1 2 9 0 UNIVERSIDADE D COIMBRA

vis team isr-uc

JULHO DE 2022

Prova de vida para imagens de faces

Carolina Pedro | Orientadores: Nuno Gonçalves e Luiz Schirmer

Autenticação Biométrica

O que é ?

Validação da identidade do utilizador através de características distintivas do corpo humano.

Aplicações

Pagamentos, check-ins, desbloqueio de telemóveis ou computadores, identificação de pessoas.

Desvantagem

Vulnerabilidade a ataques de apresentação.

Reconhecimento facial

1. Interação pouco intrusiva
2. Rapidez
3. Baixos custos

O número de aplicações de reconhecimento facial tem aumentado substancialmente.

Surgem cada vez mais tentativas de contornar este tipo de sistemas.

Existe uma grande necessidade de desenvolver técnicas para detetar e impedir ataques de apresentação.

Objetivos do trabalho

1. Abordar a aplicação de métodos de deep learning nos métodos de prova de vida.

2. Desenvolver um sistema de prova de vida usando redes neuronais convolucionais (CNN's).

3. Avaliar a eficácia do sistema, utilizando a base de dados WMCA.

Base de dados

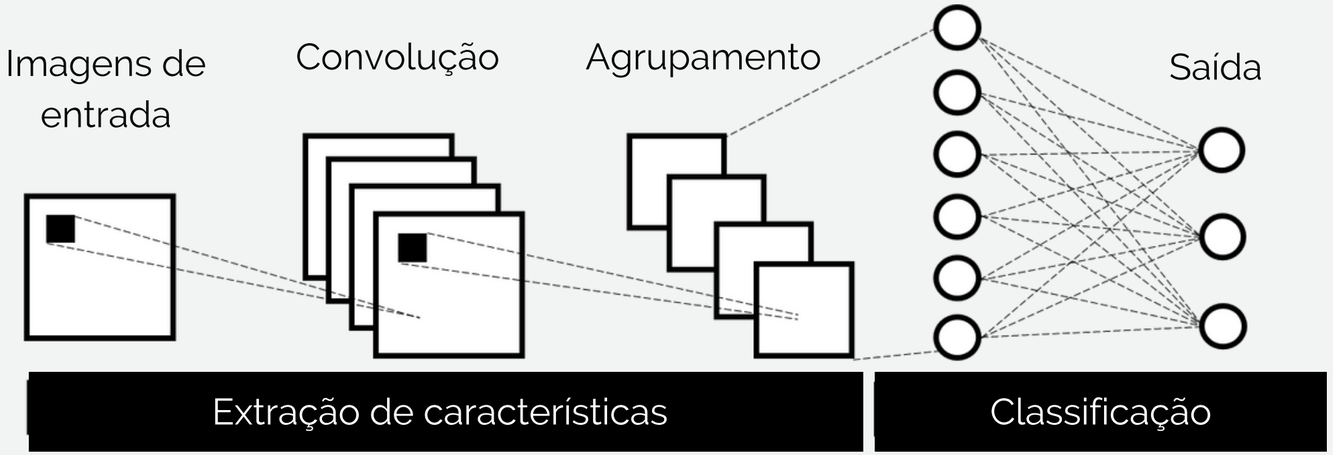
DISPONIBILIZADA PELO IDIAP

1679 pequenos vídeos

347 apresentações genuínas

1332 ataques de apresentação

Redes Neuronais Convolucionais



Elevado desempenho com dados multidimensionais, como é o caso das imagens.

Unidade básica: camadas

Redes Neuronais Convolucionais

Aumento do número de camadas

Aumento da profundidade

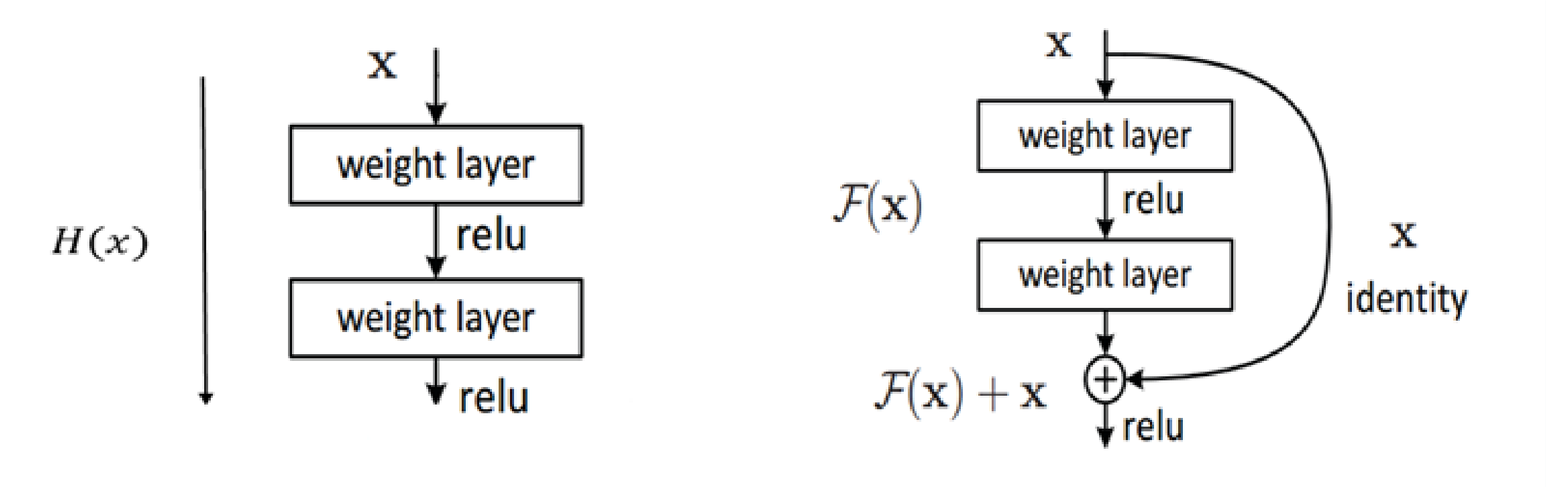
Vantagens

Permite a computação de características mais discriminantes.

Desvantagens

São difíceis de treinar. Apresentam erros de treino crescentes devido a questões na função de otimização e no gradiente de fuga.

ResNets



As ResNets têm a capacidade de resolver estes problemas devido à introdução de uma nova camada de rede neuronal, o bloco residual.

Transfer learning

- Acelera o processo de treino.
- Diminui os custos computacionais.
- Reduz o sobre-ajuste de redes de grandes dimensões.

Desafio

Escolher acertadamente o algoritmo para adaptar ao novo problema, sem alterar nenhum aspeto fundamental da rede original que possa influenciar negativamente o resultado.

ResNet-18 pré-treinada

72 camadas no total

18 camadas profundas

Nome da camada	Tamanho da saída	ResNet-18
Conv1	112 x 112 x 64	7x7, passo 2
Conv2_x	56 x 56 x 64	3x3 Max pool, passo 2
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix}$ x 2
Conv3_x	28 x 28 x 128	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix}$ x 2
Conv4_x	14 x 14 x 256	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix}$ x 2
Conv5_x	7 x 7 x 512	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix}$ x 2
	1 x 1 x 512	Average pool, 1000-d fc, softmax
FLOP's		$1{,}8 \times 10^9$

ABORDAGEM DE FRAME ÚNICO

Processamento de cada frame individualmente.

ABORDAGEM SEQUECIAL

Formação de sequências de frames para cada vídeo, havendo consideração dos frames anteriores em cada etapa.

Classificação multiclasse

Pré-processamento

Imagens de treino

Redimensionamento para 224 x 224
Rotação na horizontal aleatória
Normalização
Transformação em tensor

Imagens para validação

Redimensionamento para 224 x 224
Normalização
Transformação em tensor

Classificação multiclasse

Implementação do modelo

Abordagem de frame único

Utilização de duas camadas totalmente conectadas e uma função ReLU

Alteração da saída para 4, devido às 4 classes

Abordagem sequencial (early fusion)

Alteração na entrada da primeira convolução

Adição de uma camada de dropout de probabilidade 0,25

Alteração da saída para 4, devido às 4 classes

Classificação multiclasse

Implementação do modelo

Optimizador: ADAM com learning rate de 0,0005

Função perda: Cross Entropy Loss

Métricas de avaliação

Exatidão

$$\frac{TP + TN}{TP + FP + TN + FN}$$

Sensibilidade

$$\frac{TP}{TP + FN}$$

Precisão

$$\frac{TP}{TP + FP}$$

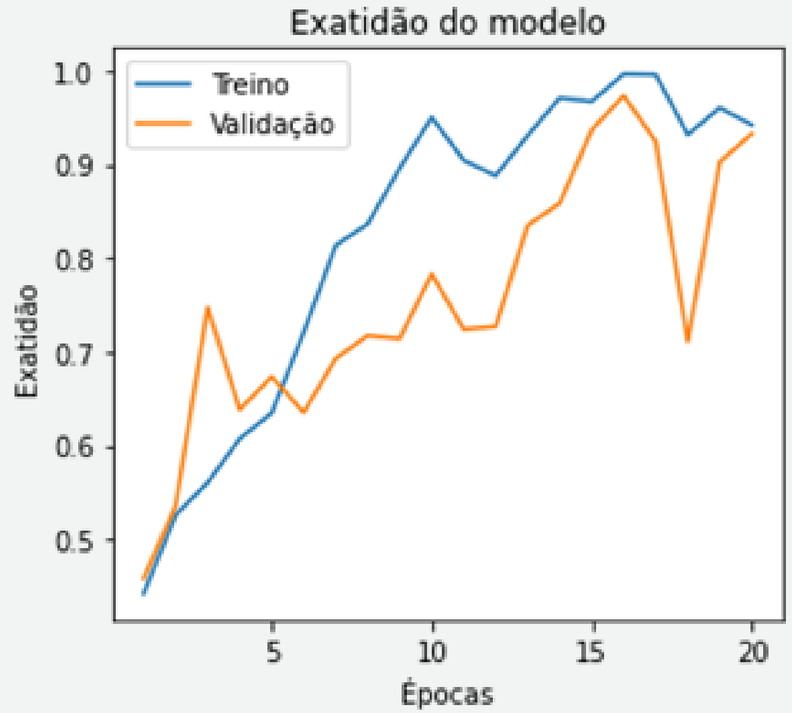
F1-score

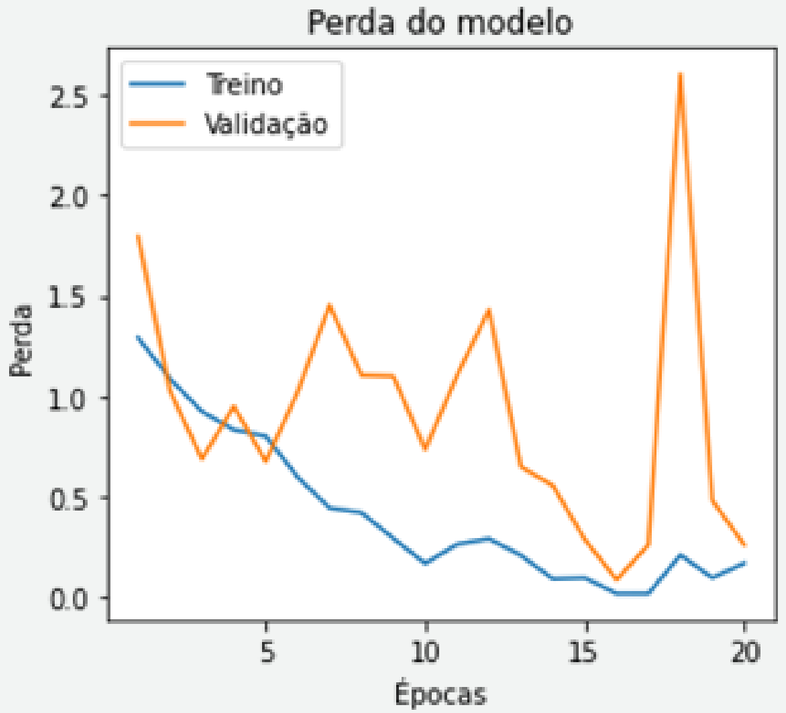
$$2 \times \frac{\text{precisão} \times \text{sensibilidade}}{\text{precisão} + \text{sensibilidade}}$$

Classificação multiclasse

Análise do modelo

Abordagem de frame único



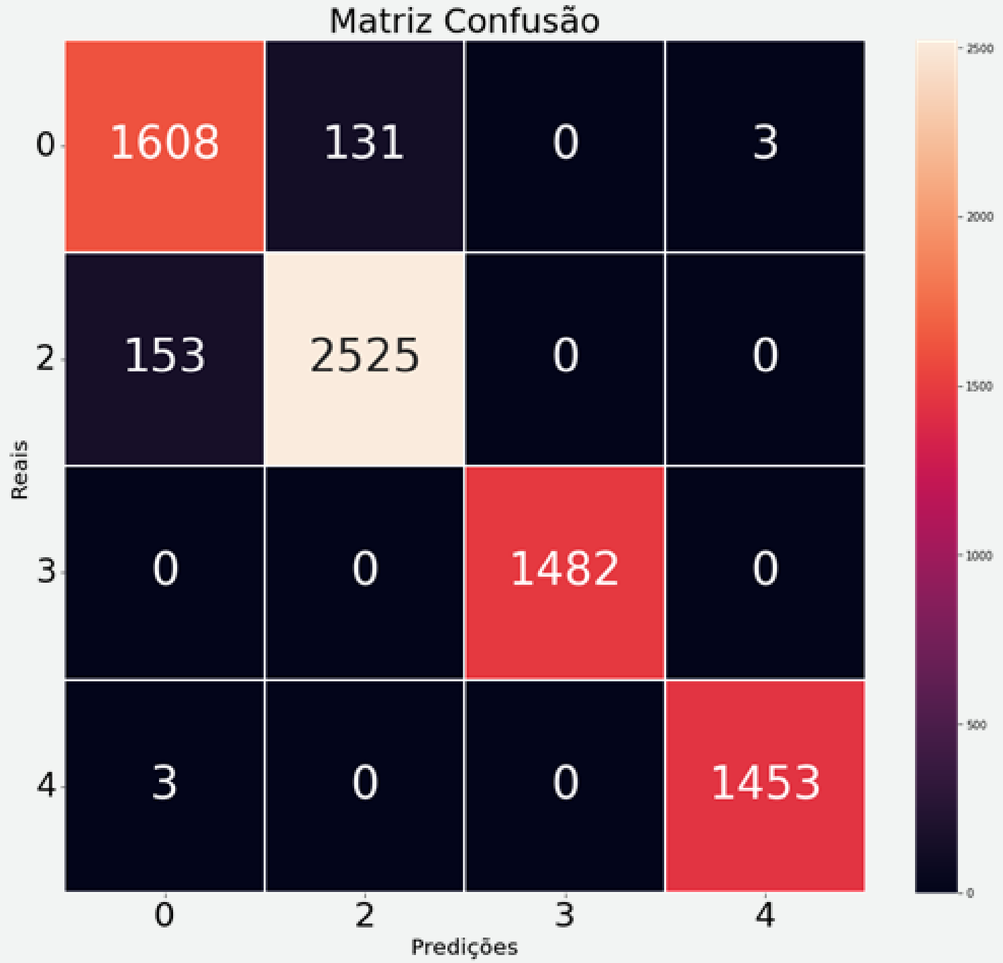


Classificação multiclasse

Análise do modelo

Abordagem de frame único

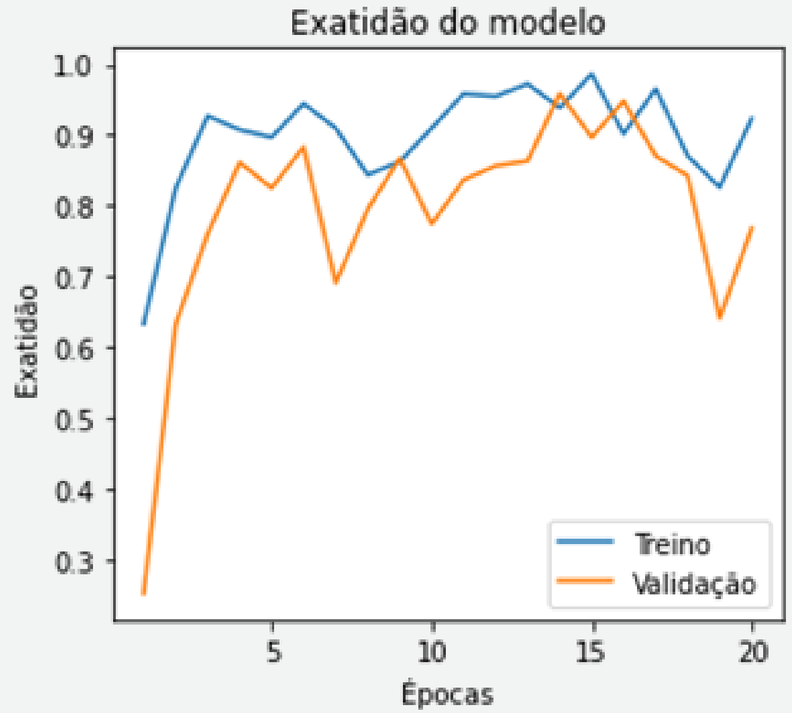
Classes	Precisão	Sensibilidade	F1-score
0	0,91	0,92	0,92
2	0,95	0,94	0,95
3	1,00	1,00	1,00
4	1,00	0,99	1,00

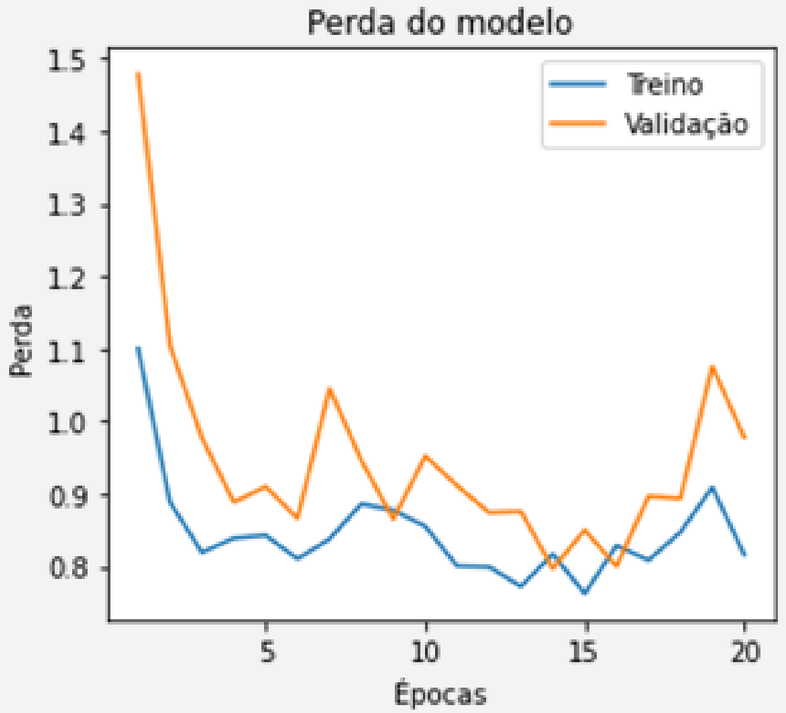


Classificação multiclasse

Análise do modelo

Abordagem sequencial



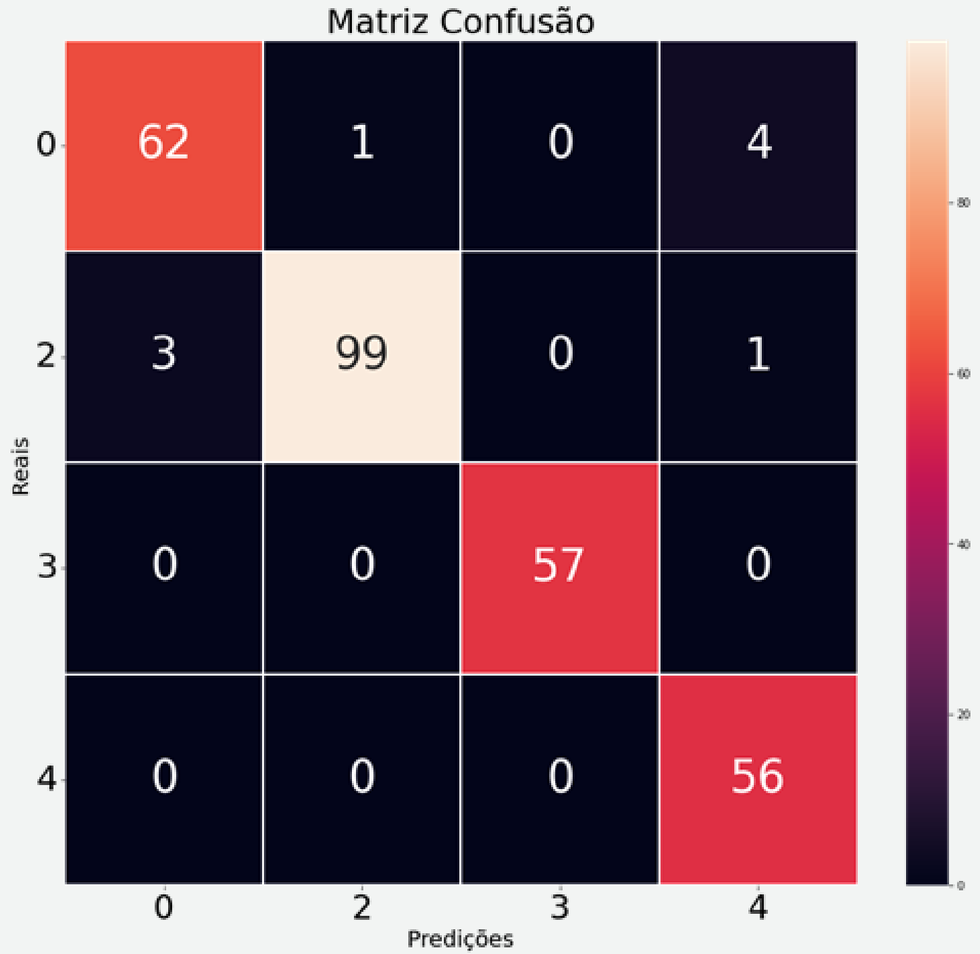


Classificação multiclasse

Análise do modelo

Abordagem sequencial

Classes	Precisão	Sensibilidade	F1-score
0	0,95	0,93	0,94
2	0,99	0,96	0,98
3	1,00	1,00	1,00
4	0,92	1,00	0,96



Classificação binária

Pré-processamento

Imagens de treino

Redimensionamento para 224 x 224
Rotação na horizontal aleatória
Normalização
Transformação em tensor

Imagens para validação

Redimensionamento para 224 x 224
Normalização
Transformação em tensor

Classificação binária

Implementação do modelo

Alterações na rede pré-treinada:

Alteração na entrada da primeira convolução

Adição de uma camada de dropout de probabilidade 0,25

Alteração da saída para 2, devido às 2 classes

Optimizador: ADAM com learning rate de 0,0005

Função perda: Cross Entropy Loss

Métricas de avaliação

Exatidão

Precisão

Sensibilidade

F1-score

ACER - Taxa Média de Erros de Classificação

$$\frac{APCER + BPCER}{2}$$

BPCER

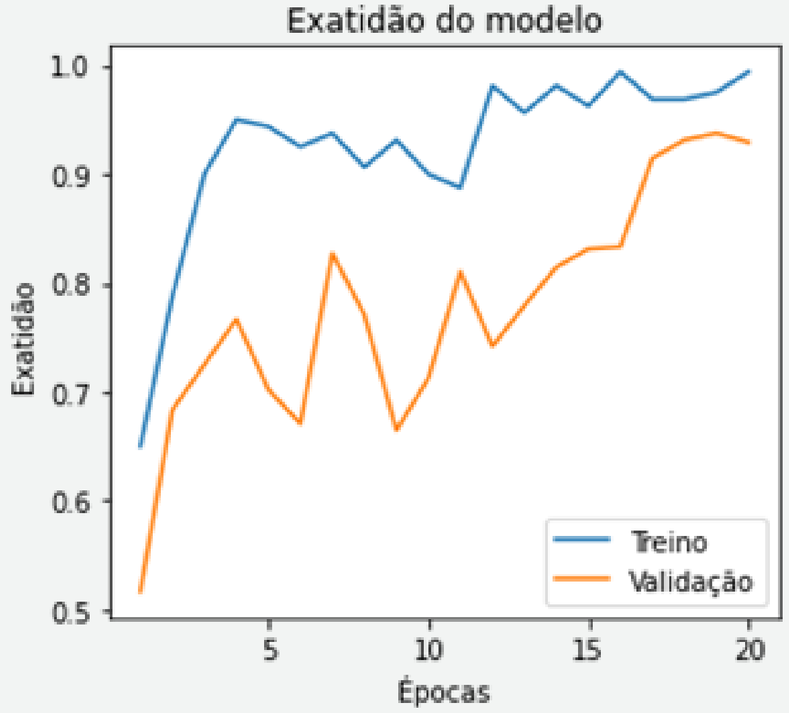
$$\frac{FN}{FN + TP}$$

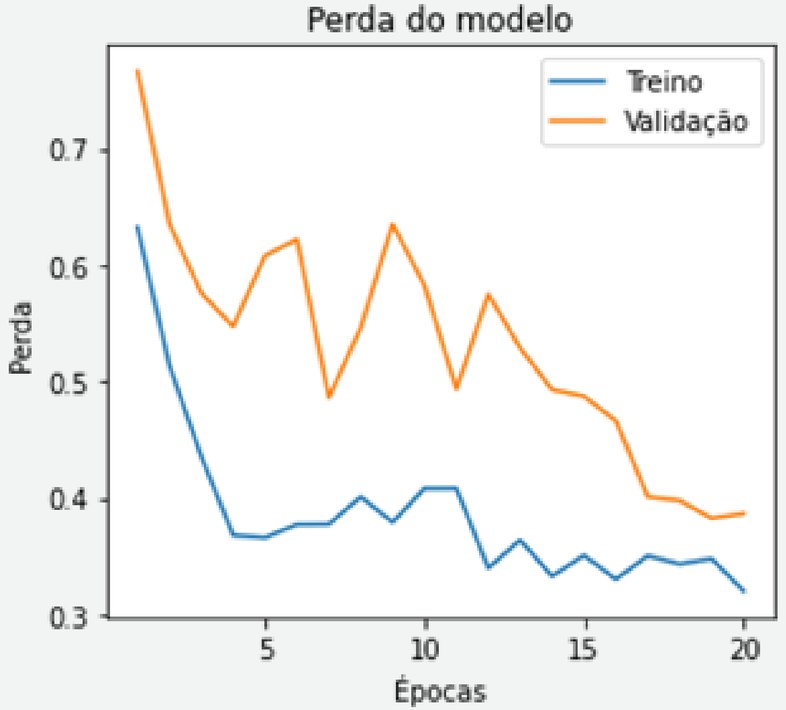
APCER

$$\frac{FP}{FP + TN}$$

Classificação binária

Análise do modelo





Classificação binária

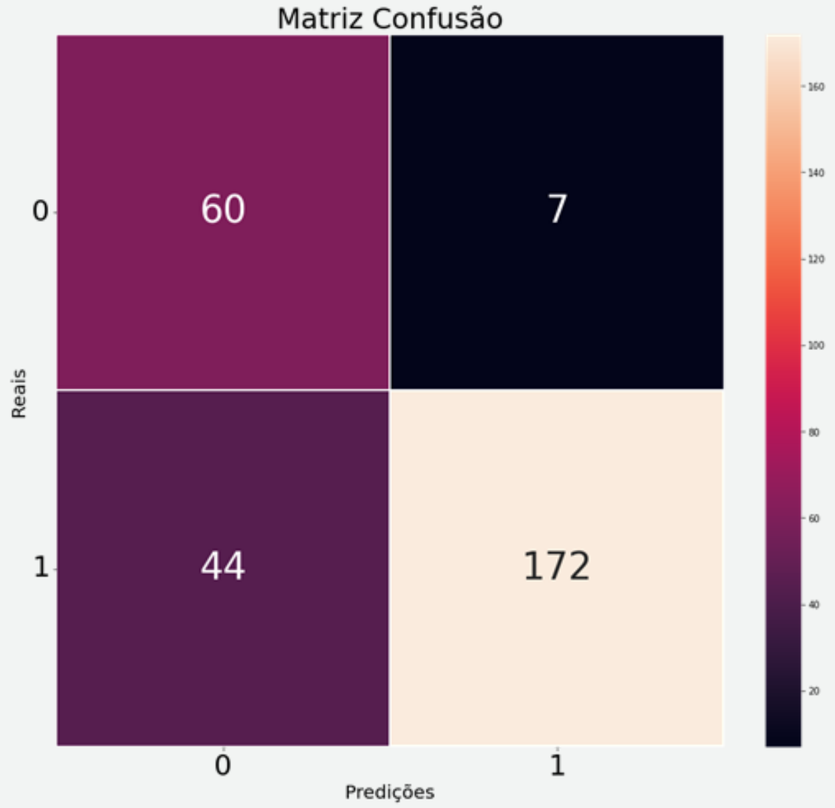
Análise do modelo

Classes	Precisão	Sensibilidade	F1-score
0	0,58	0,90	0,70
1	0,96	0,80	0,87

APCER = 0,204

BPCER = 0,154

ACER = 0,104



Conclusão

Os modelos evidenciam irregularidades na fase treino e validação.
Na fase de teste, os modelos de classificação multiclasse apresentam resultados satisfatórios, já o modelo binário apresenta uma incapacidade de assumir uma grande parte de ataques de apresentação como tal.
Poderiam ser alvo de algumas melhorias, o que reflete o facto de prova de vida ser ainda um problema em aberto, havendo uma constante necessidade de evolução.