

# **MODELO DE MACHINE LEARNING PARA DETECÇÃO E CONTAGEM DE PÉS DE CAFÉ (*Coffea sp.*) POR ANÁLISE DE VÍDEO.**

ANDREW GABRIEL MARCOLINO FERNANDES<sup>1</sup>; FELIPE GABRIEL NOVAIS JONAS<sup>1</sup>; JOÃO RICARDO FAVAN<sup>2</sup>

<sup>1</sup>Discentes do curso Tecnólogo em Big Data no Agronegócio, Faculdade de Tecnologia “Shunji Nishimura”, Pompéia - São Paulo, felipenjonas@gmail.com, andrew8gmf@gmail.com.

<sup>2</sup>Professor Doutor, Docente em Big Data no Agronegócio na Faculdade de Tecnologia “Shunji Nishimura”, Pompéia - São Paulo, joao.favan@factec.sp.gov.br

## **RESUMO**

O setor agrícola está cada vez mais aprimorando tecnologias e construindo novas ferramentas inteligentes, dessa forma, gerando melhores índices de desempenho na produtividade, qualidade e lucratividade. Diversos tipos de sensores já fazem parte da construção de máquinas agrícolas e seus resultados têm contribuído para a conquistas das melhores métricas. Tendo em vista a importância de tecnologias com foco em monitoramento na etapa de colheita, foi estudado a implementação de um algoritmo de aprendizado de máquina para detecção dos cafeeiros presentes nas linhas de plantio, e em seguida uma contagem cumulativa por meio de uma gravação em vídeo. O desenvolvimento da pesquisa fundamentou-se nas seguintes etapas para sua realização, sendo elas: preparação das imagens, configuração dos modelos pré treinados para treinamento personalizado aos cafeeiros, treinamento do modelo, teste do modelo com imagens capturadas pelos autores na Faculdade de Tecnologia “Shunji Nishimura” na cidade de Pompeia, Implementação do modelo para detecção em de cafeeiros no formato de vídeo e desenvolvimento da contagem cumulativa desses indivíduos. O resultado gerado foi um modelo de aprendizado de máquina treinado para detecção de cafeeiros por vídeo e também sendo possível a contagem cumulativa de cada unidade na linha do plantio. Contudo, a adoção deste tipo de tecnologia pode vir a somar de forma positiva nas análises de produtividade por área plantada e trabalhar em conjunto com especialistas e empresas que queiram dispor desta tecnologia em seus equipamentos.

**Palavras-chave:** Detecção de objeto. Aprendizado de máquina. Contagem de plantas.

## 1. INTRODUÇÃO

No Brasil a cafeicultura possui importância fundamental, tanto histórica quanto econômica. Por volta de 1790, as primeiras mudas do café chegaram na região do Estado do Rio de Janeiro e São Paulo, nos anos seguintes, o plantio do café foi adentrando o interior do país sendo desenvolvido também em outros Estados como Minas Gerais e estabelecendo sua posição no mercado mundial de exportações de grãos (EXPORTAÇÃO, [s.d]). Observando dados sobre as exportações brasileiras de café por continente no período que corresponde entre os meses janeiro e junho do ano de 2021 a Europa, América do Norte e Ásia são os 3 continentes que mais importaram esta *commodity* (matéria-prima), sendo respectivamente o volume por sacas de 60 quilogramas iguais a 10.732.993 milhões, 4.691.620 milhões e 3.320.871 milhões (CONAB, 2021,p.13). No cenário econômico brasileiro as exportações em ano-safra no período de julho de 2020 a junho de 2021 foi de 45,6 milhões de sacas de 60 quilogramas, e uma receita cambial de 5,84 bilhões de dólares (CONAB, 2021, p.24).

Após breve observação dos dados, foi entendido que para alcançar estes resultados e apresentá-los de maneira sucinta, demandou esforço e empenho de muitos profissionais para realizar análises dos dados. Dessa forma, o *Big Data* (manipulação de grandes volumes de dados) tem conquistado mais relevância nos negócios e os profissionais estão desempenhando muito bem no processo de análise e tomada de decisão (PEMBERTON, 2015). Outras tecnologias como aprendizado de máquina e inteligência artificial também estão sendo utilizadas no setor agrícola para ajudar em previsões mais assertivas.

Aprendizado de máquina ou também *machine learning* é uma área dentro da inteligência artificial que permite com que sistemas ou algoritmos sejam capazes de adquirir conhecimento de maneira automatizada. Sua utilização no setor agrícola proporciona análises mais assertivas e antecipadas, oferecendo ao agricultor a oportunidade e tempo de direcionar recursos no momento e local mais necessitado (OKUMURA, 2019). Ainda mais, com intuito de apresentar algumas das tecnologias utilizadas no *Big Data* em conjunto com *machine learning*:

TensorFlow - é uma plataforma completa de código aberto para criação e implementação de aplicativos e algoritmos de *machine learning* (aprendizado de máquina). Ela permite que desenvolvedores e pesquisadores consigam criar expressões computacionais em sistemas heterogêneos, como dispositivos móveis, telefones e tablets, sistemas distribuídos em larga escala e inúmeros equipamentos computacionais como placas gráficas GPU (*Graphics Processing Units*). O objetivo dessa ferramenta é oferecer a oportunidade de acesso à tecnologias de treinamento e inferência para modelos de redes neurais profundas. Os serviços presentes nesta plataforma consistem em reconhecimento de fala, visão computacional, robótica e processamento de linguagem natural (ABAD, 2016).

Python - é uma linguagem de programação de alto nível desenvolvida sob licença de código aberto aprovada pela OSI. Dessa forma ele é totalmente distribuível até mesmo para uso comercial. Esta linguagem é fortemente recomendada para iniciantes e também é muito utilizada por desenvolvedores já experientes por fornecer bibliotecas para desenvolvimento de software, jogos, algoritmos de inteligência artificial, aprendizado de máquina, aprendizado profundo, algoritmos científicos e programação em redes de computadores (PYTHON SOFTWARE FOUNDATION, 2021).

Labelimg - é uma ferramenta gráfica de anotação de imagens. Com ele é possível criar as anotações nas imagens que serão usadas para treinamento e teste no algoritmo de detecção de objetos (LIN TZUTA, 2017).

Github - é uma plataforma de serviço baseado em nuvem capaz de hospedar um sistema de controle de versão conhecido como Git. O Git possui código fonte aberto para colaboração e foi desenvolvido para atuar em pequenos e grandes projetos (GIT, 2021). Ao utilizar a plataforma os desenvolvedores podem colaborar, realizar alterações em projetos compartilhados e manter registros detalhados dos seus progressos (GITHUB Inc, 2021).

Faster R-CNN - é um modelo pré-treinado com conjuntos de dados COCO 2017 (HUANG, 2017). Este modelo foi reutilizado e treinado novamente, mas com o conjunto de dados com imagens de cafeeiros. Esse modelo é disponibilizado gratuitamente no repositório de detecção de objetos do próprio TensorFlow (*TensorFlow Object Detection API*). Essa lista possibilita o acesso a diversos tipos de modelos com velocidade e precisão distintas, assim, foi possível escolher o modelo

que mais se adequa ao projeto e especificações da máquina. A velocidade é apresentada em milissegundos e a precisão em mAP, valor que se refere aos resultados obtidos através do uso do modelo no conjunto de dados COCO 2017 (TENSORFLOW MODEL ZOO, 2021).

OpenCV (*Open Source Computer Vision Library*) - é uma biblioteca de *software* dedicada a visão computacional e aprendizado de máquina. OpenCV oferece uma infraestrutura comum para aplicações que trabalham com visão computacional, e foi construído com licença de código aberto. Dessa forma facilitando a sua utilização por empresas que desejam criar novas soluções e alterar seu código. Seus mais de 2.500 algoritmos otimizados podem executar tarefas de detecção de objetos, classificação de ações ou reações humanas, reconhecimento facial, rastrear objetos em movimento, entre outros (OPENCV TEAM, 2021).

Para obtenção de dados como os quais foram apresentados, é fundamental que se tenha excelentes aparelhos e métodos de monitoramento agrícola. Um método utilizado pela agricultura de precisão capaz de potencializar a construção de mapas de produtividade é a captura de imagens por veículos aéreos não tripulados (VANT) ou satélites. Entretanto, a falta de mão de obra capacitada pode dificultar a realização de estimativas mais precisas para interpretação de dados (AGRICULTURA, 2018).

Por esse motivo, diferente da primeira abordagem que necessita de maior investimento financeiro em equipamentos, este projeto objetiva reconhecer e contar os cafeeiros em talhões, através de vídeos gravados com câmeras embarcadas em smartphones. Dessa maneira, esta ferramenta permite a realização de análises agrícolas na previsão de produção em área pré determinada.

## **2. MATERIAIS E MÉTODOS**

A fim de construir um modelo de detecção de objeto para reconhecer e contar de forma cumulativa os cafeeiros de um cultivo, foi elaborado um plano de execução dividido em seis fases distintas, conforme apresenta a Tabela 1.

**Tabela 1 - Plano de execução.**

---

**Atividades realizadas**

---

---

1 - Captura e preparação dos vídeos e imagens

---

2 - Escolha e configuração do modelo de detecção de objeto

---

3 - Treinamento e exportação do modelo

---

4 - Teste da eficiência do modelo em imagens

---

5 - Implementação do modelo em vídeo

---

6 - Implementação da contagem cumulativa

---

Fonte: Autor

### **2.1 Captura e preparação dos vídeos e imagens:**

A primeira fase consistiu na captura de vídeos e imagens que viriam a ser utilizadas para o treinamento e testes do modelo. Para o processo de captura foi usado a câmera embarcada em smartphone da marca Xiami e modelo Redmi note 9s, com câmera de 48 megapixels, em qualidade 4k e taxa de captura de 30 quadros por segundo. As imagens foram capturadas na lavoura de café do campo experimental da Fatec Pompéia - Shunji Nishimura, localizada na cidade de Pompéia-SP, na data de 29 de setembro de 2021.

A preparação das imagens foi feita utilizando a ferramenta, Labelimg 1.8.0 (LIN TZUTA, 2017), para anotação das imagens para rotular e criação das caixas delimitadoras de objetos em cada cafeeiro existente nas imagens, resultando em um arquivo .XML (*eXtensible Markup Language*) para cada imagem com as especificações da posição de cada objeto rotulado na imagem. Ao final dessa etapa, os arquivos .XML foram convertidos em dois arquivos em TFRecord, formato para armazenar uma sequência de registros binários, um para as imagens de treino e outro para as de teste, utilizando scripts em linguagem Python 3.9.5 (PYTHON SOFTWARE FOUNDATION, 2021).

### **2.2 Escolha e configuração do modelo de detecção de objeto:**

Na segunda fase foi realizada a instalação da API de detecção de objetos do TensorFlow 2.6.0 (ABADI et al, 2016), denominada TensorFlow *Object Detection* API (HUANG et al, 2017) que se encontra disponível em repositório Github. Foi escolhido o modelo de detecção de objeto Faster R-CNN ResNet50 V1 1024x1024 (HUANG et al, 2017) para a fase de treinamento, devido aos seus valores de velocidade e

precisão serem os melhores para os hardwares utilizados. Finalmente, foi feita a configuração do modelo, sendo especificado o poder computacional que seria utilizado durante o treinamento e o caminho para os arquivos TFRecord gerados durante a etapa anterior.

### **2.3 Treinamento e exportação do modelo:**

Na terceira fase foi feita a execução do script da API responsável pelo treinamento do modelo, que consome o poder computacional da máquina e os arquivos TFRecord gerados na fase anterior, resultando em processo de aprendizagem para detecção de cafeeiros. Após a finalização do treinamento, foi executado um *script* da API responsável pela exportação do modelo, armazenando os valores e parâmetros em forma de gráfico, apelidado pelo TensorFlow como gráfico de inferência.

### **2.4 Teste da eficiência do modelo em imagens:**

Na quarta fase, foi realizado o teste do modelo treinado a partir do gráfico de inferência nas imagens não utilizadas durante o processo de treinamento. A eficiência do modelo foi avaliada pela quantidade de cafeeiros detectadas corretamente no vídeo. Nesta etapa foi usado o TensorBoard 2.6.0 (ABADI et al, 2016), um kit de ferramentas de visualização do TensorFlow que torna possível ver os resultados em forma de gráfico e é executado em alguns arquivos gerados durante o processo de treinamento do modelo, chamados de arquivos de evento.

### **2.5 Implementação do modelo em vídeo:**

Na quinta fase foi feita a implementação do gráfico de inferência do modelo treinado em vídeo, através do uso da biblioteca OpenCV 4.5.3.56 (OPENCV TEAM, 2021) que foi responsável por coletar cada quadro do vídeo para que pudessem ser processados pelo modelo desenvolvido em TensorFlow.

### **2.6 Implementação da contagem cumulativa:**

Por fim, na sexta fase foi realizada a implementação da contagem cumulativa dos objetos detectados. Foi utilizado no primeiro passo as funções do OpenCV para criar as alterações visuais no vídeo, como a linha no centro e o número da contagem. Depois, foi usado uma API como base para a lógica da contagem,

chamada TensorFlow *Object Counting* API (OZLU, 2018), que realiza a contagem cumulativa de objetos detectados quando ultrapassam a linha vertical e centralizada na tela utilizando a linguagem Python, e as bibliotecas TensorFlow, OpenCV e Keras.

### 3. RESULTADOS E DISCUSSÕES

Visando promover mais um método de análise para a área de cultivo de café, foi desenvolvido um modelo de aprendizado de máquina que realiza o reconhecimento visual e contagem dos cafeeiros para estimação da quantidade de produção final. Conforme mostrado na Figura 1, que representa o uso do modelo em um vídeo que se movimenta da direita para a esquerda, é possível observar que os resultados obtidos com o modelo finalizado atingiram os objetivos iniciais propostos.

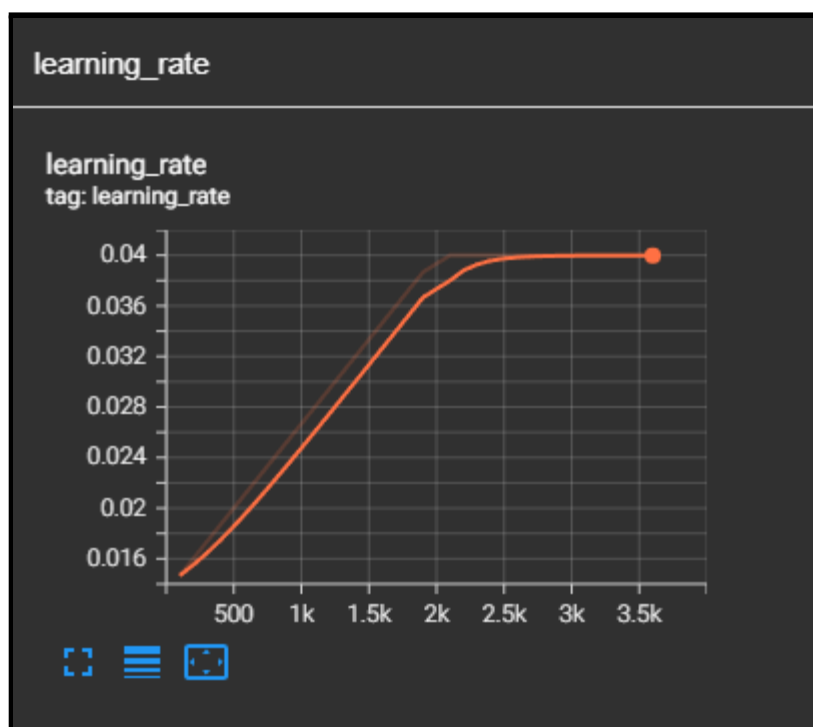
**Figura 1 - Resultados referentes ao modelo finalizado.**



Fonte: Autor

Inspirado nas duas principais APIs de detecção de objetos TensorFlow *Object Detection* API (HUANG et al, 2017) e TensorFlow *Object Counting* API (OZLU, 2018), o modelo ajustado atingiu o objetivo e de acordo com os relatórios obtidos apresenta constante evolução em seu processo de aprendizagem, mas que chega ao seu limite no fim das imagens utilizadas para o treinamento como mostrado na Figura 2, um relatório gerado através do uso do TensorBoard que apresenta a *learning rate*, ou seja, a taxa de aprendizado do modelo, onde o eixo X equivale ao tempo de treino decorrido e o eixo Y ao valor da taxa de aprendizado utilizado.

**Figura 2 - Taxa de aprendizado.**



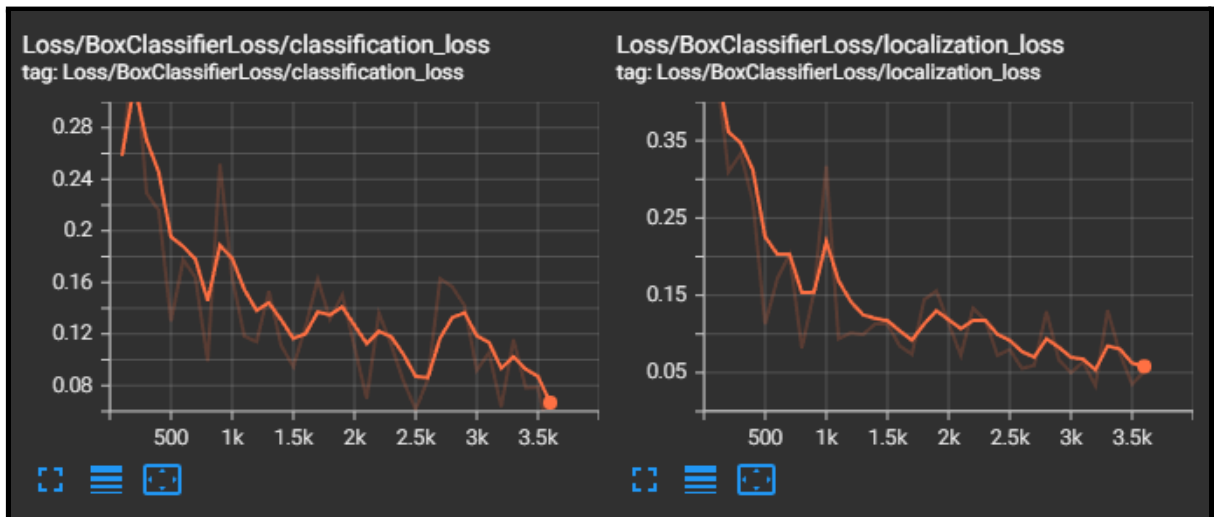
Fonte: Autor

A utilização do TensorBoard também nos possibilita a visualização de mais um tipo de relatório muito importante para medir a eficiência do modelo criado: os relatórios de perda. Eles são relatórios criados a partir da utilização de funções internas do TensorBoard chamadas de funções de perda, que são as responsáveis por representar o "preço pago pela imprecisão das previsões", portanto, quanto menor os valores de perda maior é a precisão do modelo. Como apresentado na Figura 3, o modelo criado teve uma perceptível queda nos valores de perda na classificação e localização dos cafeeiros, como mostrado a partir dos gráficos de



*classification loss* e *localization loss* que assim como os outros gráficos gerados pelo TensorBoard o eixo X é referente ao tempo de treino e o eixo Y aos valores de perda.

**Figura 3 - Relatórios referentes às taxas de perda de classificação e localização.**



Fonte: Autor

Porém, nos modelos de detecção de objeto, mesmo sendo criado com o principal objetivo a detecção em vídeo e ter atingido os resultados esperados, como apresentado pela Figura 4, sua eficiência ainda é maior em imagem. Isso acontece por conta das taxas de quadros por segundo que exige mais poder de processamento do modelo e gera algumas perdas não comumente ocorridas em imagens, como alguns quadros com ângulos não favoráveis para a detecção do objeto.

**Figura 4 - Teste do modelo em imagem.**



Fonte: Autor

#### **4. CONCLUSÕES**

Foi possível criar um modelo de detecção de objeto que atendesse a proposta inicial do projeto de reconhecer e contar os cafeeiros dos talhões presentes em um cultivo de café. E, de acordo com os resultados alcançados, o desenvolvimento do modelo pode ser mais um método de análise dentro do campo da agricultura de precisão, gerando dados valiosos principalmente para a estimativa da quantidade de produção.

#### **AGRADECIMENTOS**

Ao nosso orientador Professor Doutor João Ricardo Favan nossos sinceros agradecimentos por todos os conselhos e direcionamentos oferecidos durante nossas reuniões, sempre atencioso e paciente às nossas dúvidas e dificuldades, assim como todos os outros docentes que contribuíram para nosso aprendizado e formação. Somos gratos a todos os funcionários envolvidos em manter e conservar o campus da Faculdade de Tecnologia “Shunji Nishimura”, suas salas, laboratórios sempre em ótimas condições. Também agradecemos especialmente a nossa família, ela exerceu um papel fundamental para sustentar nosso caminho de aprendizado até o final.

## REFERÊNCIAS

ABADI, Martin et al. TensorFlow: A System for Large-Scale Machine Learning. **Computer Science**, p. 1-21, mai. 2016. Disponível em: <<https://arxiv.org/abs/1605.08695>>. Acesso em: 25 out. 2021.

AGRICULTURA digital: Quais as dificuldades para a implementação no agronegócio. **Vdibrasil**. 2018. Disponível em: <<https://www.vdibrasil.com/agricultura-digital-quais-as-dificuldades-para-a-implementacao-no-agronegocio/>>. Acesso em: 02 nov. 2021

CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira de café**, Brasília, DF, v.8 safra 2020/21, n. 3, set. 2021.

EXPORTAÇÃO. **Cecafé**. 04 set. 2017. Disponível em: <<https://www.cecafe.com.br/sobre-o-cafe/exportacao/>>. Acesso em: 02 nov. 2021.

GITHUB Inc. **GITHUB**. Disponível em: <<https://github.com>>. Acesso em: 28 out. 2021.

HAMANO, Junio. **GIT**. Versão 2.31.0. 15 mar. 2021. Sistema de controle de versões distribuído. Disponível em: <<https://git-scm.com/>>. Acesso em: 28 out. 2021.

HUANG, Jonathan et al. Speed/accuracy trade-offs for modern convolutional object detectors. **Computer Science**, p. 1-21, abr. 2017. Disponível em: <<https://arxiv.org/abs/1611.10012>>. Acesso em: 28 out. 2021.

LIN TZUTA. **LABELIMG**. Versão 1.8.0. 03 dez. 2018. Ferramenta gráfica de anotação de imagem. Disponível em: <<https://github.com/tzulin/labelimg>>. Acesso em: 28 out. 2021.

OKUMURA, Roberto. Machine learning na agricultura. **Venturus**,2019. Disponível em: <<https://www.venturus.org.br/machine-learning-na-agricultura/>>. Acesso em: 02 nov. 2021.

OPENCV TEAM. **OPENCV**. Versão 4.5.3. [Palo Alto]. 5 jul. 2021. Biblioteca de código para visão computacional. Disponível em: <<https://docs.opencv.org/4.5.3/>>. Acesso em: 28 out. 2021.

OZLU, Ahmet. **TENSORFLOW OBJECT COUNTING API**. 2018. Disponível em: <[https://github.com/ahmetozlu/tensorflow\\_object\\_counting\\_api](https://github.com/ahmetozlu/tensorflow_object_counting_api)>. Acesso em: 28 out. 2021.

PEMBERTON, Chris. Big Data Basics for Digital Marketers. **GARTNER**. 2015. Disponível em: <<https://www.gartner.com/en/marketing/insights/articles/big-data-basics-for-digital-marketers>>. Acesso em: 03 nov. 2021.

PYTHON SOFTWARE FOUNDATION. **PYTHON**. Versão 3.9.5. 03 mai. 2021. *Python Language Site*: Documentation. Disponível em:  
<<https://docs.python.org/release/3.9.5/>>. Acesso em: 28 out. 2021.