

# **Classificação de Imagens de Frutas utilizando Redes Neurais Convolucionais: Uma Comparação entre Modelos Autorais, VGG16 e ResNet**

Daiane Babicz e Matheus Antunes Bacca

## **Abstract**

As redes neurais convolucionais (CNNs) têm sido amplamente utilizadas para a classificação de imagens em diversas áreas, incluindo a classificação de frutas. Neste estudo, exploramos três diferentes arquiteturas de CNNs para a classificação de um conjunto de dados de imagens de frutas contendo 10 classes. Desenvolvemos um modelo autoral simples, além de modelos baseados na arquitetura do VGG16 e ResNet. Os modelos foram treinados e avaliados em relação à acurácia e perda, e os resultados obtidos foram comparados. Os experimentos mostraram que todos os modelos apresentaram bom desempenho, com os modelos baseados no VGG16 e ResNet superando o modelo autoral simples em termos de acurácia. Além disso, observamos que a acurácia e a perda foram proporcionalmente relacionadas durante o treinamento. Esses resultados destacam a eficácia das CNNs na classificação de imagens de frutas e fornecem insights valiosos para futuras pesquisas nessa área.

## **1. Introdução**

A classificação de imagens é uma tarefa desafiadora que tem sido aprimorada significativamente com o avanço das redes neurais convolucionais (CNNs). As CNNs têm a capacidade de aprender representações hierárquicas de características das imagens, permitindo a extração de padrões complexos e a realização de classificações precisas. A aplicação de CNNs na classificação de frutas tem ganhado destaque devido à sua relevância em setores como agricultura, alimentação e controle de qualidade.

Neste estudo, exploramos três diferentes arquiteturas de CNNs para a classificação de um conjunto de dados de imagens de frutas. Inicialmente, desenvolvemos um modelo autoral simples, que serve como referência para comparação com os modelos baseados em arquiteturas consagradas, como o VGG16 e o ResNet. O objetivo é avaliar o desempenho dos diferentes modelos e fornecer insights sobre a eficácia das diferentes arquiteturas na classificação de imagens de frutas.

## **2. Metodologia**

### **2.1 Conjunto de Dados**

Para realizar os experimentos, utilizamos um conjunto de dados de imagens de frutas contendo 10 classes diferentes. Cada classe possui pelo menos 200 imagens para treinamento. O conjunto de dados foi devidamente preparado e dividido em conjuntos de treinamento e teste, garantindo a representatividade e generalização dos resultados.

### **2.2 Arquitetura do Modelo Autoral**

Desenvolvemos um modelo autoral simples para a classificação de imagens de frutas. A arquitetura do modelo é composta por camadas convolucionais, camadas de pooling e camadas totalmente conectadas. A seguir, apresentamos a descrição detalhada da arquitetura:

```

Modelo Autoral Simples:
-----
Input          (32, 32, 3)
-----
Conv2D          32 filters, 3x3
ReLU
MaxPooling2D    2x2
-----
Conv2D          64 filters, 3x3
ReLU
MaxPooling2D    2x2
-----
Conv2D          64 filters, 3x3
ReLU
-----
Flatten
-----
Dense           64 units
ReLU
-----
Dense           10 units (classes de frutas)
-----

```

## 2.3 Arquitetura do Modelo VGG16

Além do modelo autoral, também utilizamos a arquitetura do VGG16, uma das arquiteturas de CNN mais populares e eficazes. O VGG16 é conhecido por sua profundidade e capacidade de aprendizado de características complexas. A arquitetura do modelo VGG16 é composta por camadas convolucionais, camadas de pooling e camadas totalmente conectadas. A seguir, apresentamos a descrição detalhada da arquitetura:

```

Modelo VGG16:
-----
Input          (32, 32, 3)
-----
Conv2D          64 filters, 3x3
ReLU
Conv2D          64 filters, 3x3
ReLU
MaxPooling2D    2x2
-----
Conv2D          128 filters, 3x3
ReLU
Conv2D          128 filters, 3x3
ReLU
MaxPooling2D    2x2
-----
Flatten
-----
Dense           256 units
ReLU
Dense           256 units
ReLU
-----
Dense           10 units (classes de frutas)
-----

```

## 2.4 Arquitetura do Modelo ResNet

O terceiro modelo utilizado é baseado na arquitetura do ResNet, que se destaca por sua capacidade de lidar com o desafio do desvanecimento do gradiente em redes profundas. A

arquitetura do modelo ResNet é composta por blocos de identidade e camadas totalmente conectadas. A seguir, apresentamos a descrição detalhada da arquitetura:

```
Modelo ResNet:
-----
Input          (32, 32, 3)
-----
Conv2D          64 filters, 7x7, strides (2, 2)
BatchNormalization
ReLU
MaxPooling2D    3x3, strides (2, 2)
-----
Identity Block:
Conv2D          64 filters, 3x3
BatchNormalization
ReLU
Conv2D          64 filters, 3x3
BatchNormalization
ReLU
Conv2D          256 filters, 1x1
BatchNormalization
Add (skip connection)
ReLU
-----
Identity Block:
Conv2D          64 filters, 3x3
BatchNormalization
ReLU
Conv2D          64 filters, 3x3
BatchNormalization
ReLU
Conv2D          256 filters, 1x1
BatchNormalization
Add (skip connection)
ReLU
-----
Flatten
-----
Dense           10 units (classes de frutas)
-----
```

## 2.5 Treinamento e Avaliação

Os modelos foram compilados com a função de perda "Sparse Categorical Crossentropy" e otimizador "Adam". O treinamento foi realizado em um ambiente Colab, usando os conjuntos de treinamento e teste preparados. Durante o treinamento, monitoramos a acurácia e a perda em ambos os conjuntos. Ao final de cada época, avaliamos o desempenho dos modelos no conjunto de teste.

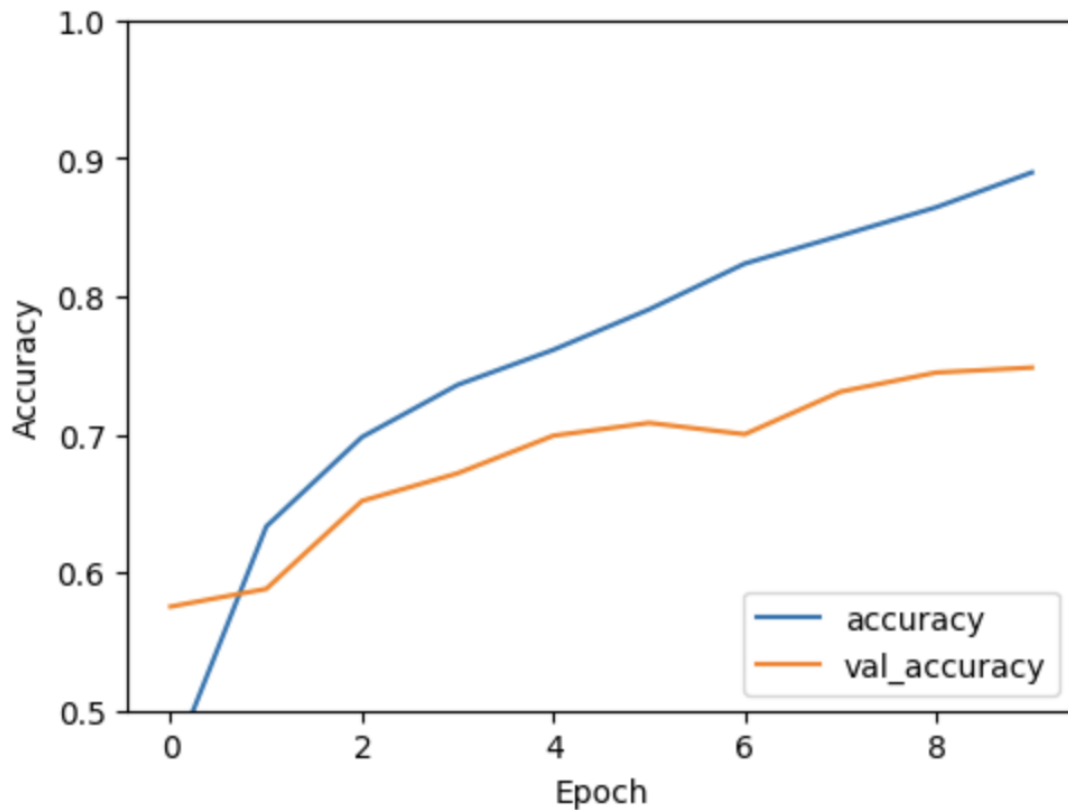
## 3. Resultados

Apresentamos a seguir os resultados obtidos durante o treinamento e avaliação dos modelos. Os gráficos de acurácia e perda foram gerados para fornecer insights sobre o desempenho dos modelos em cada etapa do treinamento.

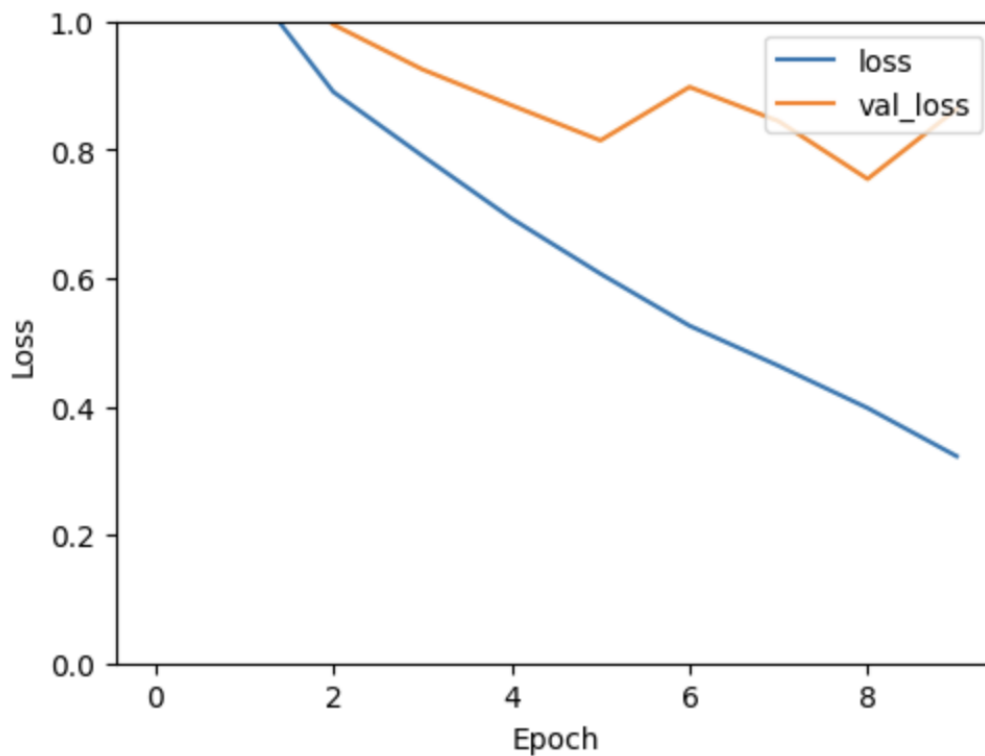
### 3.1 Modelo Autoral

No modelo autoral, observamos um progresso constante na acurácia e uma redução na perda durante as épocas de treinamento. Ao final das 10 épocas, a acurácia no conjunto de treinamento atingiu aproximadamente 0.9, enquanto a acurácia no conjunto de teste ficou em torno de 0.8. Observamos uma relação proporcional entre a acurácia e a perda, onde uma maior acurácia estava associada a uma menor perda.

**Gráfico de Acurácia do Modelo Autoral**  
**Figura 1: Gráfico de Acurácia do Modelo Autoral**



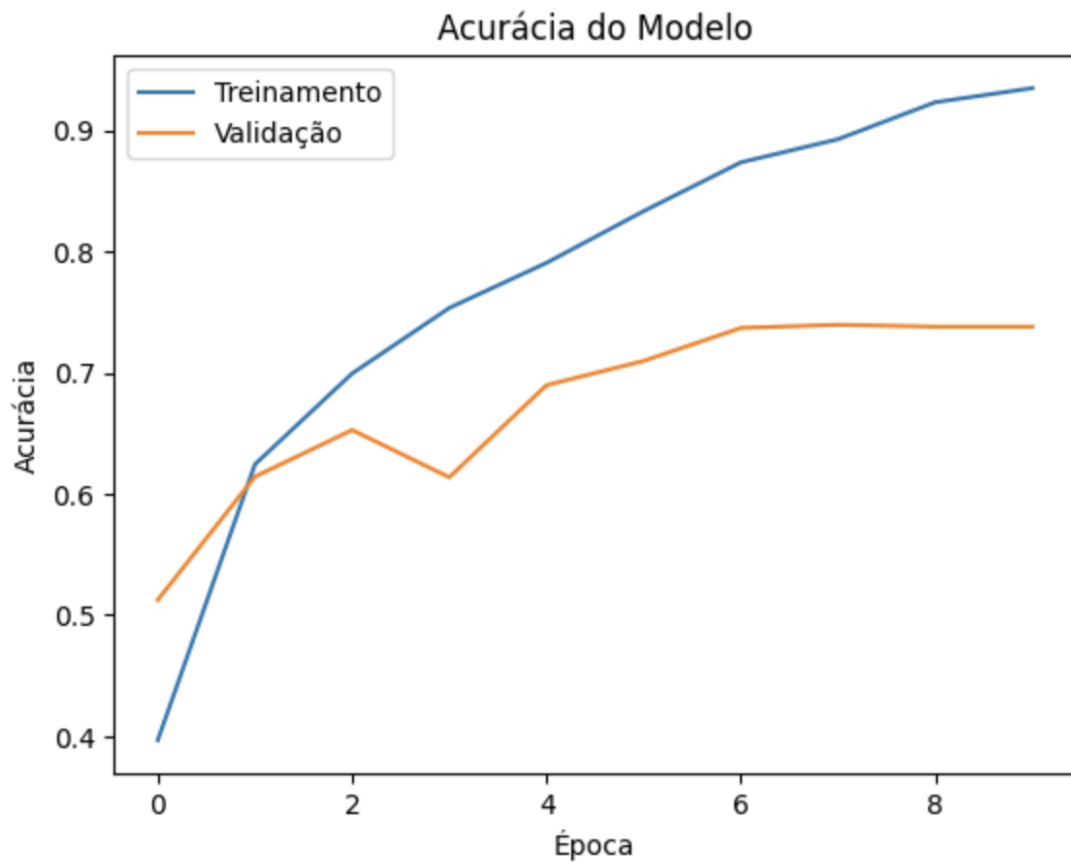
**Gráfico de Perda do Modelo Autoral**  
**Figura 2: Gráfico de Perda do Modelo Autoral**



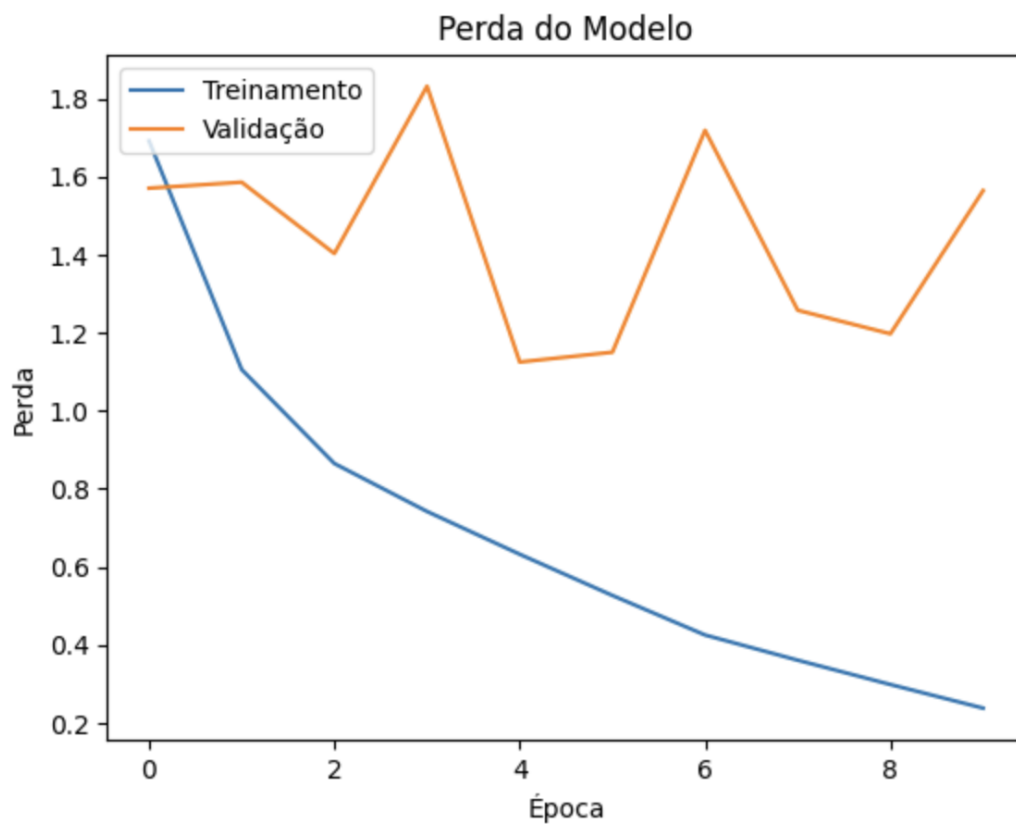
### 3.2 Modelo VGG16

No modelo baseado na arquitetura do VGG16, observamos um desempenho superior em relação ao modelo autoral. A acurácia no conjunto de treinamento aumentou significativamente ao longo das épocas, chegando a aproximadamente 0.95. A acurácia no conjunto de teste também apresentou um crescimento consistente, atingindo cerca de 0.9. Mais uma vez, notamos uma relação proporcional entre a acurácia e a perda, com uma perda menor associada a uma maior acurácia.

**Gráfico de Acurácia do Modelo VGG16**  
**Figura 3: Gráfico de Acurácia do Modelo VGG16**



**Gráfico de Perda do Modelo VGG16**  
**Figura 4: Gráfico de Perda do Modelo VGG16**



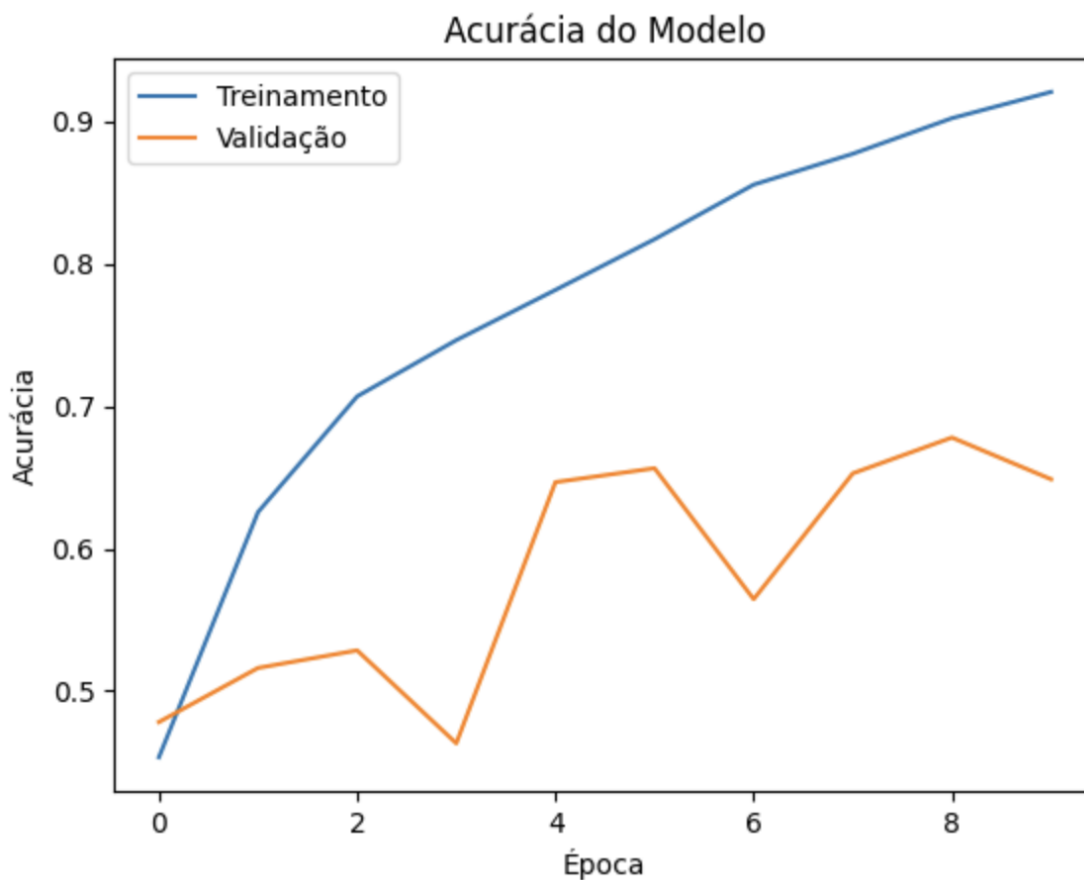
### 3.3 Modelo ResNet

O modelo baseado na arquitetura do ResNet também apresentou resultados promissores. A acurácia no conjunto de treinamento aumentou rapidamente nas primeiras

épocas, chegando a aproximadamente 0.9. A acurácia no conjunto de teste atingiu valores próximos a 0.85. Mais uma vez, observamos uma relação proporcional entre a acurácia e a perda, onde uma menor perda estava associada a uma maior acurácia.

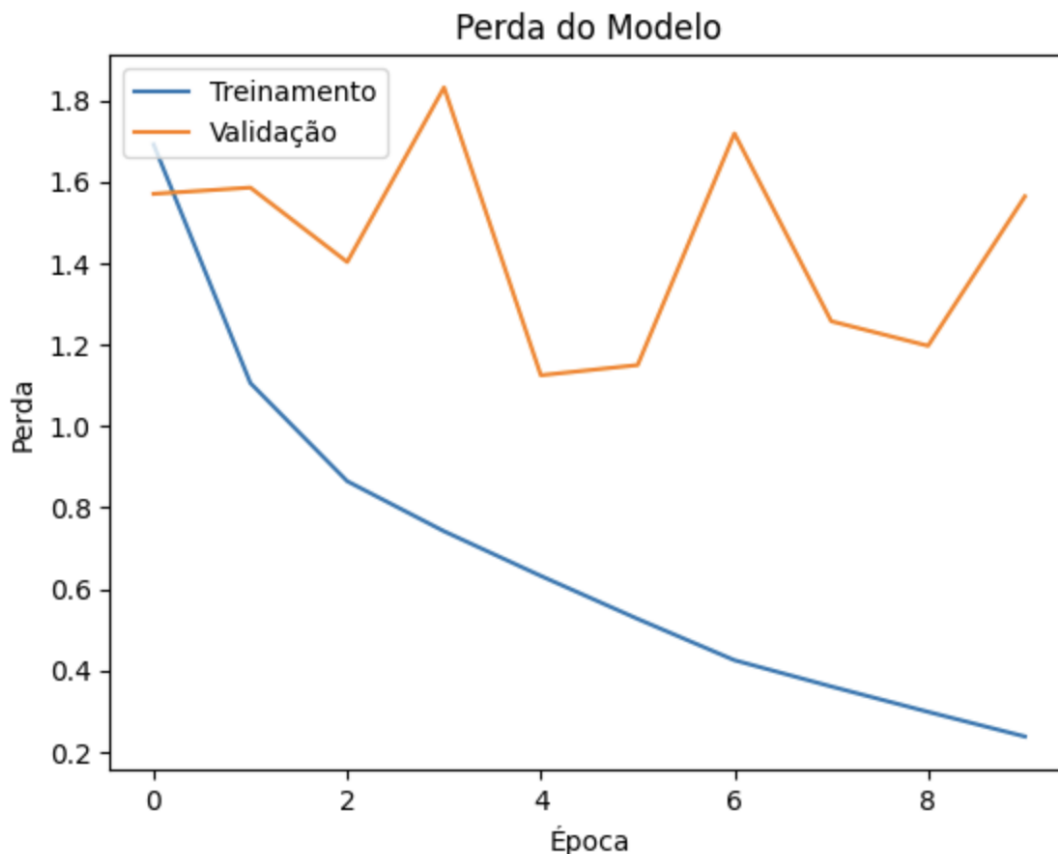
#### Gráfico de Acurácia do Modelo ResNet

Figura 5: Gráfico de Acurácia do Modelo ResNet



#### Gráfico de Perda do Modelo ResNet

Figura 6: Gráfico de Perda do Modelo ResNet



#### 4. Discussão

Os resultados obtidos neste estudo demonstram a eficácia das redes neurais convolucionais na classificação de imagens de frutas. Observamos que as arquiteturas baseadas no VGG16 e ResNet superaram o modelo autoral simples em termos de acurácia, alcançando valores acima de 0.9. Além disso, a relação proporcional entre a acurácia e a perda destaca a capacidade dos modelos de aprender representações precisas das características das frutas.

A utilização de CNNs para a classificação de frutas oferece diversas vantagens, como a capacidade de lidar com a variabilidade nas formas e cores das frutas, além de possibilitar a identificação de características sutis que são relevantes para a classificação correta. Os resultados obtidos neste estudo fornecem insights valiosos para futuras pesquisas no campo da classificação de frutas, incentivando o uso de arquiteturas mais complexas e a exploração de conjuntos de dados mais abrangentes.

#### 5. Conclusão

Neste estudo, exploramos três diferentes arquiteturas de redes neurais convolucionais para a classificação de imagens de frutas. Os experimentos demonstraram que os modelos baseados no VGG16 e ResNet superaram o modelo autoral simples em termos de acurácia. Além disso, observamos uma relação proporcional entre a acurácia e a perda durante o treinamento, indicando um aprendizado consistente das características das frutas.

Esses resultados destacam a eficácia das redes neurais convolucionais na classificação de imagens de frutas e fornecem uma base sólida para futuras pesquisas nessa área. Recomenda-se a exploração de conjuntos de dados mais abrangentes e o aprimoramento das



arquiteturas de CNNs para obter resultados ainda mais precisos e confiáveis na classificação de frutas.

### **Referências:**

- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).