetivo de este trabajo es poder ayudar para que de forma no supervisada se puedan clasificar residuos sin necesidad de interacción humana.

1.- METODOLOGÍA

Hemos aplicado una red convolucional sobre el dataset obtenido de [KAGGLE](https://www.kaggle.com/code/beyzanks/waste-classification-with-cnn/notebook) (<https://www.kaggle.com/code/beyzanks/waste-classification-with-cnn/notebook>).

Este data set contiene imágenes ya clasificadas, transformadas y listas para ser utilizadas.

O_12737.jpg O_12738.jpg O_12739.jpg O_12740.jpg O_12741.jpg

Se ha utilizado un modelo secuencial con la siguiente distribución:

Model: "sequential"		
Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 298, 298, 32)	896
max_pooling2d (MaxPooling2D)	(None, 149, 149, 32)	0
conv2d_1 (Conv2D)	(None, 147, 147, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 73, 73, 64)	0
flatten (Flatten)	(None, 341856)	0
dense (Dense)	(None, 128)	43655296
dense_1 (Dense)	(None, 10)	1290
Total params: 43,675,978		
Trainable params: 43,675,978		
Non-trainable params: 0		

```
model.compile(optimizer='adam',loss='sparse_categorical_crossentropy',metrics=['accuracy'])
```

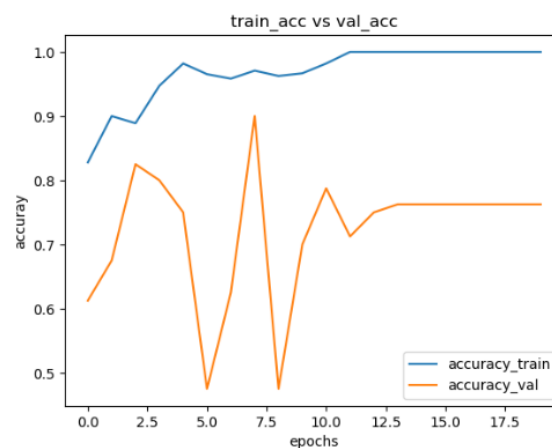
El dataset utilizado tenía más de 22.000 imágenes, al intentar normalizarlas mi ordenador tuvo problemas y decidí reducir la muestra a 1.000 imágenes (800 train + 200 test).

2.- RESULTADOS

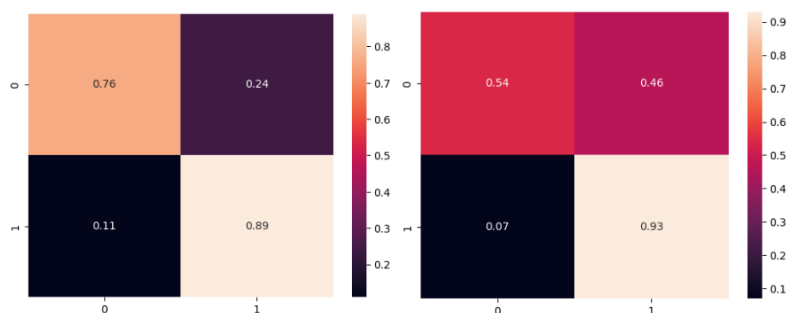
Durante los distintos entrenamientos se modificaron los parámetros de lotes y el número de épocas. Pese a eso el modelo no superaba una precisión del 85%, lo cual considero que es insuficiente. Ya que disponía de una muestra de imágenes mayor decidí tomar otra muestra de 1.000 imágenes y sustituir las que había utilizado previamente.

```
Epoch 1/20  
188/188 [=====] - 87s 471ms/step - loss: 1.0366 - accuracy: 0.8278 - val_loss: 1.2293 - val_accuracy: 0.6125  
Epoch 2/20  
188/188 [=====] - 86s 469ms/step - loss: 0.2854 - accuracy: 0.9000 - val_loss: 1.7974 - val_accuracy: 0.6750  
Epoch 3/20  
188/188 [=====] - 86s 468ms/step - loss: 0.4494 - accuracy: 0.8889 - val_loss: 0.6365 - val_accuracy: 0.8250  
Epoch 4/20  
188/188 [=====] - 84s 468ms/step - loss: 0.1338 - accuracy: 0.9472 - val_loss: 0.5146 - val_accuracy: 0.8000  
Epoch 5/20  
188/188 [=====] - 85s 470ms/step - loss: 0.0581 - accuracy: 0.9819 - val_loss: 1.2434 - val_accuracy: 0.7500  
Epoch 6/20  
188/188 [=====] - 84s 469ms/step - loss: 0.1071 - accuracy: 0.9653 - val_loss: 2.8380 - val_accuracy: 0.4750  
Epoch 7/20  
188/188 [=====] - 84s 469ms/step - loss: 0.1288 - accuracy: 0.9583 - val_loss: 1.9643 - val_accuracy: 0.6250  
Epoch 8/20  
188/188 [=====] - 84s 467ms/step - loss: 0.1152 - accuracy: 0.9708 - val_loss: 0.3053 - val_accuracy: 0.9000  
Epoch 9/20  
188/188 [=====] - 85s 474ms/step - loss: 0.1565 - accuracy: 0.9625 - val_loss: 2.6925 - val_accuracy: 0.4750  
Epoch 10/20  
188/188 [=====] - 85s 475ms/step - loss: 0.1264 - accuracy: 0.9667 - val_loss: 1.5186 - val_accuracy: 0.7000  
Epoch 11/20  
188/188 [=====] - 84s 465ms/step - loss: 0.0471 - accuracy: 0.9819 - val_loss: 1.5172 - val_accuracy: 0.7875  
Epoch 12/20  
188/188 [=====] - 84s 467ms/step - loss: 0.0031 - accuracy: 1.0000 - val_loss: 2.3487 - val_accuracy: 0.7125  
Epoch 13/20  
....  
Epoch 19/20  
188/188 [=====] - 84s 468ms/step - loss: 1.4095e-04 - accuracy: 1.0000 - val_loss: 2.5383 - val_accuracy: 0.7625  
Epoch 20/20  
188/188 [=====] - 84s 468ms/step - loss: 1.2846e-04 - accuracy: 1.0000 - val_loss: 2.5487 - val_accuracy: 0.7625  
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings.
```

[Primer entrenamiento]



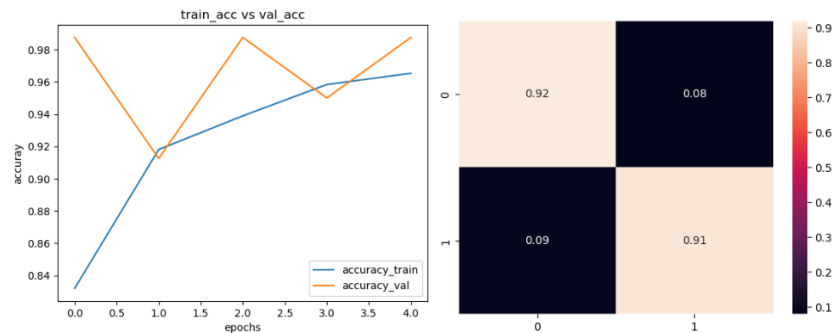
[Línea de validación]



Matrices de confusión de varios entrenamientos, presentan problemas para clasificar 0 = Orgánico.

Durante los primeros entrenamientos el modelo tenía problemas sobre todo para catalogar de forma adecuada los residuos orgánicos. Pienso que fue debido a un sesgo en las imágenes tomadas de la muestra. Con el posterior cambio de la muestra los resultados mejoraron.

Con la segunda muestra de imágenes, al entrenar el modelo este obtuvo un mejor performance. Logró superar el 91%, lo cual ya considero que es un desempeño adecuado.



Muestra de clasificación final, los valores en rojo están mal clasificados.



[O = Orgánico; R = Reciclable]

3.- CONCLUSIONES

Si dispusiéramos de una muestra con una mayor cantidad de residuos tipo podríamos utilizarla para poder segregar el flujo de material que pasa una instalación de tratamiento de residuos.

Considero que la segregación se adapta a mis necesidades y que es un modelo óptimo y escalable a otros problemas de índole similar.