

The Ember House

Inteligência Artificial Aplicada
à Qualidade do Café



Time



Felipe Israel dos Santos
R.A.: 10729570



Juliana de
Oliveira Sato
R.A.: 10727587



Yuri dos Santos Radziwill
R.A.: 10730741

Introdução

- **Café:** uma das bebidas mais consumidas do mundo.
- **Brasil:** maior exportador e 2º maior consumidor mundial.
- **Crescimento** do consumo de cafés especiais e gourmet.
- **Desafio:** avaliar qualidade dos grãos de forma rápida e precisa.
- **Solução:** IA com TensorFlow para identificar padrões visuais em grãos.



Objetivos do Projeto

- **Treinar uma IA que classifica imagens de grãos de café por tipo e torra.**
- **Otimizar reconhecimento e controle de qualidade.**
- **Aumentar a rentabilidade e fortalecer a marca no setor.**

The Ember House

A empresa une Ciência de Dados e Inteligência Artificial para transformar a forma como o café é analisado e valorizado.

Desenvolvemos soluções baseadas em análise de imagens e modelos preditivos capazes de identificar automaticamente o tipo e a qualidade dos grãos de café, auxiliando produtores, torrefadoras e distribuidores na tomada de decisões mais precisas e sustentáveis.

Missão

Oferecer soluções tecnológicas que garantam padrões elevados de qualidade do café, apoiando toda a cadeia produtiva na tomada de decisões mais rápidas, precisas e sustentáveis.

Valores

Ser reconhecida como a empresa líder em análise de dados para o setor cafeeiro, contribuindo para a modernização do mercado e para a valorização do café brasileiro no cenário global.

Visão

- Inovação
- Qualidade
- Sustentabilidade
- Parceria.

Dados

- Fonte: Coffee Bean Dataset Resized (Kaggle)
- 4 classes de torra: verde, clara, média e escura
- 1.600 imagens (400 por classe)
- Formato: PNG / CSV
- Padrão fixo: 224×224 pixels



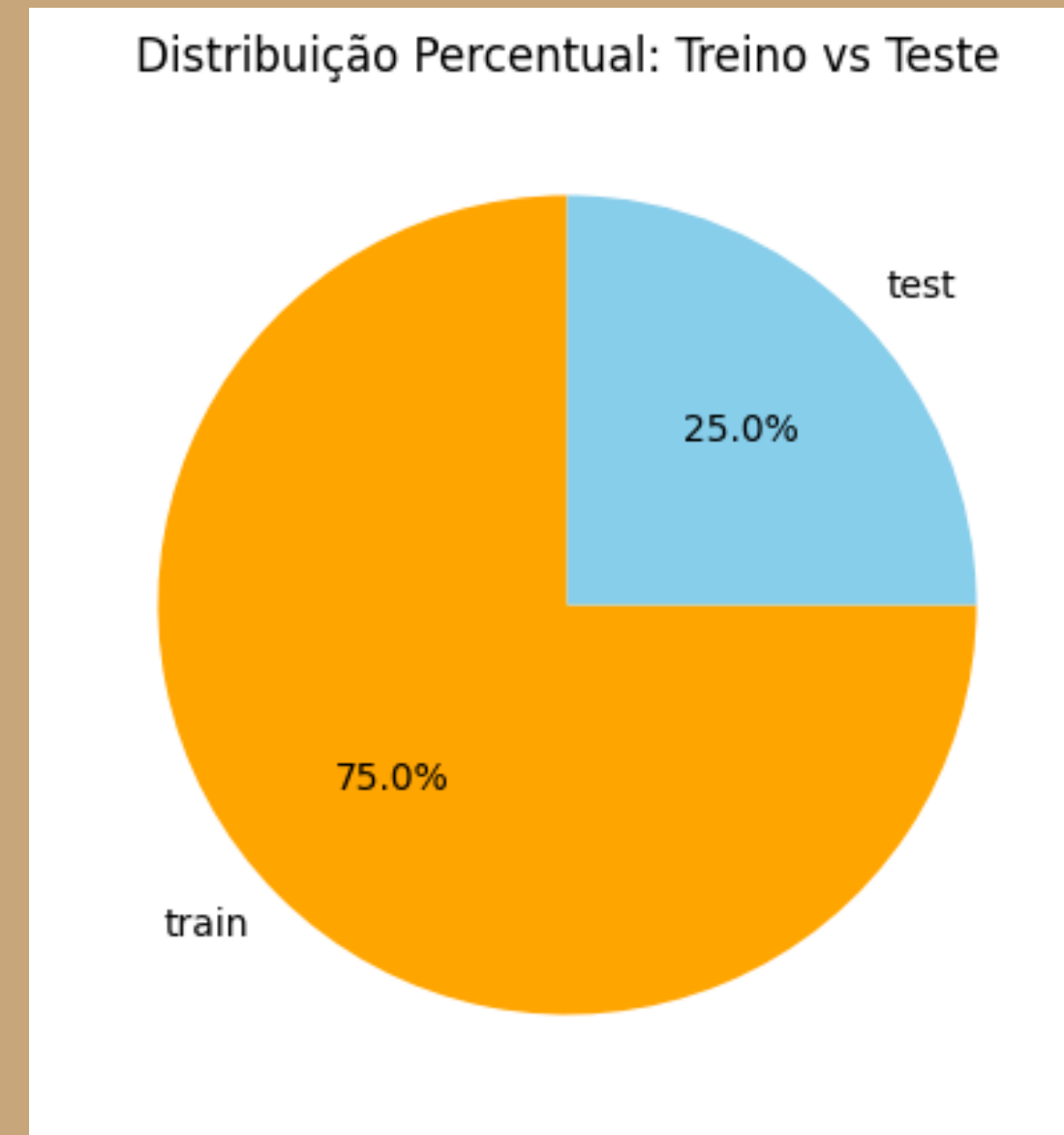
Análise Exploratória

Objetivo: Antes do treinamento do modelo, foi realizada uma análise exploratória para avaliar a consistência, integridade e balanceamento da base de imagens.



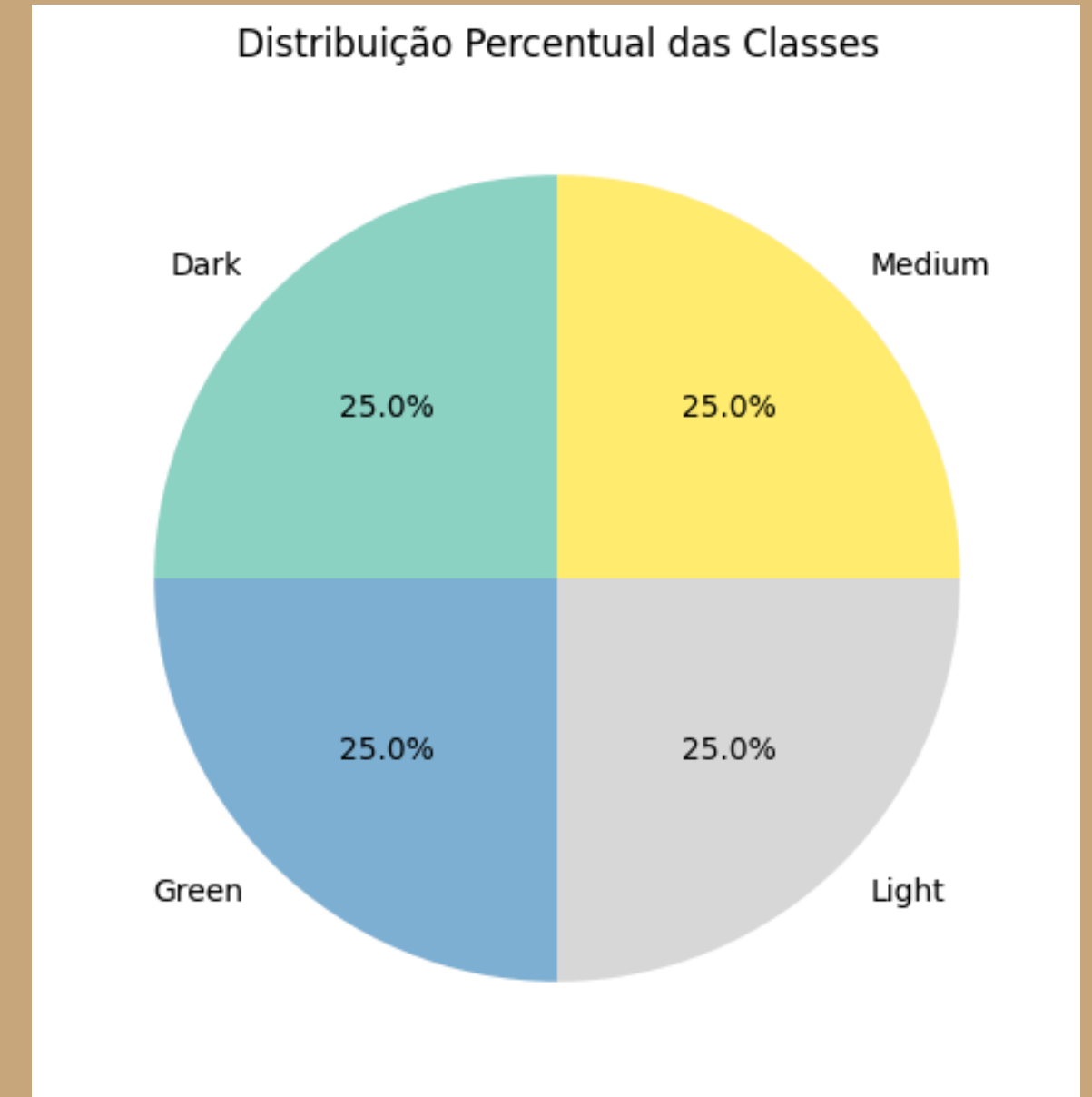
Divisão dos dados no Dataset:
75% treino (1.200 imagens)
25% teste (400 imagens)

Divisão de dados para testes:
70% treino (1.120 imagens)
30% teste (480 Imagens)



Principais pontos:

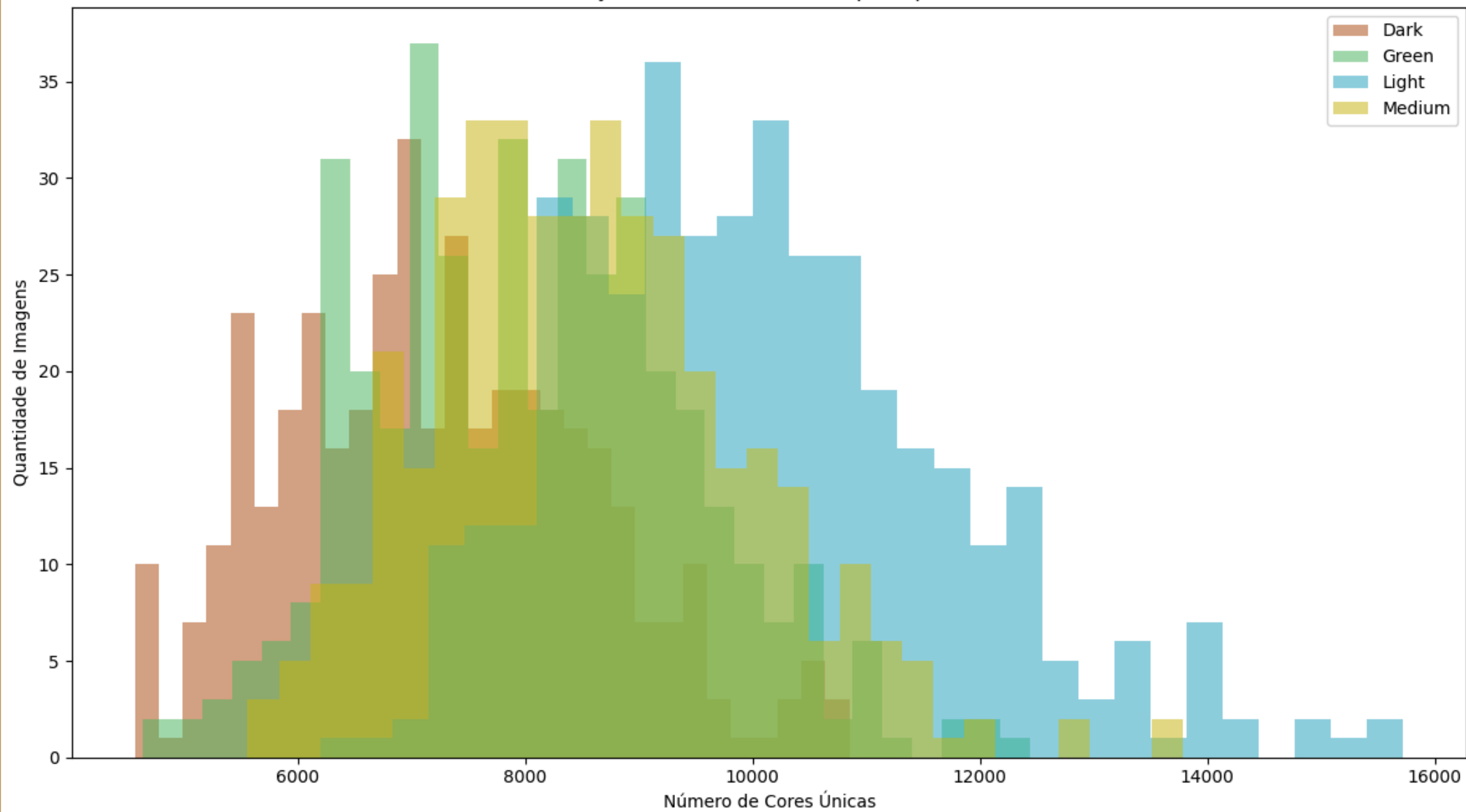
- Dataset: 1.600 imagens, divididas igualmente entre 4 classes (Green, Light, Medium, Dark).
- Balanceamento: Cada classe representa 25% do total — garantindo um aprendizado sem viés.
- Arquivos duplicados: Nenhum encontrado — dados íntegros e consistentes.
- Padrão de nomes: 26 caracteres em média — nomenclatura uniforme e ideal para manipulação programática.



Análise técnica das imagens:

- Tamanho médio: 94,68 KB (variação de apenas 4,8%) → baixa variabilidade.
- Dimensão das imagens: 224 × 224 pixels → padrão uniforme, sem necessidade de redimensionamento.
- Número médio de cores:
 - Green → 8.048
 - Light → 10.046
 - Medium → 8.548
 - Dark → 7.231
- O equilíbrio nas dimensões e cores garante melhor estabilidade no aprendizado da CNN.

Distribuição do Número de Cores por Tipo de Terra



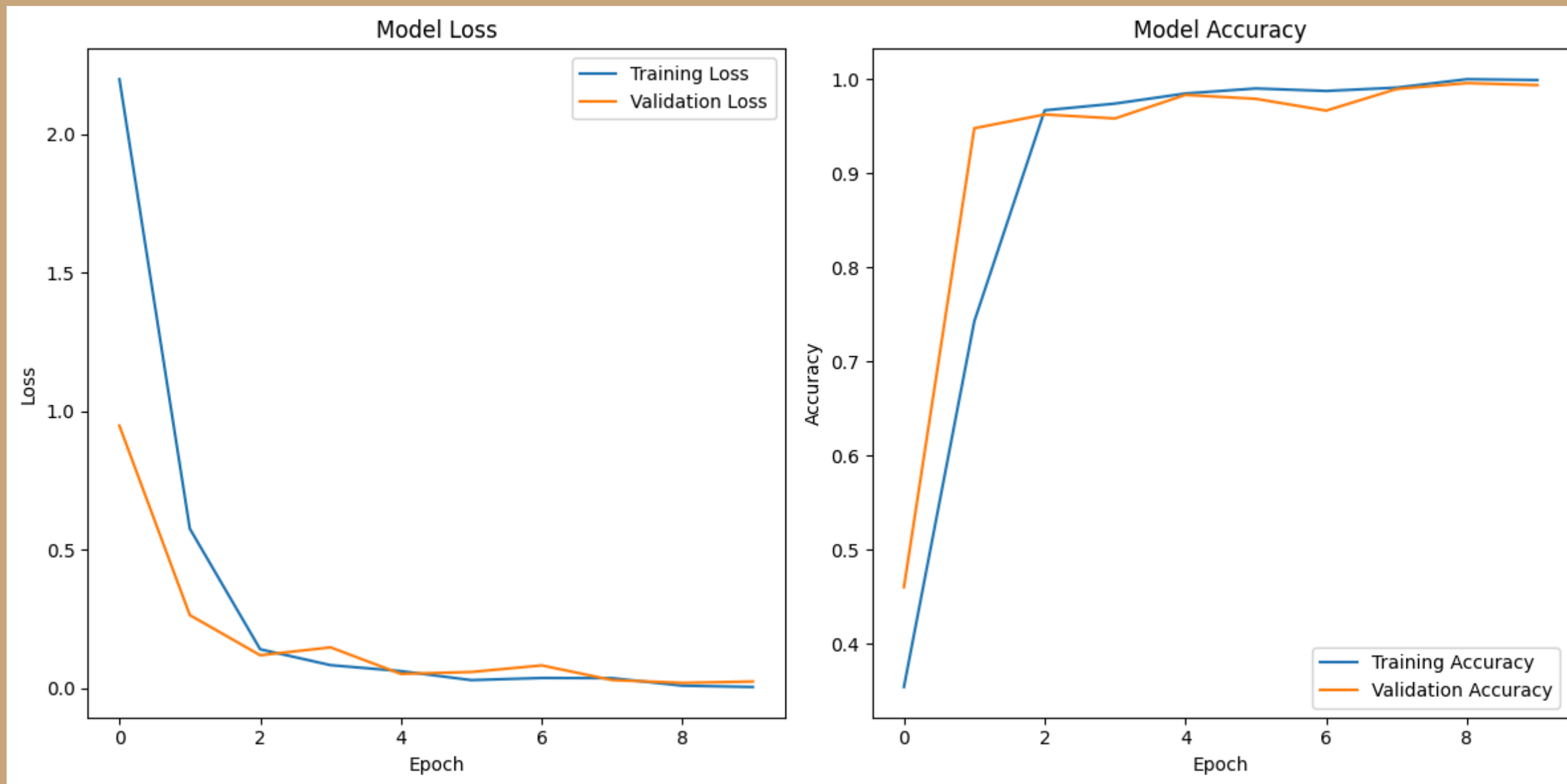
- Bibliotecas: Pandas, Matplotlib, Seaborn, TensorFlow, Keras
- Rede Neural Convolucional (CNN):
- Camadas Conv2D, Pooling, Dense
- Função de perda: `sparse_categorical_crossentropy`
- Otimizador: Adam
- Callback: Early Stop para evitar overfitting

Modelo

Treinamento

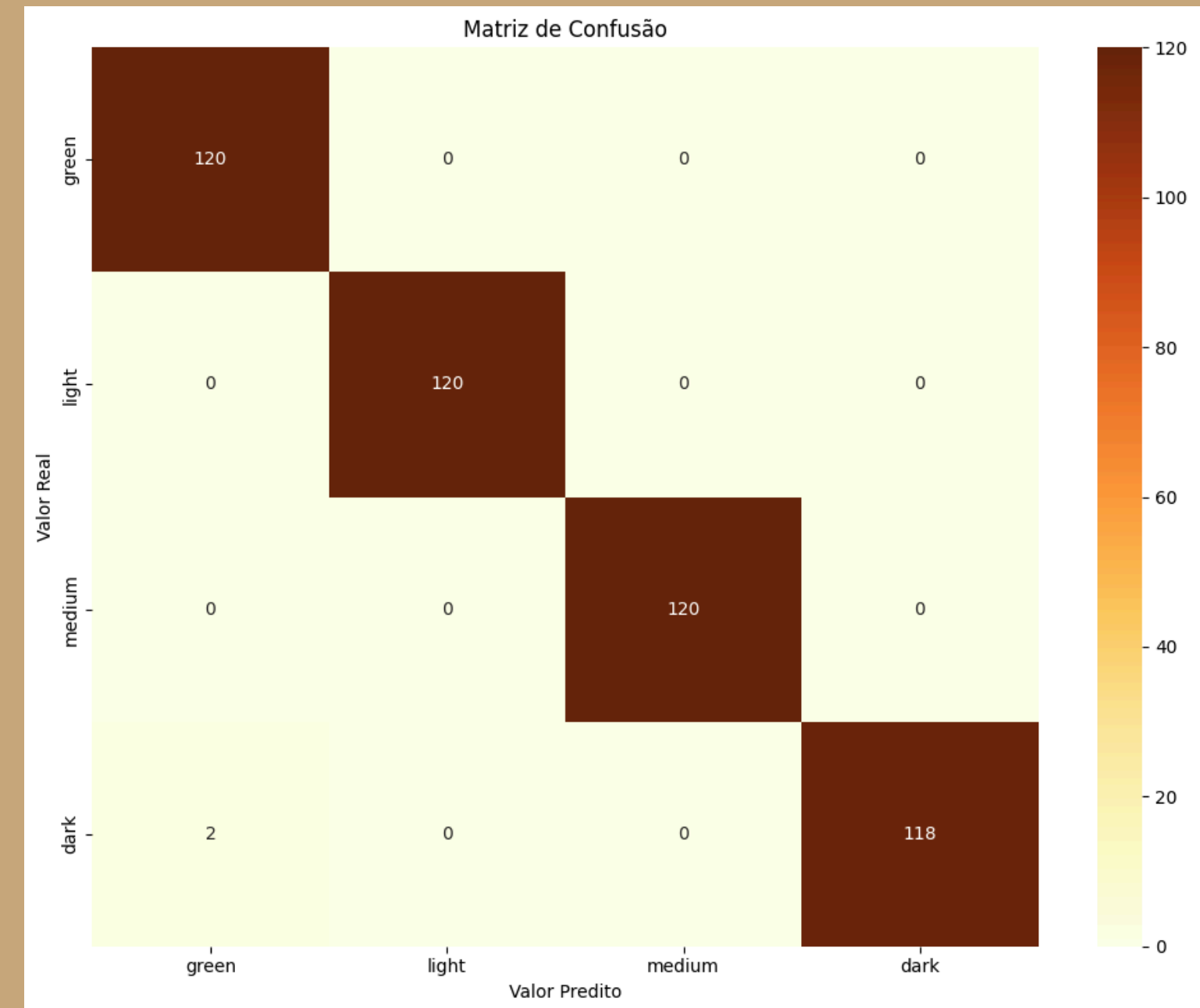
- 70% treino / 30% teste
- 10 épocas de aprendizado
- Curvas de loss e accuracy estáveis
- Variação estocástica $\approx 1,5\%$

Evolução do treinamento



Resultados e Acurácia

- Acurácia média: 99,58%
- Classes com 100%: Green, Light, Medium
- Classe com 98,3%: Dark
- Matriz de confusão



Resultados Preliminares

- IA classifica tipos de grãos com alta precisão.
- Aplicação direta em controle de qualidade automatizado.
- Modelo de negócio SaaS (assinatura para produtores e torrefadoras).
- Geração de dashboards e relatórios interativos.