

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

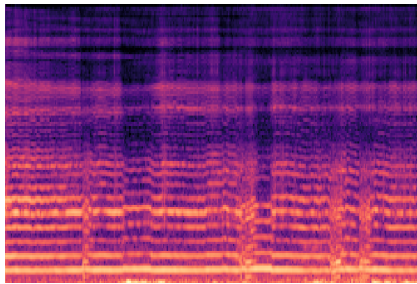
Algunos parámetros son constantes o dependen de la red neuronal utilizada. Por ejemplo, IMG_SIZE es un parámetro que cambia dependiendo de la red neuronal utilizada.

Redes utilizadas

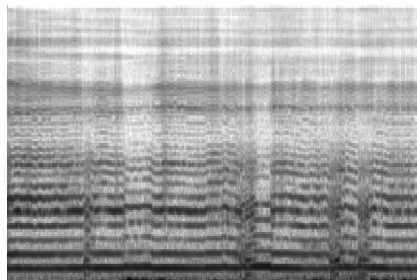
- Xception
- VGG19
- ResNet18
- ResNet50
- InceptionV3
- MobileNetV2
- DenseNet121

Datasets utilizados

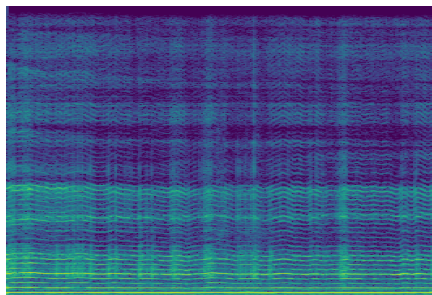
Default Color Map



Grayscale Color map



Viridis Color Map



Learning rate variable
Default Color Map

Accuracy / Loss											
Red	Prueba 1	Red	Prueba 2	Red	Prueba 3	Red	Prueba 4	Red	Prueba 5	Red	Prueba 6
MobileNetV2	94.23 (+ 4.63)	MobileNetV2	95.36 (+ 4.13)	MobileNetV2	95.73 (+ 4.99)	MobileNetV2	93.83 (+ 5.48)	ResNet50	96.15 (+ 3.84)	MobileNetV2	96.15 (+ 4.55)
	0.16		0.11		0.11		0.17		0.12		0.11
ResNet50	91.52 (+ 7.83)	ResNet50	95.36 (+ 6.83)	ResNet50	94.21 (+ 5.76)	ResNet50	93.06 (+ 7.83)	MobileNetV2	95.36 (+ 4.80)	ResNet50	95.38 (+ 6.83)
	0.15		0.11		0.10		0.18		0.13		0.12
VGG19	90.35 (+ 5.47)	DenseNet121	91.13 (+ 5.70)	DenseNet121	91.92 (+ 6.97)	DenseNet121	91.13 (+ 7.69)	DenseNet121	93.04 (+ 6.84)	Xception	91.90 (+ 6.29)
	0.25		0.23		0.21		0.28		0.23		0.19
DenseNet121	90.29 (+ 5.00)	InceptionV3	91.13 (+ 9.42)	Xception	91.15 (+ 8.77)	InceptionV3	89.95 (+ 8.29)	VGG19	90.73 (+ 5.46)	DenseNet121	91.52 (+ 6.37)
	0.27		0.21		0.19		0.23		0.22		0.24
Xception	88.83 (+ 7.76)	Xception	90.73 (+ 8.78)	InceptionV3	89.90 (+ 5.24)	Xception	89.93 (+ 4.25)	Xception	89.58 (+ 9.40)	InceptionV3	90.75 (+ 11.55)
	0.25		0.19		0.20		0.27		0.22		0.23
InceptionV3	88.43 (+ 11.78)	VGG19	88.41 (+ 7.49)	VGG19	87.98 (+ 7.19)	VGG19	88.79 (+ 7.55)	InceptionV3	89.58 (+ 8.77)	VGG19	89.18 (+ 3.71)
	0.24		0.23		0.24		0.26		0.24		0.22
ResNet18	84.15 (+ 6.23)	ResNet18	86.49 (+ 10.18)	ResNet18	83.39 (+ 8.56)	ResNet18	81.86 (+ 9.56)	ResNet18	82.98 (+ 6.19)	ResNet18	84.93 (+ 7.92)
	0.35		0.33		0.33		0.37		0.35		0.33

Learning rate variable

Grayscale Color Map

Accuracy / Loss											
Red	Prueba 1	Red	Prueba 2	Red	Prueba 3	Red	Prueba 4	Red	Prueba 5	Red	Prueba 6
ResNet50	95.36 (+ 4.48)	ResNet50	96.92 (+ 3.76)	Xception	96.92 (+ 4.14)	MobileNetV2	93.83 (+ 5.48)	ResNet50	95.36 (+ 4.80)	ResNet50	96.53 (+ 5.56)
	0.11		0.087		0.10		0.17		0.13		0.096
MobileNetV2	94.99 (+ 5.17)	MobileNetV2	96.153 (+ 3.84)	ResNet50	96.92 (+ 4.80)	ResNet50	92.29 (+ 7.08)	MobileNetV2	95.36 (+ 4.80)	MobileNetV2	96.15 (+ 4.55)
	0.17		0.12		0.086		0.17		0.13		0.11
Xception	94.23 (+ 5.76)	Xception	93.83 (+ 7.72)	MobileNetV2	95.36 (+ 4.48)	Xception	89.61 (+ 9.88)	Xception	93.84 (+ 5.21)	Xception	96.15 (+ 5.70)
	0.19		0.14		0.10		0.24		0.15		0.13
InceptionV3	90.33 (+ 6.47)	DenseNet121	91.83 (+ 5.64)	DenseNet121	91.52 (+ 8.20)	DenseNet121	89.59 (+ 7.49)	VGG19	90.73 (+ 5.46)	DenseNet121	89.55 (+ 7.88)
	0.21		0.26		0.26		0.30		0.22		0.26
DenseNet121	89.56 (+ 5.72)	InceptionV3	89.96 (+ 7.32)	VGG19	89.58 (+ 7.29)	VGG19	88.79 (+ 7.55)	InceptionV3	89.95 (+ 8.44)	VGG19	89.18 (+ 3.71)
	0.28		0.23		0.25		0.26		0.21		0.22
VGG19	88.03 (+ 6.97)	VGG19	87.99 (+ 5.53)	InceptionV3	89.56 (+ 7.28)	InceptionV3	87.27 (+ 12.85)	DenseNet121	89.56 (+ 6.18)	InceptionV3	86.09 (+ 11.65)
	0.28		0.28		0.25		0.26		0.26		0.26
ResNet18	83.35 (+ 8.77)	ResNet18	86.86 (+ 7.31)	ResNet18	86.07 (+ 6.66)	ResNet18	81.86 (+ 9.56)	ResNet18	82.98 (+ 6.19)	ResNet18	84.93 (+ 7.92)
	0.33		0.30		0.30		0.37		0.35		0.33

Learning rate variable
Viridis Color Map

Accuracy / Loss											
Red	Prueba 1	Red	Prueba 2	Red	Prueba 3	Red	Prueba 4	Red	Prueba 5	Red	Prueba 6
ResNet50	95.36 (+ 5.38)	ResNet50	95.76 (+ 5.82)	ResNet50	98.06 (+ 3.11)	ResNet50	92.30 (+ 8.06)	ResNet50	95.38 (+ 7.05)	ResNet50	96.15 (+ 5.95)
	0.15		0.11		0.067		0.18		0.11		0.11
MobileNetV2	93.03 (+ 5.38)	MobileNetV2	91.90 (+ 8.66)	MobileNetV2	96.15 (+ 2.97)	InceptionV3	91.50 (+ 6.81)	DenseNet121	90.76 (+ 6.92)	MobileNetV2	94.58 (+ 5.50)
	0.18		0.16		0.13		0.23		0.26		0.14
InceptionV3	91.53 (+ 7.45)	Xception	90.75 (+ 8.79)	Xception	90.38 (+ 9.91)	MobileNetV2	90.75 (+ 6.91)	InceptionV3	90.75 (+ 7.52)	Xception	92.30 (+ 9.10)
	0.23		0.21		0.18		0.23		0.24		0.19
Xception	88.39 (+ 8.27)	DenseNet121	90.70 (+ 3.93)	InceptionV3	90.30 (+ 9.01)	VGG19	86.86 (+ 5.17)	MobileNetV2	90.35 (+ 13.66)	InceptionV3	91.49 (+ 6.41)
	0.28		0.25		0.23		0.28		0.19		0.23
DenseNet121	88.03 (+ 5.24)	VGG19	89.58 (+ 5.95)	DenseNet121	90.29 (+ 4.69)	DenseNet121	86.86 (+ 6.43)	VGG19	90.32 (+ 4.91)	VGG19	90.35 (+ 6.69)
	0.28		0.23		0.23		0.29		0.24		0.23
VGG19	87.90 (+ 7.24)	InceptionV3	88.79 (+ 8.65)	VGG19	88.79 (+ 6.03)	Xception	85.72 (+ 8.93)	Xception	89.21 (+ 9.53)	DenseNet121	89.92 (+ 6.00)
	0.27		0.23		0.24		0.32		0.24		0.24
ResNet18	83.75 (+ 7.06)	ResNet18	84.49 (+ 8.13)	ResNet18	85.67 (+ 5.95)	ResNet18	81.44 (+ 9.79)	ResNet18	82.21 (+ 7.67)	ResNet18	82.92 (+ 6.31)
	0.36		0.34		0.33		0.38		0 .35		0.34

Epocas variable
Default Color Map

Accuracy / Loss							
Red	Prueba 1	Red	Prueba 2	Red	Prueba 3	Red	Prueba 4
MobileNetV2	95.35 (+ 5.40)	ResNet50	97.30 (+ 3.86)	MobileNetV2	95.36 (+ 3.34)	MobileNetV2	97.30 (+ 4.88)
	0.13		0.09		0.15		0.10
ResNet50	94.23 (+ 7.54)	MobileNetV2	96.89 (+ 3.39)	ResNet50	94.19 (+ 3.12)	ResNet50	95.75 (+ 4.99)
	0.15		0.077		0.15		0.10
Xception	92.29 (+ 6.65)	Xception	91.10 (+ 8.06)	DenseNet121	90.32 (+ 5.77)	DenseNet121	91.89 (+ 4.33)
	0.24		0.17		0.27		0.21
DenseNet121	91.83 (+ 8.24)	DenseNet121	90.75 (+ 6.91)	InceptionV3	89.55 (+ 8.07)	Xception	91.52 (+ 7.24)
	0.24		0.22		0.22		0.17
VGG19	89.16 (+ 9.53)	InceptionV3	89.18 (+ 10.00)	Xception	88.81 (+ 8.47)	InceptionV3	89.98 (+ 8.27)
	0.24		0.24		0.27		0.23
InceptionV3	89.20 (+ 12.37)	VGG19	88.76 (+ 6.52)	VGG19	87.98 (+ 6.38)	VGG19	88.38 (+ 5.13)
	0.25		0.23		0.25		0.21
ResNet18	84.52 (+ 8.40)	ResNet18	83.75 (+ 6.55)	ResNet18	84.95 (+ 7.53)	ResNet18	84.52 (+ 5.09)
	0.34		0.33		0.36		0.32

Epocas variable
Grayscale Color Map

Accuracy / Loss							
Red	Prueba 1	Red	Prueba 2	Red	Prueba 3	Red	Prueba 4
ResNet50	94.23 (+ 7.54)	ResNet50	97.30 (+ 3.86)	ResNet50	94.19 (+ 3.12)	ResNet50	95.75 (+ 4.99)
	0.15		0.09		0.15		0.10
MobileNetV2	94.21 (+ 7.54)	MobileNetV2	95.36 (+ 6.61)	MobileNetV2	93.06 (+ 6.38)	Xception	95.76 (+ 8.32)
	0.16		0.12		0.19		0.12
Xception	91.92 (+ 9.33)	Xception	92.69 (+ 11.33)	Xception	91.12 (+ 9.58)	MobileNetV2	95.0 (+ 6.89)
	0.19		0.14		0.21		0.14
DenseNet121	89.52 (+ 5.49)	DenseNet121	89.92 (+ 5.53)	VGG19	88.39 (+ 7.33)	DenseNet121	91.90 (+ 5.54)
	0.27		0.26		0.29		0.27
VGG19	88.04 (+ 9.63)	VGG19	89.56 (+ 3.88)	InceptionV3	88.046 (+ 12.08)	VGG19	88.78 (+ 7.42)
	0.29		0.26		0.25		0.25
InceptionV3	87.64 (+ 12.11)	ResNet18	87.63 (+ 4.45)	DenseNet121	87.99 (+ 6.74)	InceptionV3	87.24 (+ 7.28)
	0.24		0.29		0.29		0.28
ResNet18	86.10 (+ 9.41)	InceptionV3	84.89 (+ 9.16)	ResNet18	81.87 (+ 10.99)	ResNet18	84.16 (+ 10.33)
	0.32		0.31		0.36		0.30

Epocas variable
Viridis Color Map

Accuracy / Loss							
Red	Prueba 1	Red	Prueba 2	Red	Prueba 3	Red	Prueba 4
ResNet50	95.75 (+ 7.77)	ResNet50	96.153 (+ 6.88)	ResNet50	95.76 (+ 5.82)	ResNet50	98.07 (+ 3.10)
	0.15		0.096		0.17		0.074
MobileNetV2	91.52 (+ 9.06)	Xception	94.61 (+ 6.0)	MobileNetV2	92.64 (+ 6.34)	MobileNetV2	96.12 (+ 4.57)
	0.19		0.13		0.19		0.11
InceptionV3	89.95 (+ 10.61)	MobileNetV2	93.41 (+ 5.71)	DenseNet121	88.43 (+ 7.86)	Xception	91.92 (+ 10.09)
	0.21		0.13		0.29		0.19
DenseNet121	89.52 (+ 3.53)	InceptionV3	89.21 (+ 7.63)	InceptionV3	87.24 (+ 13.42)	DenseNet121	90.75 (+ 6.23)
	0.27		0.30		0.25		0.23
VGG19	89.10 (+ 6.50)	DenseNet121	89.19 (+ 7.81)	VGG19	86.50 (+ 9.42)	VGG19	90.30 (+ 5.51)
	0.24		0.25		0.28		0.22
Xception	88.07 (+ 11.97)	VGG19	87.20 (+ 9.31)	Xception	86.84 (+ 7.70)	InceptionV3	88.83 (+ 8.49)
	0.27		0.25		0.31		0.24
ResNet18	83.69 (+ 5.74)	ResNet18	84.12 (+ 6.76)	ResNet18	81.79 (+ 2.90)	ResNet18	83.35 (+ 9.42)
	0.35		0.33		0.38		0.34