

INTRODUÇÃO

Este projeto explora o uso de métodos de Deep Learning para a classificação de gênero a partir de imagens faciais. Essa tarefa possui relevância em diversos contextos, como análise de mídia social, publicidade direcionada e segurança. Identificar o gênero com base em características faciais é um desafio, especialmente devido às variações na iluminação, expressões faciais e qualidade das imagens. Os objetivos e motivações específicas deste trabalho incluem:

- **Objetivos e Motivações:** A identificação precisa do gênero tem potencial para revolucionar áreas como marketing e análise demográfica, fornecendo insights valiosos e personalização. Este projeto investiga abordagens robustas e eficientes para enfrentar este problema complexo e altamente aplicável.

FIGURA 1



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

METODOLOGIA

- **A. Conjunto de Dados:** CelebFaces Attributes (CelebA) Dataset (Liu et al., 2018) contém imagens de 500 rostos de 500 celebridades diferentes (250 homens e 250 mulheres). Cada imagem está associada a um rótulo de gênero (0 para mulheres, 1 para homens). As imagens têm um tamanho de 178×218 pixels. O conjunto original foi dividido em três conjuntos: treinamento, validação e teste, sendo 55%, 15% e 30% respectivamente.
- **B. Arquitetura da Rede:** A arquitetura selecionada foi a ResNet-18 (He et al., 2016). Esse tipo de arquitetura introduz as “conexões de salto”, que permitem que o modelo pule uma ou mais camadas. Essa abordagem torna possível treinar a rede em milhares de camadas sem afetar o desempenho. Para a linha de base, foram utilizados pesos pré-treinados da base de dados ImageNet (Deng et al., 2009).
- **C. Hiperparâmetros:**
 - **Loss:** Binary Cross-Entropy (BCE) loss function:

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

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

- **Optimizer:** Adaptive Moment (Adam) estimation optimizer
- **Learning rate:** 0.0001
- **Epochs:** 30

FIGURA 2

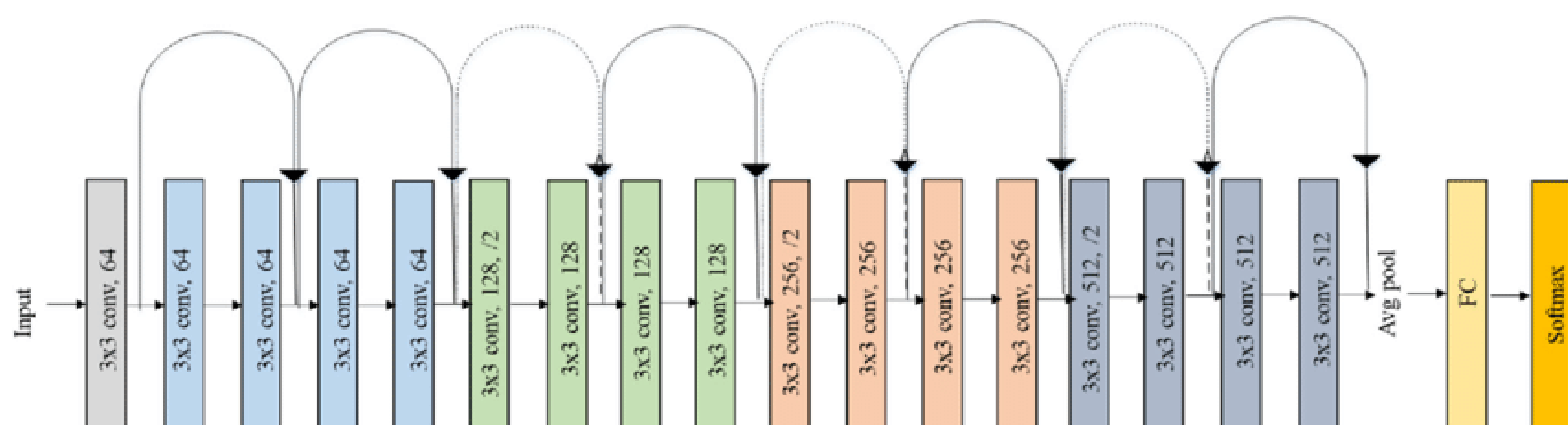


Figura 2: Arquitetura ResNet-18.

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

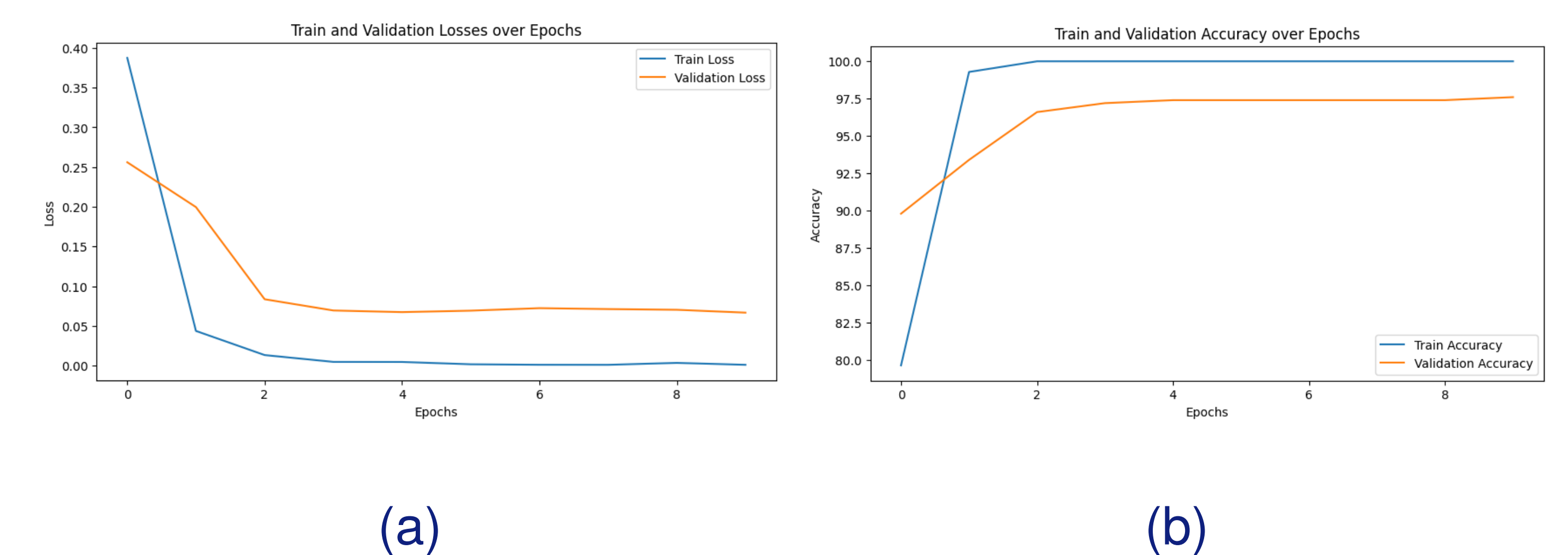


Figura 3: (a) Evolução do loss de treino e validação. (b) Evolução da precisão de treino e validação.

FIGURA 4

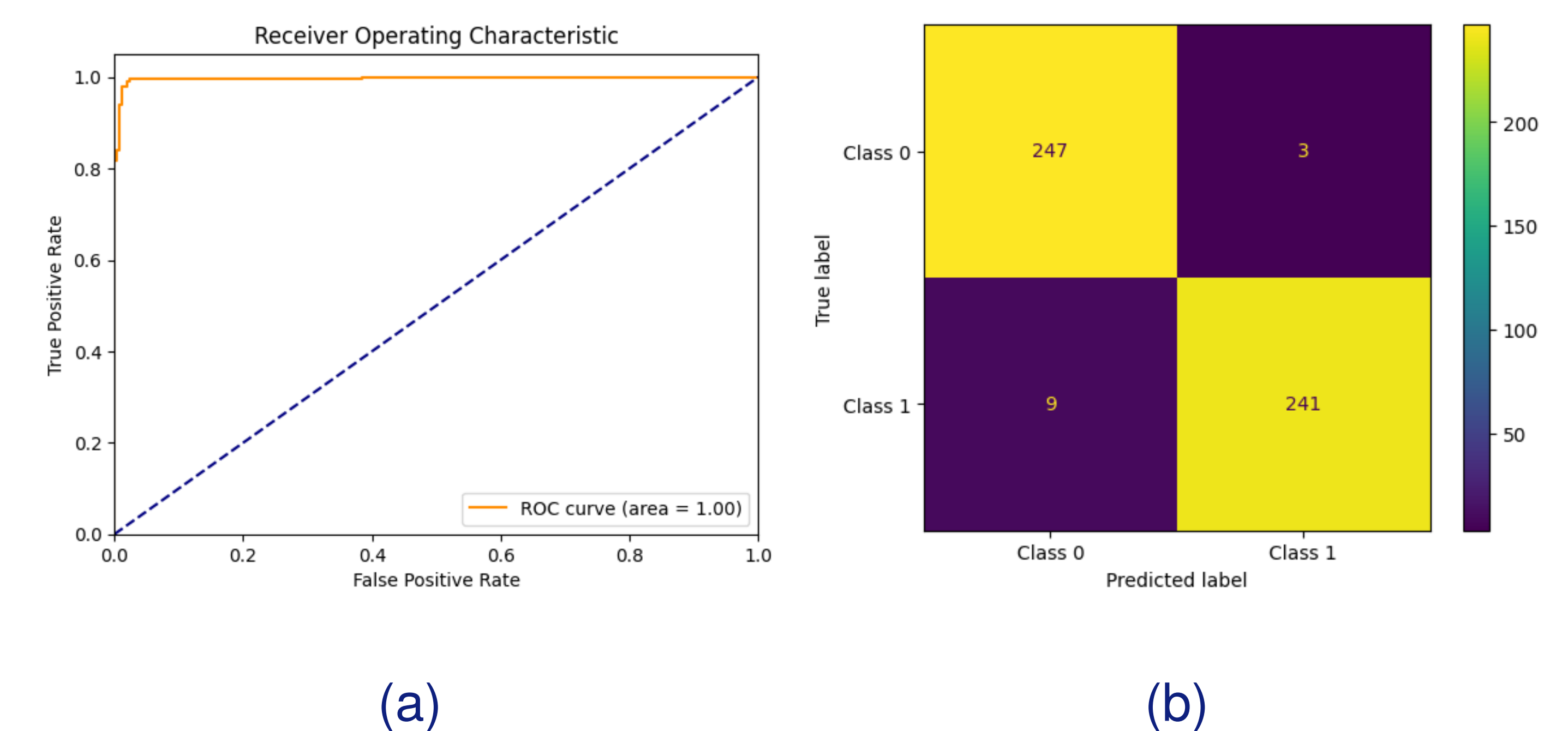


Figura 4: (a) Curva ROC. (b) Matriz de confusão para a fase de teste do modelo.

FIGURA 5

	Precision	Recall	F-Score
ResNet-18	0.99	0.96	0.98

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.
- Liu, Ziwei et al. (2018). “Large-scale celebfaces attributes (celeba) dataset”. Em: *Retrieved August 15.2018*, p. 11.