

Comparación del rendimiento de diferentes redes neuronales pre-entrenadas

Procedimiento

A continuación se detalla el procedimiento realizado para la comparación:

Sobre los espectrogramas

- En base a los audios grabados en la UNLAM, se obtuvieron distintos tipos de “datasets” de espectrogramas. En cada dataset, una característica de los espectrogramas varía. Por ejemplo, algunos son en escala de grises, otros a color, otros con ejes de tiempo y frecuencia, etc.
- Se generaron espectrogramas tanto en escala de mel, como en escala logarítmica.
- No se aplicó ningún tipo de pre-procesamiento a los espectrogramas en la generación de los mismos.

Sobre las redes neuronales

- Se utilizó la técnica de validación cruzada de 10 folds (tenth fold). Dicha técnica nos ayuda a obtener una mejor aproximación del rendimiento de la red neuronal cuando el set de datos es muy reducido, como en nuestro caso.
- Se usaron algunas de las redes pre-entrenadas fácilmente disponibles a través de Keras. En este enlace se puede verificar cuáles redes Keras tiene disponibles: <https://keras.io/api/applications/>
- Se realizaron 4 pruebas con diferentes parámetros en cada dataset.

Tipos de pruebas

Learning rate variable

Estas pruebas presentan un learning rate variable, pero mantienen constantes las épocas y el batch_size.

Parámetros	Prueba 1	Prueba 2	Prueba 3
batch_size	16	16	16
learning_rate	0.0002	0.0005	0.0009
initial_epochs	100	100	100

Parámetros	Prueba 4	Prueba 5	Prueba 6
batch_size	32	32	32
learning_rate	0.0002	0.0005	0.0009
initial_epochs	100	100	100

Número de épocas variable

El número de épocas varía, pero el resto de parámetros no.

Parámetros	Prueba 1	Prueba 2
batch_size	16	16
learning_rate	0.0005	0.0005
initial_epochs	50	250

Parámetros	Prueba 3	Prueba 4
batch_size	32	32
learning_rate	0.0005	0.0005
initial_epochs	50	250

Consideraciones

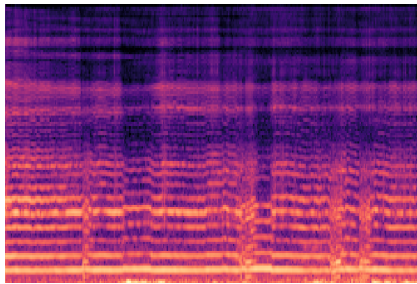
Algunos parámetros son constantes o dependen de la red neuronal utilizada. Por ejemplo, IMG_SIZE es un parámetro que cambia dependiendo de la red neuronal utilizada.

Redes utilizadas

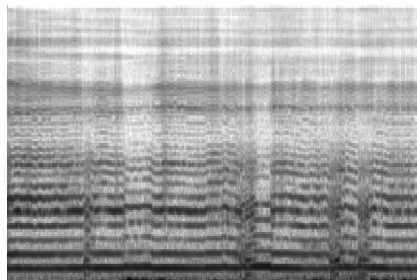
- Xception
- VGG19
- ResNet18
- ResNet50
- InceptionV3
- MobileNetV2
- DenseNet121

Datasets utilizados

Default Color Map



Grayscale Color map



Viridis Color Map

