

Yasmim Santana,¹ Victor Sarrís, Maria Clara,¹ Isac Brito¹, Rafael Leite¹, Joaquim de Moura²

¹Instituto Federal de Educação, Ciência e Tecnologia do Piauí (IFPI), Floriano, Brasil

² Grupo VARPA, INIBIC, Universidade da Coruña (UDC), A Coruña, Espanha

INTRODUÇÃO

Este projeto tem como foco a aplicação de técnicas de Deep Learning na classificação de gênero a partir de imagens faciais. Trata-se de uma tarefa relevante em diferentes contextos, como análise de mídias sociais, publicidade direcionada e segurança. O desafio central está nas variações de iluminação, expressões faciais e qualidade das imagens, que podem impactar a precisão do processo. ►

Objetivos e Motivações: A correta identificação do gênero pode gerar avanços significativos em áreas como marketing e análise demográfica, oferecendo insights mais precisos e experiências personalizadas. Assim, este trabalho busca investigar métodos robustos e eficientes capazes de lidar com a complexidade e aplicabilidade desse problema.

FIGURA 1



Figura 1: Amostra de imagens do conjunto de dados, com sua classificação correspondente.

METODOLOGIA

- **A. Conjunto de Dados:** Foi utilizado o *CelebFaces Attributes* (CelebA) Dataset (Liu et al., 2018), composto por 500 imagens de celebridades (250 homens e 250 mulheres). Cada amostra possui um rótulo de gênero binário (0 para mulheres e 1 para homens), com resolução de 178×218 pixels. O conjunto foi dividido em três partes:

55

- **B. Arquitetura da Rede:** A rede escolhida foi a ResNet-18 (He et al., 2016), que introduz as chamadas *skip connections*, permitindo ao modelo pular camadas durante o treinamento. Essa característica possibilita treinar redes muito profundas sem perda de desempenho. Para a linha de base, foram utilizados pesos pré-treinados no ImageNet (Deng et al., 2009).

- **C. Hiperparâmetros:**

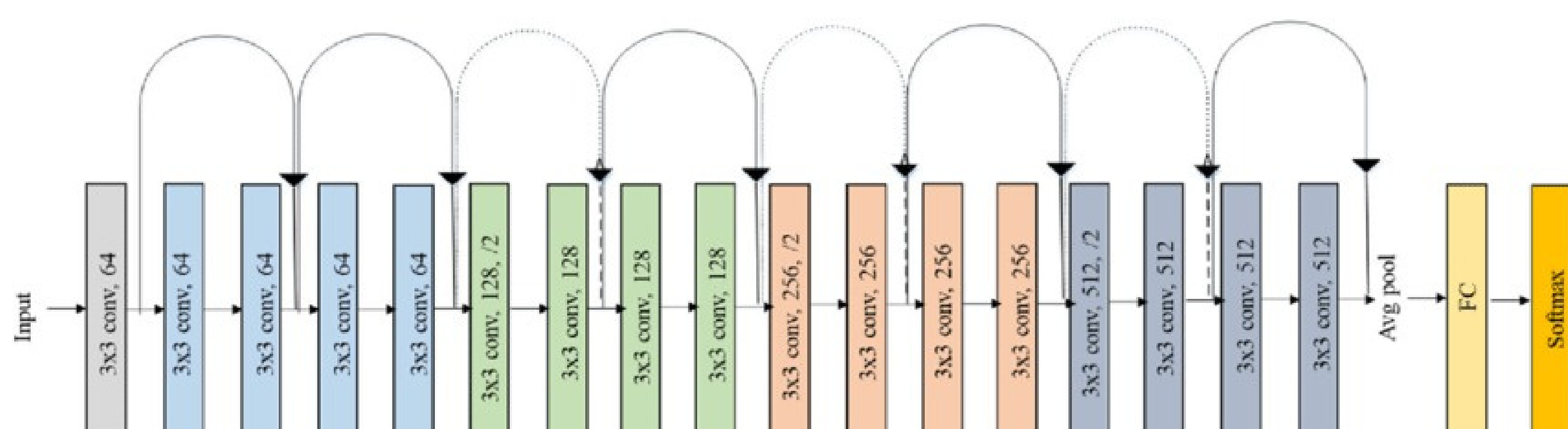
- **Função de perda (Loss):** Binary Cross-Entropy (BCE), definida como:

$$BCE = -\frac{1}{N} \sum_{i=1}^N (y_i \log(p_i) + (1 - y_i) \log(1 - p_i))$$

onde y_i representa o rótulo verdadeiro, p_i a probabilidade prevista, e N o número de amostras.

- **Otimizador:** Adam (Adaptive Moment Estimation)
- **Taxa de aprendizado:** 0.0001
- **Épocas:** 30

FIGURA 2



MÉTRICAS DE AVALIAÇÃO

- **Precision:**

$$\text{Precision} = \frac{TP}{TP + FP}$$

onde TP é o número de verdadeiros positivos e FP é o número de falsos positivos.

- **Recall:**

$$\text{Recall} = \frac{TP}{TP + FN}$$

onde TP é o número de verdadeiros positivos e FN é o número de falsos negativos.

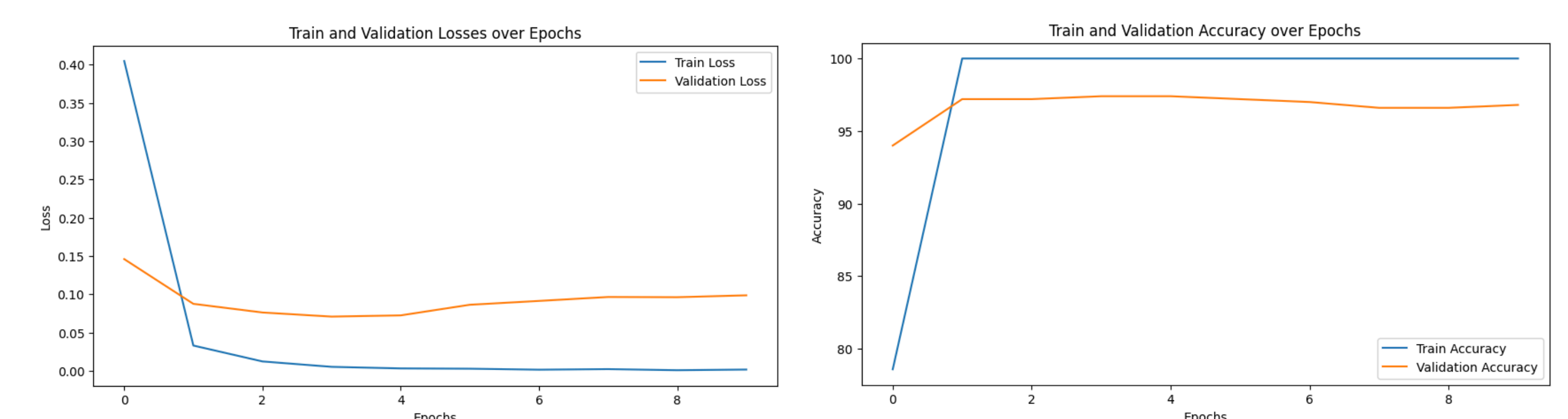
- **F-Score:**

$$\text{F-Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

onde Precision é a precisão calculada como acima, e Recall é a exatidão calculada como acima.

RESULTADOS E DISCUSSÃO

FIGURA 3

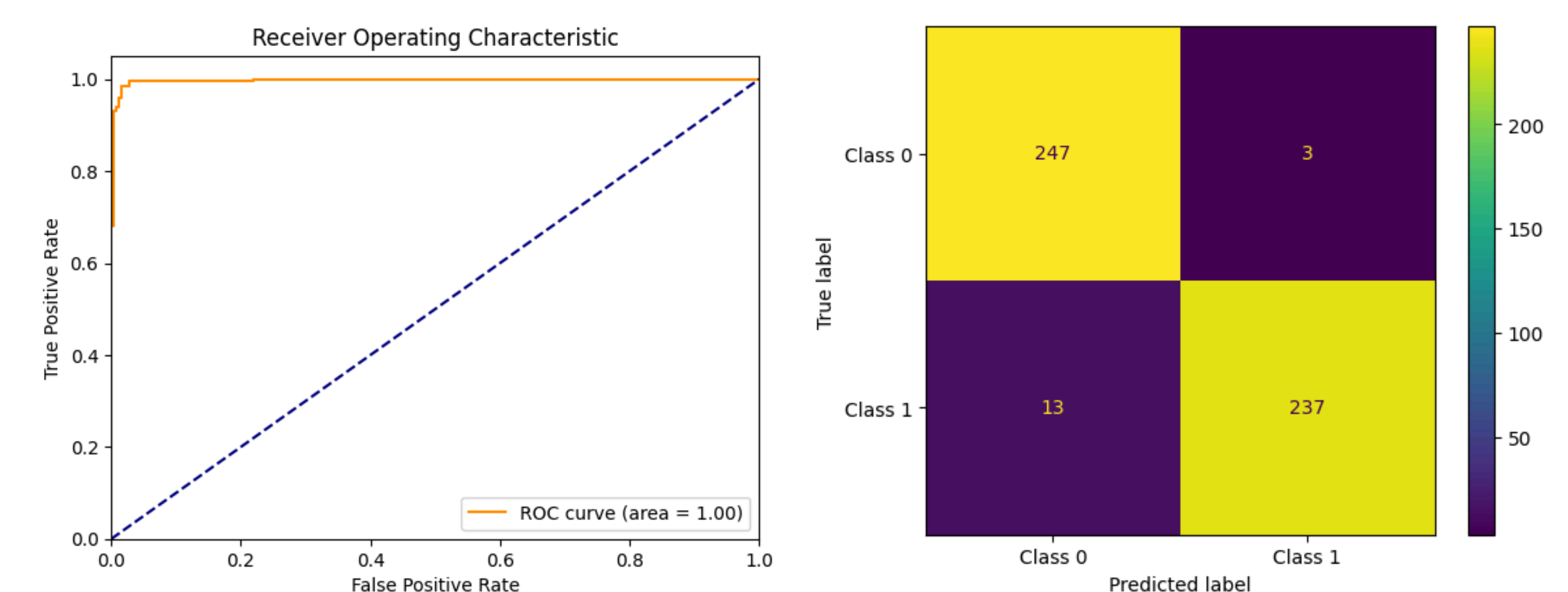


(a)

(b)

Figura 3: (a) Evolução do loss de treino e validação. (b) Evolução da precisão de treino e validação.

FIGURA 4



(a)

(b)

Figura 4: (a) Curva ROC. (b) Matriz de confusão para a fase de teste do modelo.

FIGURA 5

	Precision	Recall	F-Score
ResNet-18	0.99	0.95	0.97

Figura 5: Resultados de teste para o modelo ResNet-18.

REFERÊNCIAS

Deng, Jia et al. (2009). "Imagenet: A large-scale hierarchical image database". Em: *2009 IEEE conference on computer vision and pattern recognition*. Ieee, pp. 248–255.

He, Kaiming et al. (2016). "Deep residual learning for image recognition". Em: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.