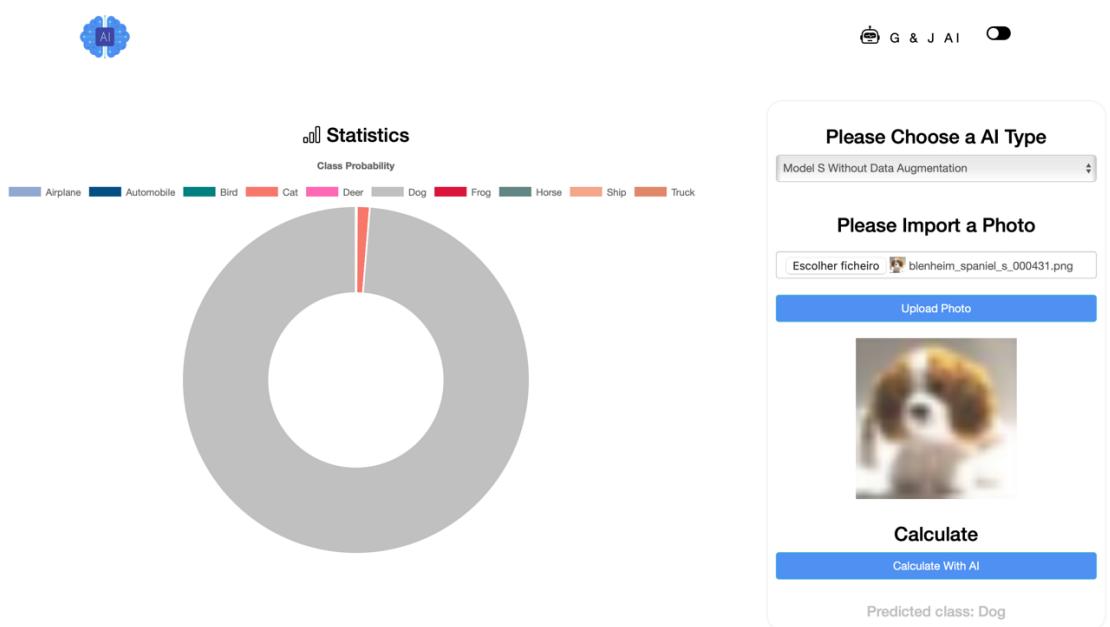


# ModelS\_without\_dataAugmentation Results

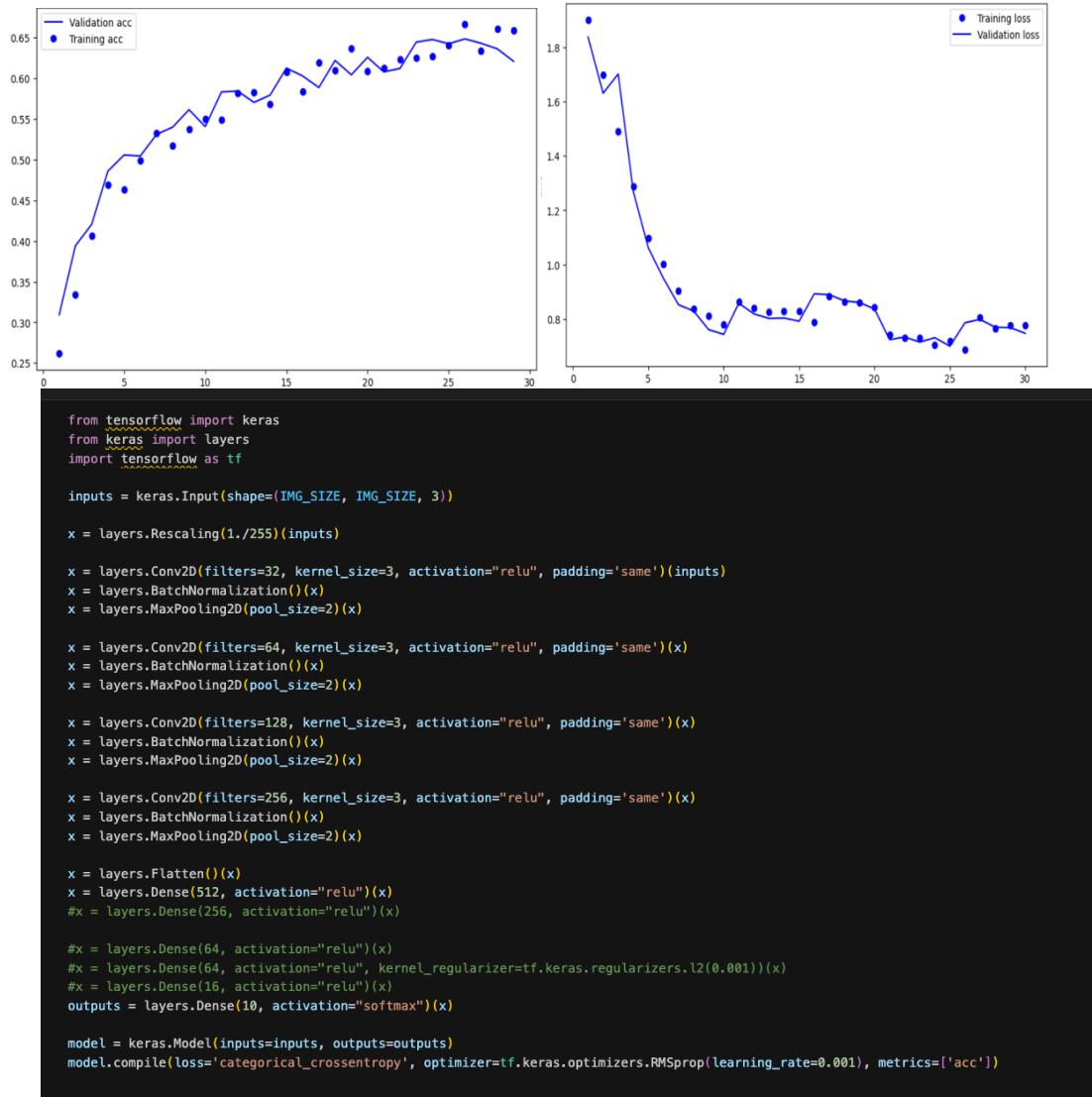


Powered By:

Gonçalo Ferreira, nº 2222051  
José Delgado, nº 2222049

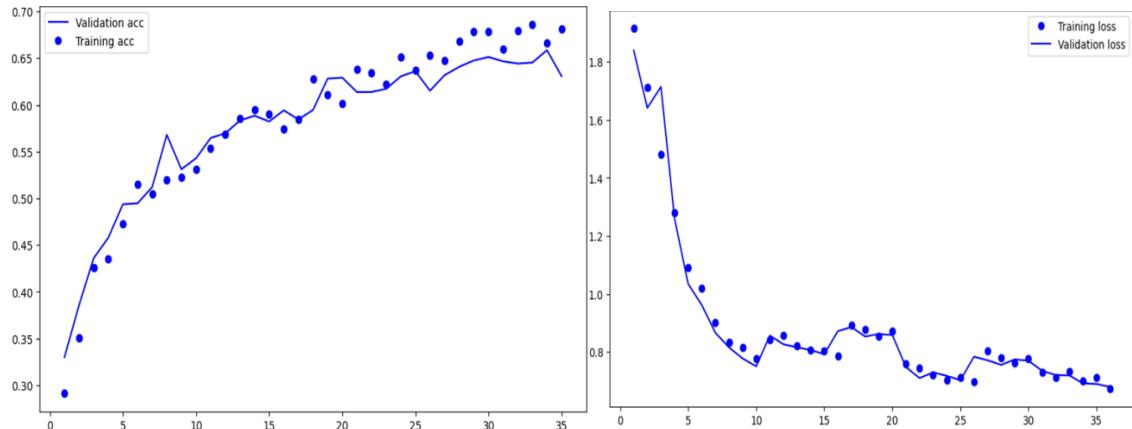
## Results

Batch\_Size 256, 30 épocas Optimizer RMSProp Sem DropOut  
Sem Data Augmentation



Com esta nossa primeira rede, conseguimos uma accuracy de sensivelmente 0.67 e uma loss de 0.6. Com um batch\_size de 256, o treino foi relativamente rápido, tendo em conta o tamanho do dataset de treino. Foi utilizado o optimizer utilizado nas aulas e não foi utilizado nem dropout nem data augmentation.

## Batch\_Size 128, 36 épocas Optimizer RMSProp Sem DropOut Sem Data Augmentation



```

from tensorflow import keras
from keras import layers
import tensorflow as tf

inputs = keras.Input(shape=(IMG_SIZE, IMG_SIZE, 3))

x = layers.Rescaling(1./255)(inputs)

x = layers.Conv2D(filters=32, kernel_size=3, activation="relu", padding='same')(inputs)
x = layers.BatchNormalization()(x)
x = layers.MaxPooling2D(pool_size=2)(x)

x = layers.Conv2D(filters=64, kernel_size=3, activation="relu", padding='same')(x)
x = layers.BatchNormalization()(x)
x = layers.MaxPooling2D(pool_size=2)(x)

x = layers.Conv2D(filters=128, kernel_size=3, activation="relu", padding='same')(x)
x = layers.BatchNormalization()(x)
x = layers.MaxPooling2D(pool_size=2)(x)

x = layers.Conv2D(filters=256, kernel_size=3, activation="relu", padding='same')(x)
x = layers.BatchNormalization()(x)
x = layers.MaxPooling2D(pool_size=2)(x)

x = layers.Flatten()(x)
x = layers.Dense(512, activation="relu")(x)
#x = layers.Dense(256, activation="relu")(x)

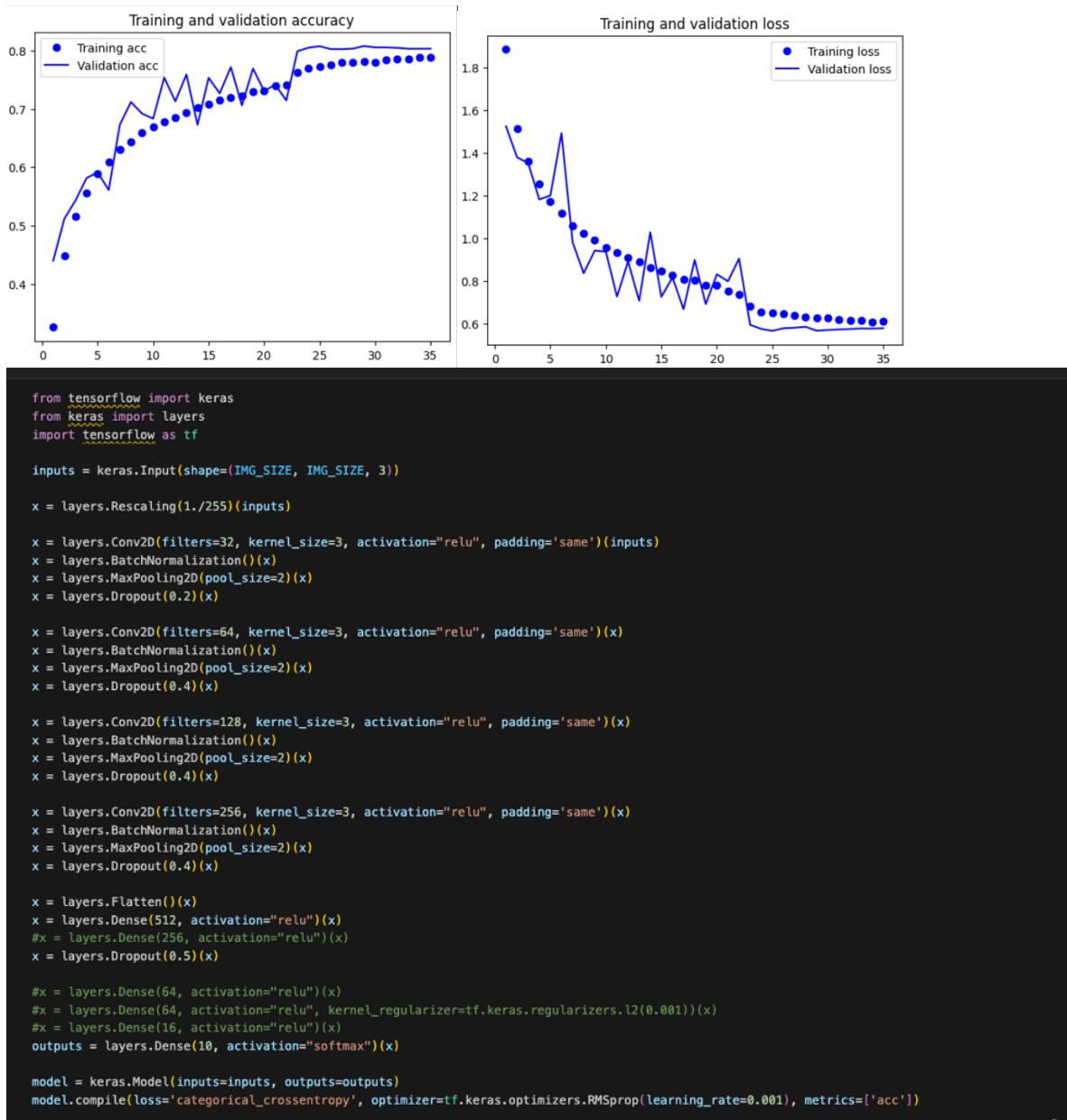
#x = layers.Dense(64, activation="relu")(x)
#x = layers.Dense(64, activation="relu", kernel_regularizer=tf.keras.regularizers.l2(0.001))(x)
#x = layers.Dense(16, activation="relu")(x)
outputs = layers.Dense(10, activation="softmax")(x)

model = keras.Model(inputs=inputs, outputs=outputs)
model.compile(loss='categorical_crossentropy', optimizer=tf.keras.optimizers.RMSprop(learning_rate=0.001), metrics=['acc'])

```

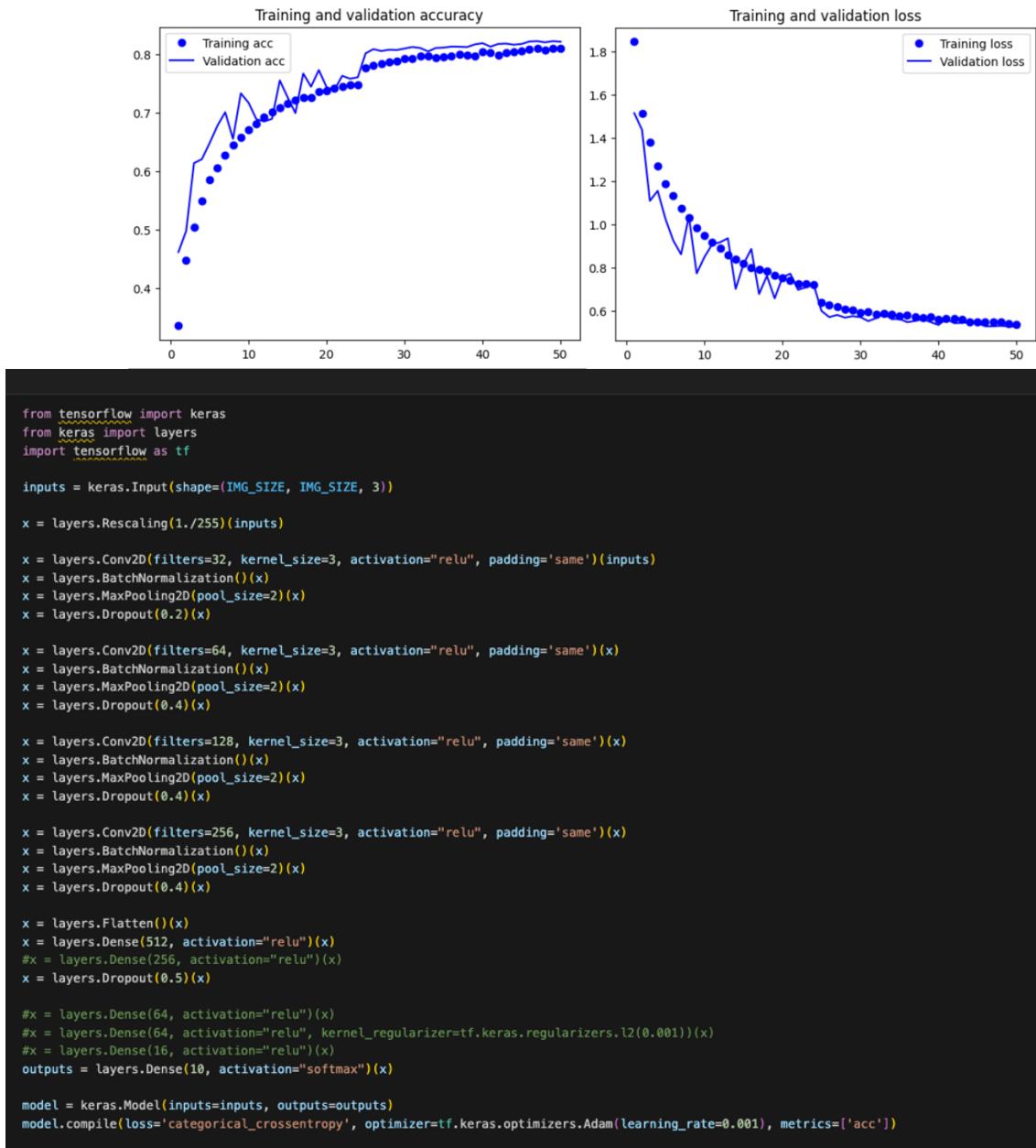
Com esta segunda versão da rede, conseguimos uma accuracy de sensivelmente 0.69 e uma loss de 0.56. Com um batch\_size de 128, o treino foi relativamente rápido, tendo em conta o tamanho do dataset de treino, no entanto mais lento e demorado, comparando com a versão anterior. Foi utilizado o optimizer utilizado nas aulas e não foi utilizado nem dropout nem data augmentation.

## Batch\_Size 64, 35 épocas Optimizer RMSProp Com DropOut Sem Data Augmentation



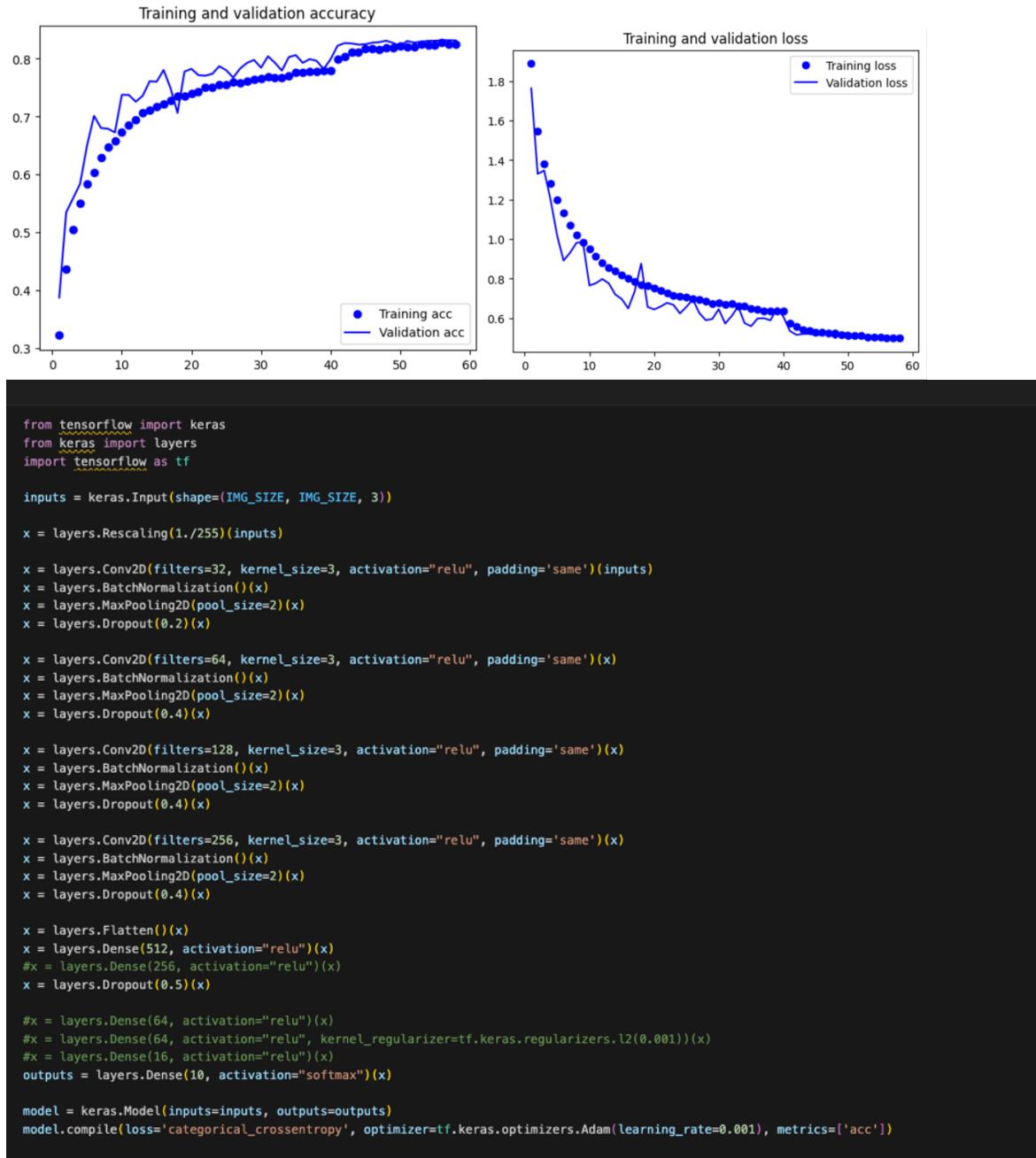
Com esta terceira versão da rede, conseguimos uma accuracy de sensivelmente 0.8 e uma loss de 0.58. Com um batch\_size de 64, o treino foi ainda mais demorado, pois os “lotes de imagens” ficavam cada vez maiores. Foi utilizado o optimizer utilizado nas aulas e não foi utilizado data augmentation. Porém introduzimos nesta rede o dropout, que desliga aleatoriamente as ligações entre neurónios entre as respetivas camadas. Efetivamente, conseguimos resultados muito melhores utilizando o dropout e um batch\_size menor.

## Batch\_Size 32, 50 épocas Optimizer Adam Com DropOut Sem Data Augmentation



Com esta quarta versão da rede, conseguimos uma accuracy de sensivelmente 0.82 e uma loss de 0.57. Com um batch\_size de 32, o treino foi ainda mais demorado que na versão anterior pelo mesmo motivo. Foi utilizado o optimizer Adam, que ajusta a taxa de aprendizagem (learning-rate) de cada parâmetro individualmente, tornando-o muito mais eficaz em situações onde a escala dos gradientes varia. Não foi utilizado data augmentation. Porém introduzimos também nesta rede o dropout, que desliga aleatoriamente as ligações entre neurónios entre as respetivas camadas. Efetivamente, conseguimos resultados muito melhores utilizando o dropout, um batch\_size menor e o optimizer Adam.

## Batch\_Size 16, 58 épocas Optimizer Adam Com DropOut Sem Data Augmentation



Com esta quinta e ultima versão da rede, conseguimos uma accuracy de sensivelmente 0.84 e uma loss de 0.5. Com um batch\_size de 16, o treino foi ainda mais demorado que na versão anterior pelo mesmo motivo. Foi utilizado o optimizer Adam, que ajusta a taxa de aprendizagem (learning-rate) de cada parâmetro individualmente, tornando-o muito mais eficaz em situações onde a escala dos gradientes varia. Não foi utilizado data augmentation. Porém introduzimos também nesta rede o dropout, que desliga aleatoriamente as ligações entre neurónios entre as respetivas camadas. Efetivamente, conseguimos resultados muito melhores utilizando o dropout, um batch\_size menor e o optimizer Adam.

## Número de Épocas (Dinâmico)

Como descrito no notebook destinado a esta rede, o número de épocas definido inicialmente foi de 80 épocas. No entanto, foram utilizadas outras tecnologias, estando elas relatadas no respetivo notebook.

```
early_stopping = EarlyStopping(monitor='val_loss', patience=10, restore_best_weights=True)
reduce_lr = ReduceLROnPlateau(monitor='val_loss', factor=0.1, patience=5, min_lr=0.00001)
history = model.fit(train_dataset, epochs=80, validation_data=validation_dataset, callbacks=[early_stopping, reduce_lr])
```