

**PROGRAMA DE PÓS-GRADUAÇÃO EM PRODUÇÃO VEGETAL**

**Ana Clara Gonçalves Fernandes**

**FENOTIPAGEM DE ALTA EFICIÊNCIA NO MELHORAMENTO GENÉTICO DA BATATA-DOCE POR  
ANÁLISE COMPUTACIONAL DE IMAGENS**

**Montes Claros**

**2022**

**Ana Clara Gonçalves Fernandes**

**Fenotipagem De Alta Eficiência No Melhoramento Genético Da Batata-Doce Por Análise Computacional De Imagens**

Tese apresentada ao Programa de Pós-Graduação em Produção Vegetal da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Doutor em Produção Vegetal.

**Orientador:** Alcinei Místico Azevedo

Montes Claros  
Outubro de 2022

Fernandes, Ana Clara Gonçalves.

F363f  
2022

Fenotipagem de alta eficiência no melhoramento genético da batata-doce por análise computacional de imagens [manuscrito]/ Ana Clara Gonçalves Fernandes. Montes Claros, 2022.  
67f. : il

Tese (doutorado) - Área de concentração em Produção Vegetal. Universidade Federal de Minas Gerais / Instituto de Ciências Agrárias.

Orientador: Alcinei Místico Azevedo

Banca examinadora: Valter Carvalho de Andrade Júnior, Nermy Ribeiro Valadares, Tiago Olivoto, Samy Pimenta.

Inclui referências: f. 25-29; f. 39-41; f. 52-54; f. 64-66.

1. Batata doce – Teses. 2. Inteligência Computacional – Teses. 3. Processamento de imagens – Teses. 4. Genética vegetal – Teses. 5. Hortaliças – Melhoramento genético – Teses. I. Azevedo, Alcinei Místico. II. Universidade Federal de Minas Gerais. Instituto de Ciências Agrárias. III. Título.

CDU: 635.1/8

ELABORADA PELA BIBLIOTECA UNIVERSITÁRIA DO ICA/UFMG  
Nádia Cristina Oliveira Pires / CRB-6/2781

ATA DE DEFESA DE TESE

Aos 28 dias do mês de outubro de 2022, às 8:00 horas, sob a Presidência do Professor Alcinei Místico Azevedo, D. Sc. (Orientador UFMG/ICA) e com a participação dos Professores Valter Carvalho de Andrade Júnior, D. Sc. (UFLA), Nermy Ribeiro Valadares, D. Sc. (IFNMG – Campus Araçuaí), Tiago Olivoto, D. Sc. (UFSC) e Samy Pimenta, D. Sc. (Unimontes – Campus Janaúba), reuniu-se, por videoconferência, a Banca de tese de **Ana Clara Gonçalves Fernandes**, aluna do Curso de Doutorado em Produção Vegetal. O resultado da defesa de tese intitulada: "FENOTIPAGEM DE ALTA EFICIÊNCIA NO MELHORAMENTO GENÉTICO DA BATATA-DOCE POR ANÁLISE COMPUTACIONAL DE IMAGENS" sendo a aluna considerada (aprovada/reprovada) **Aprovada**. E, para constar, eu, Professor Alcinei Místico Azevedo, Presidente da Banca, lavrei a presente ata que depois de lida e aprovada, será assinada por mim e pelos demais membros da Banca examinadora.

OBS.: A aluna somente receberá o título após cumprir as exigências do **ARTIGO 65** do regulamento do Curso de Doutorado em Produção Vegetal, conforme apresentado a seguir:

**Art. 65** Para dar andamento ao processo de efetivação do grau obtido, o candidato deverá, após a aprovação de sua Dissertação ou Tese e da realização das modificações propostas pela banca examinadora, se houver, encaminhar à secretaria do Colegiado do Programa, com a anuência do orientador, 1 (um) exemplar impresso e 1 (um) exemplar eletrônico da dissertação ou tese, no prazo de 60 (sessenta) dias.

Montes Claros, 28 de outubro de 2022.

Valter Carvalho de Andrade Júnior

Membro



Documento assinado digitalmente

TIAGO OLIVOTO

Data: 02/11/2022 12:03:49-0300

CPF: \*\*\* 889.590-\*\*

Verifique as assinaturas em <https://v.ufsc.br>

Tiago Olivoto

Membro

*Nermy Ribeiro Valadares*

Nermy Ribeiro Valadares

Membro

*Samy Pimenta*

Samy Pimenta

Membro

*Alcinei Místico Azevedo*

Alcinei Místico Azevedo

Orientador

Dedico este trabalho aos meus pais Ione e Helio e ao meu irmao Helio Filho, pelo apoio e dedicaao em todos os momentos da minha vida.

## AGRADECIMENTOS

A Deus pela sabedoria e discernimento para conseguir contornar as dificuldades e alcançar meus objetivos.

Aos meus pais Ione e Hélio que sempre me apoiaram, incentivaram e nunca pouparam esforços para que eu conseguisse alcançar meus ideais, ao depositar sua confiança em mim e fazendo das minhas vitórias as suas.

Aos meus irmãos Rafael e Hélio, em especial, a Hélio pelo companheirismo, amizade, união e apoio incondicional.

Ao Professor Alcinei pela orientação, confiança, paciência e ensinamentos compartilhados que me possibilitaram chegar ao final desta etapa.

Aos colegas do Grupo de Estudo em Experimentação Agrícola (GEEA), em especial, à Nermy e ao Clóvis Henrique pela amizade, companheirismo e auxílio na condução das pesquisas.

A todos os meus amigos e colegas, que de alguma forma não me deixaram desistir e me ajudaram e incentivaram nesses anos.

Aos professores do Programa de Pós-Graduação pela contribuição e ensinamentos nas disciplinas cursadas.

À banca pela disponibilidade.

À Universidade Federal de Minas Gerais, em especial, ao Instituto de Ciências Agrárias pela oportunidade da graduação em Engenharia Agrícola e Ambiental e ao Programa de Pós-Graduação em Produção Vegetal, pela oportunidade de realização do mestrado e doutorado.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior- Brasil (CAPES), ao Conselho Nacional de Desenvolvimento Científico (CNPQ) e à Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) pelo apoio a este estudo.

Meu muito obrigada!

## FENOTIPAGEM DE ALTA EFICIÊNCIA NO MELHORAMENTO GENÉTICO DA BATATA-DOCE POR ANÁLISE COMPUTACIONAL DE IMAGENS

### RESUMO

A batata-doce (*Ipomoea batatas* (L.) Lam..) destaca-se entre as hortaliças mais plantadas no Brasil. Esforços para o melhoramento genético da cultura são necessários, visando ao aumento de produtividade e qualidade das raízes. Porém, na seleção dos melhores genótipos, é necessário avaliar um grande número de características, sendo um processo oneroso e subjetivo, dificultando a análise do melhorista. Nesse sentido, a adoção de novas tecnologias ao processo de fenotipagem representa um avanço e, entre as possibilidades, tem-se a análise de imagens associada à inteligência computacional. Sendo assim, possibilita a avaliação de características de interesse, em menor espaço de tempo, reduzindo a mão de obra no melhoramento genético da cultura. Além disso, possibilita a classificação de caracteres qualitativos, reduzindo a subjetividade que existe nas notas dadas pelos avaliadores. Diante do exposto, objetiva-se aplicar metodologias, para a automação computacional da avaliação da produção e qualidade de raízes de batata-doce, por meio da análise de imagens digitais. Foram avaliadas 16 progênies de meios-irmãos de batata-doce, em delineamento em blocos casualizados, com quatro repetições e dez plantas por parcela. Para a aquisição das imagens, foi retirado o excesso de solo das raízes, e as imagens foram adquiridas por câmera digital e pré-processadas no software R. Foram avaliadas, em nível de planta, as seguintes características de produção e qualidade: peso individual de raízes, formato das raízes, danos causados por insetos, coloração da casca e da polpa. A utilização da análise de imagens associada à visão computacional, bem como ao aprendizado profundo, é uma ferramenta eficiente em programas de melhoramento genético da batata-doce, auxiliando na fenotipagem da cultura, como também na tomada de decisões. Além disso, podem ser utilizadas no desenvolvimento de aplicativos e aparelhos que auxiliam o produtor na classificação de raízes de batata-doce. Isso possibilita a utilização da hortaliça aos seus mais diversos fins, reduzindo perdas e desperdício no processo produtivo da cultura.

**Palavras-Chave:** *Ipomoea batatas* (L.) Lam.. Inteligência Computacional. Melhoramento genético de hortaliças. Análise de imagens. Fenotipagem de alto rendimento.

## HIGH-EFFICIENCY PHENOTYPING IN THE GENETIC IMPROVEMENT OF SWEET POTATO BY COMPUTATIONAL IMAGE ANALYSIS

### ABSTRACT

Sweet potato (*Ipomoea batatas* (L.) Lam.) stands out among the most planted vegetables in Brazil. Efforts for the genetic improvement of the crop are necessary, aiming to increase the productivity and quality of the roots. However, in selecting the best genotypes, it is necessary to evaluate a large number of characteristics, being an expensive and subjective process, making it difficult for the breeder to analyze it. In this sense, the adoption of new technologies to the phenotyping process represents a breakthrough and, among the possibilities, there is the image analysis associated with computational intelligence. Thus, it enables the evaluation of characteristics of interest, in a shorter period of time, reducing labor in the genetic improvement of the crop. In addition, it enables the classification of qualitative characters, reducing the subjectivity that exists in the grades given by the evaluators. Given the above, the objective is to apply methodologies for the computational automation of the evaluation of production and quality of sweet potato roots, through the analysis of digital images. They were evaluated sixteen half-sib progenies of sweet potato in a randomized block design, with four replications and ten plants per plot. For image acquisition, it was removed excess soil from the roots, and the images were acquired by digital camera and pre-processed in R software. They were evaluated at the plant level the following production and quality characteristics: individual root weight, root shape, damage caused by insects, coloring of peel and pulp. The use of image analysis associated with computer vision, as well as deep learning, is an efficient tool in sweet potato genetic improvement programs, helping in the phenotyping of the crop, as well as in decision making. In addition, they can be used in the development of applications and devices that help the producer in the classification of sweet potato roots. This makes it possible to use the vegetable for its most diverse purposes, reducing losses and waste in the production process of the crop.

**Keywords:** *Ipomoea batatas* (L.) Lam.. Computational Intelligence. Genetic improvement of vegetables. Image analysis. High-throughput phenotyping.

## LISTA DE ILUSTRAÇÕES

### REFERENCIAL TEÓRICO

Figura 1- <b>(A)</b> Esteira utilizada para manejos pós-colheita na cultura da batata-doce; <b>(B)</b> possíveis classificações para o formato de raízes obtidas por análise computacional de imagem.....	19
Figura 2- Rede Neural Artificial.....	21
Figura 3- Imagem RGB (Red – Green – Blue) que foi separada por seus três planos coloridos – Vermelho, Verde e Azul.....	22
Figura 4- Exemplo de CNN para classificação de espécies de animais, com camadas de convolução, max pooling e uma camada totalmente conectada. Sua saída é a classificação da imagem de entrada em alguma classe.....	23

### CAPÍTULO I: Viabilidade da visão computacional no melhoramento genético da produção de raízes de batata-doce

Figura 1- Aquisição de imagens: A) “estúdio” para a aquisição das imagens; B) Iluminação artificial com lâmpada fluorescente; C) Imagem adquirida no estúdio.....	32
Figura 2- Fluxograma das etapas realizadas a partir da obtenção das imagens.....	33
Figura 3- Regressão ajustada na predição do peso da raiz (g), a partir do número de pixels em imagens de 2233 raízes, oriundas de 16 progênes de meios-irmãos de batata-doce.....	34
Figura 4- Dispersão gráfica entre os valores de peso total preditos e observados em nível de parcela (A) e de família (B) em 16 progênes de meios-irmãos de batata-doce.....	35
Figura 5- Dispersão em gráfico radar e correlação de Spearman (r) dos valores genéticos aditivos de 16 progênes de meios-irmãos de batata-doce.....	36

### CAPÍTULO II: Melhoramento genético da cor da polpa de batata-doce por técnicas de visão computacional e random Forest

Figura 1- Aquisição de imagens: A) “estúdio” para a aquisição das imagens; B) Iluminação artificial com lâmpada fluorescente; C) Imagem adquirida no estúdio.....	44
Figura 2- Paletas para a classificação das polpas de batata-doce quanto à cor.....	45
Figura 3- Fluxograma da metodologia de avaliação da coloração de polpas de raízes de batata-doce.....	46

Figura 4- Demonstração da classificação de pixels quanto à coloração em polpas de batata-doce.....	48
Figura 5- Dendrograma obtido pelo método UPGMA para progênes de meios-irmãos de batata-doce em função da proporção de pixels com cada uma das seis classificações de cor.....	49
Figura 6- Dispersão das duas primeiras variáveis canônicas no estudo da coloração de raízes de batata-doce.....	50

### **CAPÍTULO III: Redes Neurais convolucionais no melhoramento qualitativo de raízes de batata-doce**

Figura 1- Aquisição de imagens: A) “estúdio” para a aquisição das imagens; B) Iluminação artificial com lâmpada fluorescente; C) Imagem adquirida no estúdio.....	58
Figura 2- Etapas para a individualização de uma raiz por imagem na fenotipagem de raízes de batata-doce por análise computacional de imagens.....	58
Figura 3- Classificação de raízes selecionadas aleatoriamente nas amostras pertencentes ao conjunto de imagens de teste por RCNN InceptionResNetV2 na classificação de raízes de batata-doce quanto ao formato, danos ocasionados por insetos e cor da casca.....	62

## LISTA DE TABELAS

### REFERENCIAL TEÓRICO

Tabela 1 - Características de cada arquitetura.....	24
---	----

### **CAPÍTULO I: Viabilidade da visão computacional no melhoramento genético da produção de raízes de batata-doce**

Tabela 1- Estimativas de parâmetros genéticos no estudo do peso observado e predito em 16 progênies de meios-irmãos de batata-doce.....	35
Tabela 2- Ganho de seleção direto (negrito) e indireto em progênies de meios-irmãos de batata-doce.....	37

### **CAPÍTULO II: Viabilidade da visão computacional no melhoramento genético da produção de raízes de batata-doce**

Tabela 1- Tabela de confusão para a classificação de pixels correspondentes a diferentes tons na paleta de cores.....	47
Tabela 2- Avaliadores da classificação do ajuste da metodologia de florestas randômicas na classificação de pixels em seis colorações de polpas de batata-doce.....	47

### **CAPÍTULO III: Redes Neurais convolucionais no melhoramento qualitativo de raízes de batata-doce**

Tabela 1- Escala de notas relacionada ao formato, cor predominante da casca e danos causados por insetos para progênies de meios-irmãos de batata-doce ( <i>Ipomoea batatas</i> (L.) Lam).....	57
Tabela 2- Número de imagens de raízes de batata-doce utilizadas para a fenotipagem da cor, formato e danos por insetos (originais) e número de imagens destinadas ao ajuste das RCNNs (Ajuste) e avaliação da qualidade do ajuste (Teste).....	59
Tabela 3- Número de épocas, tempo de treinamento e classificações realizadas por diferentes arquiteturas de RCNNs para raízes de batata-doce quanto ao formato, danos ocasionados por insetos e cor da casca.....	60
Tabela 4- Avaliadores da qualidade do ajuste de RCNNs com diferentes arquiteturas na classificação de raízes de batata-doce quanto ao formato, danos ocasionados por insetos e cor da casca.....	61

## SUMÁRIO

1	INTRODUÇÃO GERAL .....	13
2	OBJETIVOS.....	15
2.1	Objetivo Geral.....	15
2.2	Objetivos Específicos .....	15
3	REVISÃO DE LITERATURA .....	16
3.1	Cultura da batata-doce .....	16
3.2	Melhoramento Genético de Batata-Doce .....	17
3.3	Fenotipagem de alta eficiência.....	18
3.4	Análise computacional de imagens .....	20
3.5	Redes Neurais Convolucionais na análise de imagens .....	21
3.5.1	Redes Neurais Artificiais .....	21
3.5.2	Redes Neurais Artificiais Convolucionais.....	22
3.6	REFERÊNCIAS .....	25
4	ARTIGOS.....	30
4.1	CAPÍTULO I:Visão computacional no melhoramento genético para a produção de raízes de batata-doce .....	30
	Resumo .....	30
4.1.1	Introdução.....	30
4.1.2	Material e Métodos .....	31
4.1.2.1	Instalação e avaliação do experimento .....	31
4.1.2.2	Aquisição e processamento de imagens .....	32
4.1.2.3	Predição do peso das raízes e estimação de parâmetros genéticos e ganhos de seleção .....	33
4.1.3	Resultados.....	34
4.1.4	Discussão .....	37
4.1.5	Conclusões .....	38
4.1.6	Agradecimentos.....	38
4.1.7	Referências .....	39
4.2	CAPÍTULO II: Diversidade genética da cor da polpa de batata-doce por visão computacional e random Forest.....	42
	Resumo: .....	42
4.2.1	Introdução.....	42
4.2.2	Material e Métodos .....	43
4.2.2.1	Instalação e avaliação do experimento .....	43
4.2.2.2	Aquisição e processamento de imagens .....	44
4.2.3	Resultados.....	46
4.2.4	Discussão .....	50
4.2.5	Conclusões.....	51
4.2.6	Agradecimentos.....	52
4.2.7	REFERÊNCIAS .....	52

4.3 CAPÍTULO III: Redes neurais convolucionais no melhoramento qualitativo de raízes de batata-doce .....	55
4.3.1 Introdução.....	55
4.3.2 Material e Métodos .....	56
4.3.2.1 Instalação e avaliação do experimento .....	56
4.3.2.2 Aquisição e processamento de imagens .....	57
4.3.3 Resultados.....	60
4.3.4 Discussão .....	62
4.3.5 Conclusões.....	64
4.3.6 Agradecimentos.....	64
4.3.7 Referências .....	64
5 CONSIDERAÇÕES FINAIS .....	67

## 1 INTRODUÇÃO GERAL

O uso de variedades obtidas por meio do melhoramento genético é considerado uma das mais importantes estratégias na agricultura moderna. As variedades melhoradas são mais produtivas, apresentam resistência a estresses bióticos e abióticos, maior qualidade nutracêutica e, conseqüentemente, menor impacto ambiental e custo de produção (FANG; XIONG, 2015; RINCKER *et al.*, 2014; HUANG *et al.*, 2016). O melhoramento genético teve início com a seleção fenotípica de indivíduos que se destacavam positivamente durante a domesticação pelo homem (SOUSA *et al.*, 2019). Busca obter indivíduos com características genéticas superiores aos genitores, almejando materiais de ampla variabilidade. Para isso, os melhoristas devem ter amplo conhecimento acerca da cultura que está sendo melhorada (FERREIRA; GRATTAPAGLIA, 1998; SOUSA *et al.*, 2007). Nesse sentido, os programas de melhoramento de diversas culturas, como a batata-doce, são de suma importância para atender a demanda alimentar da humanidade.

A batata-doce (*Ipomoea batatas* (L.) LAM) é uma hortaliça pertencente à família Convolvulaceae, de raízes tuberosas, originária da América Tropical. Essa hortaliça se caracteriza por possuir indivíduos hexaploides ( $2n=6x=90$ ), fato responsável pela alta variabilidade genética observada nessa espécie (RITSCHER *et al.*, 2010). Essa característica proporciona à cultura uma grande quantidade de genótipos, que se tornam foco de estudos para a seleção de novas cultivares melhoradas geneticamente (AZEVEDO *et al.*, 2015a).

A batata-doce está entre as olerícolas mais consumidas no Brasil por ser uma rica fonte de carboidratos, fibras, vitaminas, minerais, antioxidantes e pelo baixo conteúdo glicêmico (RAHAJENG *et al.*, 2020). Além dos múltiplos usos, a batata-doce é uma hortaliça de papel importante no abastecimento de matérias-primas para a alimentação humana e animal (WERA *et al.*, 2014; PRADHAN *et al.*, 2015). Com isso, a crescente demanda por bioprodutos da cultura aguça o interesse na diversificação agrícola e estimula a pesquisa de matérias primas para a produção de amidos, açúcares e outros produtos de origem vegetal (HENRY, 2012). Assim, visando ao aumento de produtividade e qualidade do produto, faz-se necessário investir em informações e novas tecnologias associadas ao melhoramento genético da hortaliça (RAHAJENG *et al.*, 2020).

Contudo o melhoramento genético da batata-doce torna-se um trabalho oneroso, em razão, principalmente, do grande número de características analisadas, uso intensivo de mão de obra, tempo, além de vários genótipos possuírem características semelhantes. A partir disso, a adoção de novas tecnologias relacionadas ao processo de fenotipagem de plantas torna-se uma vertente, em que, entre elas, pode-se destacar o uso de imagens associadas à inteligência computacional. Nesse sentido, a análise de imagens se tornou um componente-chave, para preencher a lacuna do genótipo-fenótipo, em que a medição e contagem automatizada permitiram maior rendimento, desenvolvimento de sistemas consistentes e imparciais, além de caracterizar a fenotipagem não destrutiva da planta (WALTER; LIEBISCH; HUND, 2015). Assim, um requisito fundamental, para ferramentas de fenotipagem baseadas em imagens, é transformar automaticamente essas imagens em medições fenotípicas confiáveis e precisas (UBBENS, STAVNESS, 2017).

Fenótipos gerados pela associação de imagens com programas de computador permitem o futuro genético de várias espécies, uma vez que podem ser gerados bancos de dados com as imagens, para serem fenotipados e consultados, ao longo dos anos, de acordo com a conveniência do pesquisador. Já

para as medidas manuais, a fenotipagem deve ser realizada, em um espaço de tempo pré-definido, ou seja, enquanto as plantas estão vivas. Além disso, essas novas tecnologias que visam aumentar a acurácia e rapidez das mensurações fenotípicas vêm sendo alvo de intensas pesquisas, nos últimos anos (MASSRUHÁ *et al.*,2014), como no desenvolvimento de plataformas de fenotipagem (UBBENS, STAVNESS, 2017), sistemas automatizados de fenotipagem de alta eficiência (LEE *et al.*,2018), bem como na caracterização de fenótipos em batata-doce, usando imagens de alto rendimento (HAQUE *et al.*, 2021).

Observada a necessidade de melhoria à obtenção e análise de dados fenotípicos em batata-doce, principalmente, pela grande variedade genética e fenotípica verificada para a cultura, faz-se necessário estudar as possibilidades de uso de imagens tanto para as características de produção quanto para as qualitativas, a fim de facilitar a fenotipagem da cultura.

## **2 OBJETIVOS**

### **2.1 Objetivo Geral**

Aplicar metodologias para a automação computacional da avaliação da produção e qualidade de raízes de batata-doce por meio da análise de imagens digitais.

### **2.2 Objetivos Específicos**

- Desenvolver uma metodologia para a fenotipagem da produção de raízes, para o melhoramento genético de progênies de meios irmãos de batata-doce, por meio da análise computacional de imagens.
- Verificar a eficiência da análise computacional de imagens na avaliação da coloração de polpas de raízes de batata-doce e sua aplicabilidade no melhoramento genético.
- Estudar a dissimilaridade genética da coloração das polpas das raízes por meio de técnicas estatísticas e de inteligência computacional.
- Verificar o potencial de redes neurais convolucionais na classificação de raízes de batata-doce quanto ao formato, cor da casca e danos ocasionados por insetos.

### 3 REVISÃO DE LITERATURA

#### 3.1 Cultura da batata-doce

A batata-doce (*Ipomoea batatas* (L.) LAM) é uma hortaliça, de raízes tuberosas, pertencente à família Convolvulaceae e originária da América Tropical. Essa cultura se destaca pela sua rusticidade, facilidade de tratos culturais, ampla adaptação às adversidades ambientais e baixo custo de cultivo, já que não necessita de cuidados intensivos (ANDRADE JÚNIOR *et al.*, 2012). Além disso, como o tempo de colheita pode ser diversificado e mecanizado, seu crescimento prostrado e grande massa foliar dá destaque à eficiência dessa espécie no controle de erosão do solo, favorecendo sua produção (SILVA; LOPES; MAGALHÃES, 2008; SOUZA, 2000).

O alto potencial de produção da cultura está intimamente ligado ao aproveitamento total da planta (raízes e ramas), uma vez que possui raízes ricas em açúcares, carboidratos, sais minerais e vitaminas, tornando-se uma alternativa importante na alimentação humana e animal (AZEVEDO *et al.*, 2002), além de ser matéria-prima para indústrias de alimentos, tecidos, papéis, cosméticos, etanol, entre outros (CARDOSO *et al.*, 2005). Contudo a cultura tem sido foco de muitos trabalhos, possibilitando a seleção de genótipos para inúmeros propósitos (AZEVEDO *et al.*, 2015a).

Assim, pode-se inferir que a produtividade de raízes e ramas variam de acordo com o genótipo utilizado e as regiões de cultivo da hortaliça. Andrade Junior *et al.* (2012), ao avaliarem 12 genótipos de batata-doce, em Diamantina-MG, quanto à produção de ramas e raízes, verificaram que o clone BD-45(UFVJM05) foi o que mais se destacou na produção de massa verde (19,7 t ha<sup>-1</sup>), produção de massa seca (3,3t ha<sup>-1</sup>), produtividade total (32,9 t ha<sup>-1</sup>) e comercial de raízes (29,5 t ha<sup>-1</sup>). Essas produtividades de raízes foram inferiores aos obtidos por Santos Neto *et al.* (2017) que, ao avaliarem a adubação nitrogenada, em três clones de batata-doce em Sergipe, obtiveram produtividades totais de raízes de 55,18 e 39,96 t ha<sup>-1</sup>, para os clones IPB-075 e IPB-149, respectivamente. Já para alimentação animal, as ramas de batata-doce apresentaram características satisfatórias, podendo ser fornecidas frescas, secas, na forma de silagem (MONTEIRO *et al.*, 2007; FIGUEIREDO *et al.*, 2012; PEDROSA *et al.*, 2015) e de feno (DONATO *et al.*, 2020).

O interesse no potencial da batata-doce na indústria de processamento está aumentando (TODD *et al.*, 2015), já que a crescente demanda por bioprodutos e o crescente interesse, no desenvolvimento de culturas alternativas, para promover a diversificação agrícola, estimularam a pesquisa de matérias-primas para a produção de amidos, açúcares e outros produtos naturais de origem vegetal (HENRY,2012).

A batata-doce é rica em amido e aproximadamente 80% da matéria seca são carboidratos bioconvertidos anaerobicamente, em hidrogênio e etanol (LAY *et al.*, 2012), sendo possível observar estudos que computaram a produção de um volume entre 17.354 e 20.007 L de etanol por hectare de batata-doce (PAVLAK *et al.*, 2011).

A batata-doce possui um alto valor nutritivo por seu conteúdo de carboidratos, além disso, possui ampla versatilidade sensorial, principalmente em termos de sabor, texturas e cores de polpa (VIZZOTO *et al.*, 2017). O interesse por cultivares de batata-doce de polpa colorida está crescendo em função do seu potencial nos compostos bioativos e nutricionais e dos benefícios à saúde humana (CAMIRE *et al.*, 2009). Dessa forma, pode-se dizer que as diferenças nas cores das polpas estão associadas às quantidades de β-caroteno, antocianinas, compostos fenólicos, fibra dietética, ácido ascórbico, ácido fólico e de sais minerais presentes nas raízes de batata-doce (WOOLFE, 2008).

Nesse sentido, a batata-doce é considerada uma excelente fonte de vitaminas do complexo B, vitamina A e de minerais, como ferro, cálcio, potássio, enxofre e magnésio (LOW *et al.*, 2007). E podem ser utilizadas, como alimento promissor à erradicação de carências nutricionais, em grupos populacionais com deficiência de consumo em calorias ou nutrientes (KEHOE *et al.*, 2015). Além disso, a batata-doce também pode ter potencial como fonte de antocianinas para uso como corante alimentar natural e ingredientes alimentares funcionais pelas suas propriedades antioxidantes e anticancerígenas reconhecidas (PHILPOTT *et al.*, 2004; TEOW *et al.*, 2007).

Apesar da importância da batata-doce como fonte de minerais, são escassos os estudos sobre sua composição no Brasil. Nessa ótica, destaca-se o melhoramento genético para a cultura, permitindo uma melhora da qualidade da cadeia produtiva da batata-doce e possibilitando aumentar a qualidade nutricional da hortaliça, além de agregar valor comercial ao alimento (BOVELL-BENJAMIN, 2007).

### 3.2 Melhoramento Genético de Batata-Doce

O melhoramento genético teve início com a seleção fenotípica de indivíduos que se destacavam positivamente durante a domesticação pelo homem (SOUSA *et al.*, 2019). Busca obter indivíduos com características genéticas superiores aos genitores, almejando materiais de ampla variabilidade. Para isso, os melhoristas devem ter amplo conhecimento acerca da cultura que está sendo melhorada (FERREIRA; GRATTAPAGLIA, 1998; SOUSA *et al.*, 2007). Nessa perspectiva, os programas de melhoramento de diversas culturas, como a batata-doce, são de suma importância para atender a demanda alimentar da humanidade.

A batata-doce pode ser propagada tanto vegetativamente quanto por sementes botânicas, em que, para culturas propagadas vegetativamente, a domesticação é encarada como um processo longo, envolvendo ciclos repetidos de seleção, recombinação e fluxo gênico, conduzindo a transição planta silvestre-domesticada (ROULLIER *et al.*, 2013).

A batata-doce se caracteriza por possuir indivíduos hexaploides ( $2n=6x=90$ ), fato responsável pela alta variabilidade genética observada nessa espécie (RITSCHER *et al.*, 2010), pelo alto nível de ploidia e alogamia. A alogamia da batata-doce, aliada à propagação vegetativa, faz com que o melhoramento da batata-doce seja mais fácil, pois, diante de grande variabilidade genética, a obtenção de uma única planta superior agronomicamente é o suficiente, visto que poderia ser propagada facilmente pelo método vegetativo. Mas, em virtude de seu alto nível de ploidia, os estudos genéticos em batata-doce ficam atrás das principais culturas diploides (MOLLINARI *et al.*, 2020).

Para o melhoramento da batata-doce, visando à alimentação animal, humana e produção de bioprodutos, como etanol, algumas características devem ser observadas. Para a alimentação humana, deseja-se selecionar indivíduos que apresentem maior quantidade e qualidade de raízes. Para alimentação animal, maior produtividade de ramos e, para a produção de etanol, raízes com maior teor de amido.

Entre as principais finalidades, para o melhoramento genético da batata-doce, destacam-se a resistência a pragas e doenças, melhor qualidade nutricional (AZEVEDO *et al.*, 2002; ANDRADE JÚNIOR *et al.*, 2012; AZEVEDO *et al.*, 2014), aumento no teor de matéria seca e biomassa (GONÇALVES NETO *et al.*, 2011; AZEVEDO *et al.*, 2014); desenvolvimento de cultivares, para a produção de etanol (GONÇALVES NETO *et al.*, 2012), maior produção de ramos para a alimentação animal (VIANA *et al.*,

2011; ANDRADE JÚNIOR *et al.*, 2012), bem como a investigação do potencial de cruzamentos de genótipos de batata-doce para a produção de amido e antocianina no Sudeste dos Estados Unidos (TODD *et al.*, 2015).

O melhoramento genético da batata-doce é essencialmente desenvolvido pelo setor público, não sendo objetivo para grandes empresas. Esse fato está associado à facilidade de propagação da espécie e, dessa forma, o agricultor precisaria comprar sementes apenas uma vez, descartando a dependência do produtor quanto a essas empresas para expandir sua área de cultivo. Além disso, a cultura possui menor valor agregado quando comparada a outras hortaliças. Assim, cabe às instituições governamentais o desenvolvimento de pesquisas e obtenção de genótipos de alto desempenho. Nesse contexto, a automação da fenotipagem, por meio da análise computacional de imagens, pode ser uma ferramenta de grande utilidade para o melhoramento dessa cultura.

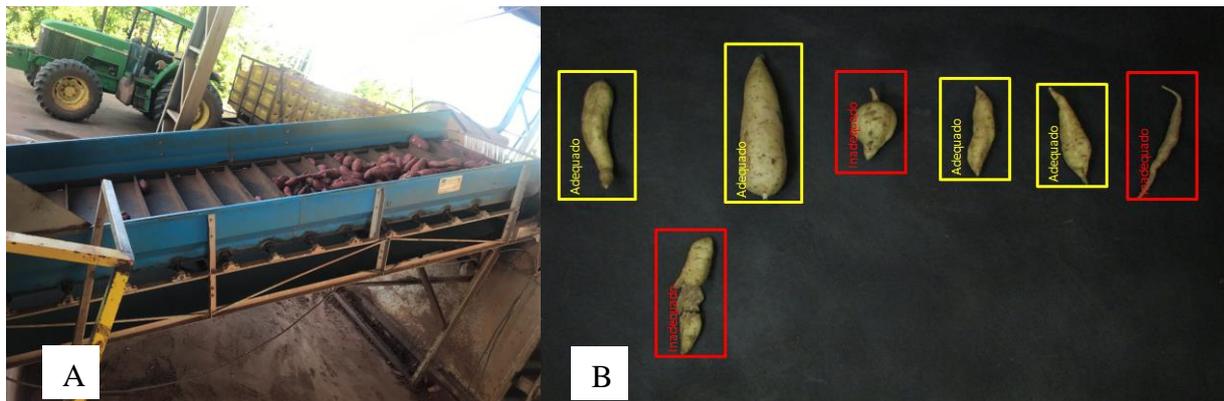
### 3.3 Fenotipagem de alta eficiência

A fenotipagem de plantas refere-se à descrição quantitativa das propriedades anatômicas, ontogenéticas, fisiológicas e bioquímicas da planta (WALTER; LIEBISCH; HUND, 2015). Para essa tarefa, a análise de imagens é uma ferramenta de grande utilidade. A análise de imagens se tornou um componente-chave, para preencher a lacuna do genótipo-fenótipo, em que a medição e contagem automatizada permitiram maior rendimento, desenvolvimento de sistemas consistentes e imparciais, além de caracterizar a fenotipagem não destrutiva da planta (WALTER; LIEBISCH; HUND, 2015). Assim, um requisito fundamental, para ferramentas de fenotipagem baseadas em imagens, é transformar automaticamente essas imagens em medições fenotípicas confiáveis e precisas (UBBENS; STAVNESS, 2017). Em uma perspectiva de análise de imagem, os recursos fenotípicos podem ser categorizados com base em sua complexidade (UBBENS; STAVNESS, 2017). Os fenótipos podem ser amplamente separados naqueles que são simplesmente funções lineares de intensidades de pixel de imagem ou tipos mais complexos que são funções não lineares de intensidades de pixel, que podem ser descrições geométricas ou não geométricas (UBBENS; STAVNESS, 2017).

Fenótipos gerados pela associação de imagens com programas de computador podem ser uma alternativa, para o futuro genético detalhado de várias espécies, podendo ser consultado, ao longo dos anos, de acordo com a conveniência do pesquisador. Ao contrário das medidas manuais que podem precisar serem executadas, durante um espaço de tempo pré-definido, enquanto as plantas estão vivas.

Na cultura da batata-doce, esteiras (**Figura 1-A**) poderiam ser adaptadas com a adição de câmeras permitindo a obtenção de imagens de várias raízes rapidamente. A partir dessas imagens, características de interesse agrônomo poderiam ser estimadas nas raízes, como o peso médio, produtividade, cor, formato (**Figura 1-B**) e danos ocasionados por pragas e insetos, por exemplo.

Figura 1- **(A)** Esteira utilizada para manejos pós-colheita na cultura da batata-doce; **(B)** possíveis classificações para o formato de raízes obtidas por análise computacional de imagem



Fonte: **(A)** <https://www.embrapa.br/en/hortalicas/batata-doce/pos-colheita>; **(B)** Do autor, 2022.

A fenotipagem de alta eficiência, ou fenômica, é a área da biologia que tem como objetivo a caracterização e mensuração das características físicas e bioquímicas de um indivíduo (FRITSCHENETO; BORÉM, 2015). A fenômica concentra-se na aquisição de dados fenotípicos, em larga escala e curto espaço de tempo, oferecendo a oportunidade de visualizar, em vários níveis, o fenótipo de um organismo (MASSRUHÁ *et al.*, 2014). Logo as técnicas tradicionais de fenotipagem de plantas avaliam características integrativas, facilmente mensuráveis, relacionadas ao peso e ao tamanho dos órgãos e estruturas vegetais, como parte aérea e as raízes das plantas, sendo esta última mais difícil de fenotipar em campo (DE SOUSA, 2014).

Essas novas tecnologias que visam aumentar a acurácia e rapidez das mensurações fenotípicas vêm sendo alvo de intensas pesquisas nos últimos anos (MASSRUHÁ *et al.*, 2014). Esse esforço recente visa diminuir o “gargalo de fenotipagem”, composta pela lacuna entre a quantidade e qualidade de dados genotípicos e fenotípicos disponíveis (SANTOS; YASSITEPE, 2014). Para a caracterização de genomas, técnicas correntes de sequenciamento tornaram um processo eficiente e barato, gerando uma enorme quantidade de dados para pesquisa. Porém esse avanço ainda vem em curtos passos, na caracterização em larga escala de fenótipos, retardando o avanço no entendimento das associações genótipo versus fenótipo (MASSRUHÁ *et al.*, 2014).

Para alcançar a mesma velocidade observada atualmente na caracterização genômica, é necessário incorporar à fenotipagem avanços nas áreas de visão computacional e análise de imagens, ressaltando a importância de associar a fenotipagem em campo a outras metodologias de caracterização. Entre as vantagens dessas técnicas, está a possibilidade de as medições serem realizadas, em qualquer horário, sem a necessidade de protocolo ou conhecimentos profundos, para a obtenção e interpretação dos dados, além de ser um método não destrutivo ou pouco destrutivo se comparado aos tradicionais (DE SOUSA, 2014).

Entre os estudos de fenotipagem de alta eficiência, pode-se destacar a quantificação de vitamina A em banana, utilizando redes neurais artificiais e dados colorimétricos (AQUINO; SALOMÃO; AZEVEDO, 2016), arquitetura de raízes por imagens (CLARK *et al.*, 2011), fenotipagem assistida por imagens digitais, em programa de reprodução de mamão (CORTES *et al.*, 2017), desenvolvimento de plataformas de

fenotipagem (UBBENS; STAVNESS, 2017), sistemas automatizados de fenotipagem de alta eficiência (LEE *et al.*, 2018), entre outros.

Observada a necessidade de melhoria na obtenção e análise de dados fenotípicos em batata-doce, principalmente pela grande variedade genética e fenotípica encontrada para a cultura, faz-se necessário estudar as possibilidades de uso de imagens tanto para as características de produção (quantitativas) quanto para as qualitativas, a fim de facilitar a fenotipagem da cultura.

### 3.4 Análise computacional de imagens

As análises visuais fazem parte da caracterização fenotípica de plantas e são realizadas de forma empírica desde os primórdios da agricultura (BARBOSA *et al.*, 2016). Contudo, com o avanço tecnológico e uso de câmeras digitais, bem como de scanners, o uso de imagens passam a ter espaço importante na análise de plantas (BARBOSA *et al.*, 2016). O aumento da disponibilidade de recursos computacionais permitiu estudos, em larga escala, utilizando análise de imagens (GAGE *et al.*, 2017), as quais são utilizadas a diversas finalidades, como avaliação do vigor de sementes de girassol (ROCHA; SILVA; CÍCERO, 2015), área do cladódio de clones de palma forrageira (DA SILVA *et al.*, 2014).

A análise de imagens possui vantagens sobre os métodos tradicionais, entre as quais pode-se destacar o fato de ser uma metodologia não destrutiva, rápida e reprodutível (BARBOSA *et al.*, 2016), além de permitir uma análise futura, uma vez que as imagens podem ser armazenadas em bancos de dados (ELICEIRI *et al.*, 2012). A fim de obter eficiência na análise das imagens digitais, é de suma importância o auxílio de programas computacionais que permitam a extração de informações (FRENCH *et al.*, 2009; ELMASRY *et al.*, 2009).

Uma imagem é composta por pixels, que representam o menor ponto que compõe uma imagem digital (BARBOSA *et al.*, 2016). Essas imagens podem ter espectro bidimensional, ou seja, em níveis de cinza, ou tridimensionais, contendo as três cores primárias, vermelho, verde e azul (RGB), em que a interpolação dessas cores representa a cor do pixel (BARBOSA *et al.*, 2016).

Para a extração de informações, as imagens devem ser submetidas a etapas de pré-processamento, sendo possível analisar e modificar imagens, objetivando-se reconhecer, comparar e classificar os elementos que a compõem, realçando bordas, melhorando o contraste e corrigindo imperfeições que possam influenciar na obtenção de informações (GONZALES; WOODS, 2013). Em suma, o processo de análise de imagens envolve a remoção de ruído de fundo e a binarização da imagem, alisando, esqueletizando e colocando splines para a imagem binária (GAGE *et al.*, 2017).

Essa etapa de pré-processamento consiste na mais importante, já que, ao melhorar a imagem, as chances de sucesso para as etapas seguintes aumentam, facilitando assim a segmentação (RODRIGUES, 2017). A segmentação consiste em dividir a imagem de entrada, em partes constituintes, separando as regiões da imagem em fundo (*background*) e objeto (*foreground*), a fim de produzir uma imagem binária (AMARAL, 2012).

### 3.5 Redes Neurais Convolucionais na análise de imagens

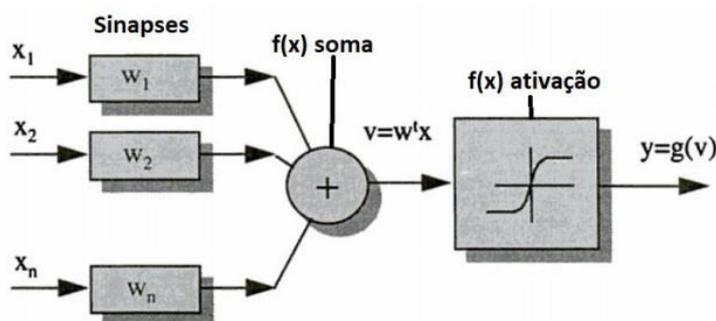
#### 3.5.1 Redes Neurais Artificiais

As redes neurais artificiais (RNA) são modelos computacionais não paramétricos que simulam o funcionamento do cérebro humano (AZEVEDO *et al.*, 2015). Capazes de reconhecer padrões e regularidades dos dados, tornam-se uma alternativa bem-sucedida para descrever, fundamentar e resolver funções complexas (GIANOLA *et al.* 2011; AZEVEDO *et al.*, 2015). As redes têm surgido como uma alternativa promissora, em diversos setores, pois apresentam bons ajustes e, diferente da regressão, conseguem utilizar variáveis quantitativas e qualitativas categóricas nas modelagens (HAYKIN, 2001).

Em função dessas características, o uso de RNAs associado a outras metodologias, pode se tornar eficaz na fenotipagem de culturas, como para a predição da produtividade e qualidade de raízes de batata-doce, além de já ter se mostrado eficiente na determinação indireta de vitaminas em banana (AQUINO; SALOMÃO; AZEVEDO, 2016). As RNAs possuem melhor desempenho, em comparação a outras técnicas de modelagem estatística, pois possuem ajuste universal de funções (GIANOLA *et al.*, 2011), admitem perda de dados e não exigem muita informação prévia sobre os fenômenos a serem modelados (AZEVEDO *et al.*, 2015), além de admitirem dados quantitativos e qualitativos como entrada (HAYKIN, 2001).

Muitas formas de RNA são propostas e usadas com variações, principalmente no número de neurônios por camada (BINOTI *et al.*, 2014). Os neurônios são elementos de processamento simples, em que cada neurônio é uma unidade de ativação, na qual é gerado um valor de saída. Nessas unidades, aplicam-se funções matemáticas aos dados (função de ativação), gerando uma única resposta (SILVEIRA, 2014) (Figura 2).

Figura 2- Rede Neural Artificial.



Fonte: KOVÁCS,2002.

Uma Rede Neural Artificial Multicamadas é tipicamente composta de camadas alinhadas de neurônios, em que a primeira camada é composta pelas variáveis de entrada. Essa camada distribui as informações às camadas intermediárias ou ocultas da rede. E a última camada é chamada de saída, na qual é obtida a solução do problema (DE OLIVEIRA *et al.*, 2010).

Outro fator que interfere significativamente no tempo de treinamento e na eficiência de uma RNA são os algoritmos utilizados, bem como os parâmetros e coeficientes utilizados em seu treinamento (BINOTI *et al.*, 2014). Em suma, a etapa de treinamento consiste em um problema de minimização não

linear sem restrições, em que os pesos sinápticos da rede são iterativamente modificados, para minimizar o erro quadrático médio entre a resposta desejada, a partir dos dados de entrada e a saída obtida no neurônio de saída (DE OLIVEIRA *et al.*, 2010). Qualquer problema a ser solucionado pelas RNAs, necessita-se conhecer valores de entrada e saída, devendo-se separar em conjunto treinamento e validação (DE OLIVEIRA *et al.*, 2010).

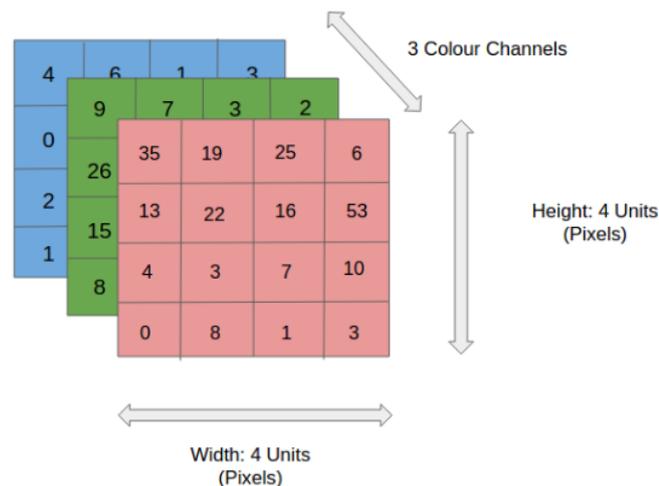
### 3.5.2 Redes Neurais Artificiais Convolucionais

A rede neural convolucional (ConvNets ou CNN's) trata-se de uma técnica de aprendizado profundo, baseada em multicamadas e permite captar uma imagem de entrada, atribuir importância (pesos e vieses que podem ser aprendidos) a vários aspectos / objetos da imagem e ser capaz de diferenciar um do outro, classificando-os e permitindo a extração de descritores de imagens (SHAHEEN *et al.*, 2016, ACADEMY, 2019).

As CNN's reconhecem as imagens como objetos tridimensionais, em vez de estruturas planas a serem medidas apenas por largura e altura. Esse fato está associado às cores das imagens que têm uma codificação RGB (Red-Green-Blue), misturando essas três cores para produzir o espectro de cores que os seres humanos percebem.

Uma rede convolucional recebe imagens como três estratos separados de cores empilhados um em cima do outro (**Figura 3**). Assim, uma rede convolucional recebe uma imagem como uma caixa retangular, cuja largura e altura são medidas pelo número de pixels, ao longo dessas dimensões e cuja profundidade é de três camadas profundas, uma para cada letra em RGB. Essas camadas de profundidade são referidas como canais (ACADEMY, 2019).

Figura 3- Imagem RGB (Vermelho – Verde – Azul) que foi separada por seus três planos coloridos – Vermelho, Verde e Azul.

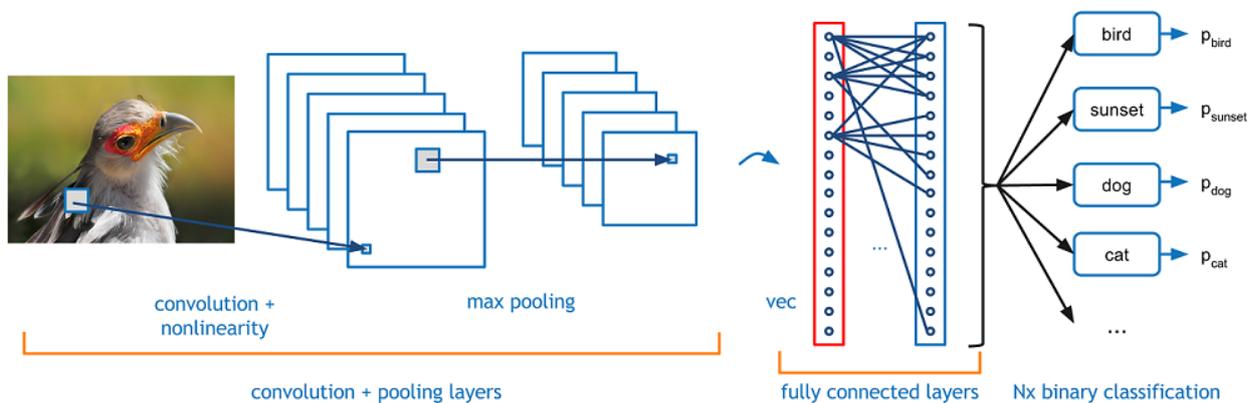


Fonte: Adaptado a partir de ACADEMY, 2019. <<https://www.deeplearningbook.com.br/introducao-as-redes-neurais-convolucionais/>>

À medida que as imagens se movem por uma rede convolucional, suas dimensões mudam até gerar uma série de probabilidades na camada de saída, sendo uma probabilidade para cada possível classe de saída. Aquela com maior probabilidade será a classe definida para a imagem de entrada (ACADEMY,

2019). Ou seja, uma CNN é composta de duas grandes etapas: a extração de características pelas convoluções e a classificação. A arquitetura de uma típica CNN possui basicamente três tipos de camadas, as convolucionais, pooling e camadas totalmente conectadas. A camada de convolução contém um conjunto de neurônios que são responsáveis por aplicar filtros em determinadas partes da imagem. Gonzalez e Woods (2009) definem a convolução como resultante do produto escalar entre a máscara e a vizinhança de cada pixel da imagem. A camada de pooling tem como principal objetivo a redução de dimensionalidade dos dados na rede, para tal, a operação mais utilizada é o max-pooling, cujas partes mais importantes da imagem são capturadas. A **Figura 4** apresenta um exemplo genérico de CNN, na qual essas etapas são denominadas convolution + pooling layers.

Figura 4: Exemplo de CNN, para a classificação de espécies de animais, com camadas de convolução, max pooling e uma camada totalmente conectada. Sua saída é a classificação da imagem de entrada em alguma classe.



Fonte: Adaptado a partir de ACADEMY, 2019.

A arquitetura de uma ConvNet é análoga àquela do padrão de conectividade de neurônios no cérebro humano e foi inspirada na organização do Visual Cortex. Os neurônios individuais respondem a estímulos apenas em uma região restrita do campo visual conhecida como Campo Receptivo. Uma coleção desses campos se sobrepõe para cobrir toda a área visual (ACADEMY, 2019). Dessa forma, a aplicação das redes neurais é baseada na captura de dependências espaciais e temporais em uma imagem através da aplicação de filtros relevantes. Assim, a arquitetura da rede executa um melhor ajuste ao conjunto de dados da imagem com a redução no número de parâmetros envolvidos e com capacidade de reutilização dos pesos (ACADEMY, 2019). Nesse sentido, diz-se que cada valor da matriz de convolução representa o peso de uma sinapse e cada iteração errada da rede (convolução) é modificada e adaptada para tentar gerar o resultado esperado (KRIZHEVSKY, 2012). Uma convolução é composta por três elementos: a entrada, que é uma matriz de valores numéricos (Imagem); o kernel, que é uma matriz de dimensões menores que a entrada; e a saída, que é o resultado da convolução (GOODFELLOW; BENGIO; COURVILLE, 2016).

As CNNs são mais fáceis de treinar, quando existe grande quantidade de amostras rotuladas, que representam as diferentes classes-alvo. Dessa forma, podem-se destacar como vantagens dessas redes a capacidade de extrair características relevantes por aprendizado de transformações (kernels) e a

capacidade de depender de menor número de parâmetros de ajustes se comparada com redes totalmente conectadas (ARAUJO *et al.*,2017). Já Vargas *et al.* (2016) citam como diferencial a possibilidade de a CNN misturar múltiplos mapas de características, ao mesmo tempo, permitindo a extração de características cada vez mais complexas. Assim, as CNNs são capazes de criar filtros extremamente complexos, aproveitando ao máximo as informações provenientes dos dados de treinamento.

As CNNs podem ser divididas em várias arquiteturas, que auxiliam na melhoria do desempenho de redes neurais profundas (SZEGEDY *et al*, 2016). Essas arquiteturas variam, de acordo com o tamanho, precisão, quantidade de parâmetros, profundidade e tempo por etapa de inferência. Dessa forma, podem-se destacar as arquiteturas como: VGG-16, Inception-v3, ResNet-50, InceptionResNetV2, EfficientNetB3 (Keras, 2022) (**Tabela 1**).

Tabela 1: Características de cada arquitetura

Variáveis	Modelo				
	VGG16	Inception V3	ResNet50	InceptionResNetV2	EfficientNetB3
Tamanho (MB)	528	92	98	215	48
Precisão principal	71,3%	77,9%	74,9%	80,3%	81,6%
Precisão Top-5	94,5%	93,7%	92,1%	95,3%	95,7%
Parâmetros (x10 <sup>6</sup> )	139,4	23,9	25,6	55,9	12,3
Profundidade	16	189	107	449	210
Tempo (ms) por etapa de inferência (CPU)	69,5	42,2	58,2	130,2	140
Tempo (ms) por etapa de inferência (GPU)	4,2	6,9	4,6	10,0	8,8

Fonte: Adaptado de Keras,2022.

### 3.6 REFERÊNCIAS

- ACADEMY, D. S. **As 10 principais arquiteturas de redes neurais**. In: Deep Learning Book. [s.n.], 2019. cap. 10. Disponível em: <<http://www.deeplearningbook.com.br>>.
- AMARAL, A.; G. **Processamento de imagens digitais para avaliação do comportamento e determinação do conforto térmico de codornas de corte**. 2012. Tese (Doutorado em Engenharia Agrícola), Universidade Federal de Viçosa- UFV, Viçosa, Minas Gerais.2012.
- ANDRADE JÚNIOR, V.C. DE; VIANA, D. J. S.; PINTO, N. A.V.D.; RIBEIRO, K. G.; PEREIRA, R. C.; NEIVA, I. P.; AZEVEDO, A. M.; ANDRADE, P. C. DE R. Características produtivas e qualitativas de ramas e raízes de batata-doce. **Horticultura Brasileira**, v.30, n.4, p.584-589, 2012.
- AQUINO, C. F.; SALOMAO, L. C. C.; AZEVEDO, A. M. High-efficiency phenotyping for vitamin A in banana using artificial neural networks and colorimetric data. *Bragantia*, Campinas, v. 75, n. 3, p. 268-274, Sept, 2016.
- ARAÚJO, F. H.; CARNEIRO, A. C.; SILVA, R. R.; MEDEIROS, F. N.; USHIZIMA, D. M. Redes neurais convolucionais com tensorflow: Teoria e prática. Sociedade Brasileira de Computação. III Escola Regional de Informática do Piauí. **Livro Anais-Artigos e Minicursos, Sociedade Brasileira de Computação**, v. 1, p. 382–406, 2017.
- AZEVEDO, A. M.; ANDRADE JÚNIOR, V. C.; FERNANDES, J. S. C.; PEDROSA, C. E.; OLIVEIRA, C. M. Parâmetros genéticos e ganho com seleção em batata-doce. **Horticultura Brasileira**, v. 33, n. 01, 2015a.
- AZEVEDO, A. M.; ANDRADE JÚNIOR, V. C. DE; PEDROSA, C. E.; OLIVEIRA, C. M. DE; DORNAS, M. F.; CRUZ, C. D.; VALADARES, N. R. Application of artificial neural networks in indirect selection: a case study on the breeding of lettuce. **Bragantia**, v.74, n.4, p.387-393, 2015.
- AZEVEDO, A.M.; ANDRADE JÚNIOR, V.C.; VIANA, D.J.S.; ELSAYED, A.Y.; PEDROSA, C.E.; NEIVA, I.P.; FIGUEIREDO, J.A. Influence of harvest time and cultivation sites on the productivity and quality of sweet potato. **Horticultura Brasileira**, v.32, p. 21-27, 2014.
- AZEVEDO, S.M.; MALUF, W.R.; SILVEIRA, M.A.; FREITAS, J.A. Reação de clones de batata-doce aos insetos de solo. **Ciência e Agrotecnologia**, v. 26, p.545 549, 2002.
- BARBOSA, J. Z.; CONSALTER, R.; PAULETTI, V.; MOTTA, A. C. Uso de imagens digitais obtidas com câmeras para analisar plantas. **Revista de Ciências Agrárias**, v.39, n.1, p.15-24, 2016.
- BINOTI, D. H. B.; DA SILVA BINOTI, M. L. M.; LEITE, H. G. Configuração de redes neurais artificiais para estimação do volume de árvores. **Revista Ciência da Madeira**, Pelotas, v.5, n. 1, p. 58-67, 2014.
- BOVELL-BENJAMIN, A. C. Sweet potato: a review of its past, present, and future role in human nutrition. **Advances in Food and Nutrition Research**, v. 52, p. 1-59, 2007.
- CAMIRE, M. E.; KUBOW, S.; DONNELLY, D. J. Potatoes and human health. **Critical Reviews in Food Science and Nutrition**, v. 49, n. 10, p. 823-840, 2009.
- CARDOSO, A.D.; VIANA, A.E.S.; RAMOS, P.A.S.; MATSUMOTO, S.N.; AMARAL, C.L.F.; SEDIYAMA, T.; MORAIS, O.M. Avaliação de clones de batata-doce em Vitória da Conquista. **Horticultura Brasileira**, v.23, p. 911-914, 2005.
- CAVALCANTE, M.; FERREIRA, P.V.; PAIXÃO, S.L.; COSTA, J.C.; PEREIRA, R.G.; MADALENA, J.A.S. Potenciais produtivo e genético de clones de batata-doce. **Acta Scientiarum Agronomy**, v.31, p.421-426. 2009.
- CLARK, R. T., MACCURDY, R. B., JUNG, J. K., SHAFF, J. E., MCCOUCH, S. R.; ANESHANSLEY, D. J.; KOCHIAN, L. V. Three-dimensional root phenotyping with a novel imaging and software platform. **Plant Physiology**, Bethesda, v. 156, n. 2, p. 455-465. 2011.

CORTES, D. F. M.; CATARINA, R. S.; BARROS, G. B. DE A.; ARÊDES, F.; A.; S.; SILVEIRA, S.; F.; DA, FERREGUETTI, G.; A.; RAMOS, H.; C.; C.; VIANA, A.; P.; PEREIRA, M.; G. Model-assisted phenotyping by digital images in papaya breeding program. **Scientia Agricola**, v.74, n.4, p.294-302, 2017.

DA SILVA, T. G.; DE MIRANDA, K. R.; DOS SANTOS, D. C.; DE QUEIROZ, M. G.; SILVA, M. D. C.; DA CRUZ NETO, J. F.; & ARAÚJO, J. E. Área do cladódio de clones de palma forrageira: modelagem, análise e aplicabilidade. **Revista Brasileira de Ciências Agrárias**, v.9, n.4, p.633-641, 2014.

DE OLIVEIRA, A. C. S.; DE SOUZA, A. A.; LACERDA, W. S.; GONÇALVES, L. R. Aplicação de redes neurais artificiais na previsão da produção de álcool. **Ciênc. agrotec.**, Lavras, v. 34, n. 2, p. 279-284, 2010.

DE SOUSA, C. A. F. Fenotipagem de plantas: uma nova abordagem para um velho problema. Embrapa Agroenergia-**Comunicado Técnico (INFOTECA-E)**. 2014.

DONATO, L. M. S.; ANDRADE JUNIOR, V. C. DE; BRITO, O. G.; FIALHO, C. M. T.; SILVA, A. J. M. DA; AZEVEDO, A. M. Uso de ramas de batata-doce para produção de feno. **Ciência Animal Brasileira**, v. 21, e-53493, 2020.

ELICEIRI, K.W.; BERTHOLD, M.R.; GOLDBERG, I.G.; IBÁÑEZ, L.; MANJUNATH, B.S.; MARTONE, M.E.; MURPHY, R.F.; PENG, H.; PLANT, A.L.; ROYSAM, B.; STUURMAN, N.; SWEDLOW, J.R.; TOMANCAK, P. E CARPENTER, A.E. Biological imaging software tools. **Nature Methods**, v. 9, n. 7, p. 697-710, 2012.

ELMASRY, G.; RADWAN, S.; ELAMIR, M. E ELGAMAL, R. Investing the effect of moisture content on some properties of peanut by aid of digital image analysis. **Food and Bioproducts Processing**, vol. 87, n. 4, p. 273-281. 2009.

FANG, Y.; XIONG, L. General mechanisms of drought response and their application in drought resistance improvement in plants. **Cellular and Molecular Life Sciences**, v.72, p.673–689, 2015.

FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3.ed. Brasília: EMBRAPA-CENARGEM, 1998. 220p.

FIGUEIREDO, J.A.; ANDRADE JÚNIOR, V.C. DE; PEREIRA, R.C.; RIBEIRO, K.G.; VIANA, D.J.S.; NEIVA, I.P. Avaliação de silagens de ramas de batata-doce. **Horticultura Brasileira**. V.30, p.708-712, 2012.

FRENCH, A.; UBEDA-TOMÁS, S.; HOLMAN, T.J.; BENNETT, M.J. E PRIDMORE, T. High-throughput quantification of root growth using a novel image-analysis tool. **Plant Physiology**, v. 150, n. 4, p. 1784-1795, 2009.

FRITSCHÉ-NETO R.; BORÉM A. **Fenômica: como a fenotipagem de próxima geração está revolucionando o melhoramento de plantas**. 1. ed. Visconde do Rio Banco, MG: Editora Suprema, 2015. 216p.

GAGE, J. L.; MILLER, N. D.; SPALDING, E. P.; KAEPLER, S. M.; DE LEON, N. TIPS: a system for automated image-based phenotyping of maize tassels. **Plant Methods**. v.13(1), n. 21, 2017.

GIANOLA, D.; OKUT, H.; WEIGEL, K. A.; ROSA, G. J. Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. **BMC Genetics**, v.12, p.87-101, 2011.

GONÇALVES NETO, A.C.; MALUF, W.R.; GOMES, L.A.A.; GONÇALVES, R.J.S.; SILVA, V.F.; LASMAR, A. Aptidões de genótipos de batata-doce para consumo humano, produção de etanol e alimentação animal. **Pesquisa Agropecuária Brasileira**, v.46, p.1513-1520, 2011.

GONÇALVES NETO, A.C.; MALUF, W.R.; GOMES, L.A.A.; MACIEL, G.M.; FERREIRA, R.P.D.; CARVALHO, R.C. Correlação entre caracteres e estimação de parâmetros populacionais para batata-doce. **Horticultura Brasileira**, v.30, p.713-719, 2012.

GONZALEZ, R.; C.; WOODS, R.; E. **Processamento de imagens digitais**. Tradução; Roberto Marcondes Cesar Junior, Luciano da Fontoura Costa. São Paulo, Editora: Blucher. ISBN:978-85-212-0264-7, 2013.

GOODFELLOW, Ian et al. **Deep learning**. Cambridge: MIT press, 2016.

HAQUE, S.; LOBATON, E.; NELSON, N.; YENCHO, G.C.; PECOTA, K.V.; MIEROP, R.; KUDENOV, M. W.; BOYETTE, M.; WILLIAMS, C.M. Abordagem de visão computacional para caracterizar fenótipos de tamanho e forma de culturas hortícolas usando imagens de alto rendimento. *Computadores e Eletrônicos na Agricultura*, v.182, 106011, 2021.

HAYKIN, S. **Redes neurais: princípios e prática**. Porto Alegre: Bookman. 2001, 900p.

HENRY, R. J. **Plant resources for food, fuel and conservation**. London: Earthscan, 2012, 153p.

HUANG, S.; WEIGEL, D.; BEACHY, R.N.; LI, J. A proposed regulatory framework for genome-edited crops. **Nature Genetics**. V.48, n.2, 2016.

KEHOE, S. H.; CHOPRA, H.; SAHARIAH, S. A.; BHAT, D.; MUNSHI, R. P.; PANCHAL, F.; YOUNG, S.; BROWN, N.; TARWANDE, D.; GANDHI, M.; MARGETTS, B. M.; POTDAR, R. D.; FALL, C. H. Effects of a food-based intervention on markers of micronutrient status among Indian women of low socio-economic status. **British Journal of Nutrition**, v. 113, n. 5, p. 813-821, 2015. PMID:25677713.

KERAS, API REFERENCE, **Keras Applications**. 2022. <<https://keras.io/api/applications/>>. Acesso em 27 de outubro de 2022.

KOVÁCS, Zsolt László. **Redes neurais artificiais**. Editora Livraria da Física, 2002.

KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. Imagenet classification with deep Convolutional neural networks. In: **Advances in neural information processing systems**. p. 1097-1105, 2012.

LAY, C. H.; LIN, H. C.; SEN, B.; CHU, C. Y.; LIN, C. Y. Simultaneous hydrogen and ethanol production from sweet potato via dark fermentation. **Journal of Cleaner Production**, v.27, p.155-164, 2012.

LEE, U.; CHANG, S.; PUTRA, G. A.; KIM, H.; KIM, D. H. An automated, high-throughput plant phenotyping system using machine learning-based plant segmentation and image analysis. **PloS one**, v. 13, n. 4, p. e0196615, 2018.

LOW, J. W.; ARIMOND, M.; OSMAN, N.; CUNGUARA, B.; ZANO, F.; TSCHIRLEY, D. A food-based approach introducing orange-fleshed sweet potatoes increased Vitamin A intake and serum retinol concentrations in young children in rural Mozambique. **The Journal of Nutrition**, v. 137, n. 5, p. 1320-1327, 2007. PMID:17449599.

MASSRUHÁ, S. M. F. S.; LEITE, M. A. A.; LUCHIARI JUNIOR, A.; ROMANI, L. A. S. **Tecnologias da Informação e Comunicação e suas relações com a agricultura**. Brasília, DF: Embrapa, p. 85–100, 2014.

MELO, P.C.T.; MELO, A.M.T.; ARAGÃO, F.A.S. Melhoramento Genético de Hortaliças no Brasil: Retrospectiva e Perspectivas. **Simpósio Nordestino de Genética e Melhoramento de Plantas**. O melhoramento genético no contexto atual: Anais / I Simpósio Nordestino de Genética e melhoramento de Plantas. Fortaleza: Embrapa Agroindústria Tropical, 2009. 210p.

MOLLINARI, M.; OLUKOLU, B. A.; PEREIRA, G. S.; KHAN, A.; GEMENET, D.; YENCHO G. C.; ZENG, Z. B. Unraveling the hexaploid sweetpotato inheritance using ultra-dense multilocus mapping. **G3: Genes, Genomes, Genetics**, v.10, n.1, p.281-292, 2020.

MONTEIRO, A.B.; MASSAROTO, J.A.; GASPARINO, C.F.; SILVA, R.R.; GOMES, L.A.A.; MALUF, W.R.; FILHO, J.C.S. Silagens de cultivares e clones de batata-doce para alimentação animal visando sustentabilidade da produção agrícola familiar. **Revista Brasileira de Agroecologia**, v.2, n.2, p.978-981, 2007.

PAVLAK, M. C. DE M.; ABREU-LIMA, T. L. DE; CARREIRO, S. C.; PAULILLO, S. C. DE L. Estudo da fermentação do hidrolisado de batata-doce utilizando diferentes linhagens de *Saccharomyces cerevisiae*. **Química Nova**, v.34, n.1, p. 82-86, 2011.

PEDROSA, C.E.; ANDRADE JÚNIOR, V.C.; PEREIRA, R.C.; DORNAS, M.F.; AZEVEDO, A.M.; FERREIRA, M.A. Yield and quality of wilted sweet potato vines and its silages. **Horticultura Brasileira**, v.33, p.283-9, 2015.

PHILPOTT, M.; GOULD, K.S.; LIM, C.; FERGUSON, L.R. In situ and in vitro antioxidant activity of sweetpotato anthocyanins **J. Agr. Food Chem**, v.52, p.1511-1513, 2004.

PRADHAN, D.M.P.; MUKHERJEE, A.; GEORGE, J.; CHAKRABARTIS, K.; VIMALA B.; NASKAR, S.K.; SAHOOB, K.; SAMAL, S. High starch, beta carotene and anthocyanin rich sweet potato: ascent to future food and nutrition security in coastal and backward areas. **International Journal of Tropical Agriculture**, v. 33, n. 2 (Part I), p.397-400, 2015.

RAHAJENG, W., RESTUONO, J., INDRIANI, F. C., PURWONO, P. Genetic Parameters of Agronomic Traits in Sweetpotato Accessions. **Biosaintifika: Journal of Biology & Biology Education**, v. 12, n. 2, p. 240-246, 2020.

Rede Neural Convolucional Resnet-50. **MathWorks**, 2022. <<https://www.mathworks.com/help/deeplearning/ref/resnet50.html>>. Acesso em 27 de outubro de 2022.

RINCKER, K.; NELSON, R.; SPECHT, J.; SLEPER, D.; CARY, T.; CIANZIO, S.R.; CASTEEL, S.; CONLEY, S.; CHEN, P.; DAVIS, V.; FOX, C.; GRAEF, G.; GODSEY, C.; HOLSHOUSER, D.; JIANG, G.L.; KANTARTZI, S.K.; KENWORTHY, W.; LEE, C.; MIAN, R.; MCHALE, L.; NAEVE, S.; ORF, J.; POYSA, V.; SCHAPAUGH, W.; SHANNON, G.; UNIATOWSKI, R.; WANG, D.; DIERS, B. Genetic Improvement of U.S. Soybean in Maturity Groups II, III, and IV. **Crop Science**. V.54, p.1419–1432, 2014.

RITSCHER, P.S.; LOPES, C.A.; HUAMÁN, Z.; FERREIRA, M.E.; FRANCA, F.J.; MENÊZES, J.E.; TEIXEIRA, D.M.C.; TORRES, A.C.; CHARCHAR, J.M.; THOMAZELLI, L. Organização do banco ativo de germoplasma de batata-doce: situação atual e perspectivas. **Recursos genéticos e melhoramento de plantas para o Nordeste brasileiro: versão**, v. 1, 2010. Disponível em:< <http://www.cpatas.embrapa.br/catalogo/livro/g/batatadoce.pdf>>. Acesso em 20 de janeiro de 2021.

ROCHA, C. R. M. da; SILVA, V. N.; CICERO, S. M. Avaliação do vigor de sementes de girassol por meio de análise de imagens de plântulas. **Cienc. Rural**, Santa Maria, v.45, n.6, p.970-976, June, 2015.

RODRIGUES, T. P.N. da S. **Processamento digital de imagens para análise do comportamento e bem-estar de frangos de corte**. Programa de Pós- graduação em Engenharia Agrícola (Tese). Recife. 2017, 81f.

ROULLIER, C.; KAMBOUO, R.; PAOFA, J.; MCKEY, D.; LEBOT, V. On the origin of sweet potato (*Ipomoea batatas* (L.) Lam.) genetic diversity in New Guinea, a secondary centre of diversity. **Heredity**, Edinburgh, v. 110, n. 6, p. 594–604, 2013.

SANTOS, T. T.; YASSITEPE, J. E. C. T. 2014. Fenotipagem de plantas em larga escala: um novo campo de aplicação para a visão computacional na agricultura. In: MASSRUHÁ, S. M. F. S.; LEITE, M. A. A.; LUCHIARI JUNIOR, A.; ROMANI, L. A. S. **Tecnologias da Informação e Comunicação e suas relações com a agricultura**, Brasília, DF: Embrapa, p. 85–100, 2014.

SANTOS NETO, A. R.; SILVA, T. O.; BLANK, A. F.; SILVA, J. O.; F FILHO, R. N. A. Produtividade de clones de batata-doce em função de doses de nitrogênio. **Horticultura Brasileira**, v.35, n.3, p.445-452, 2017.

SHAHEEN, F.; VERMA, B. E ASAFUDDOULA, M. IMPACT OF AUTOMATIC FEATURE EXTRACTION IN DEEP LEARNING ARCHITECTURE. In *Digital Image Computing: Techniques and Applications (DICTA)*. **International Conference on**, pp. 1–8. IEEE, 2016.

SILVA, J.B.C.; LOPES, C.A.; MAGALHÃES, J.S. **Batata-doce (*Ipomoea batatas*)**. Embrapa Hortaliças, 2008.

SILVEIRA, D. P. **Estimação do volume de madeira empilhada utilizando imagens digitais e redes neurais**. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa. 2014. 66f.

SOUSA, C.S.; BONETTI, A.M.; GOULART FILHO, L.R.; MACHADO, J.R.A.; LONDE, L.N.; BAFFI, M.A.; RAMOS, R.G.; VIEIRA, C.E.; KERR, W.E. Divergência genética entre genótipos de alface por meio de marcadores AFLP. **Bragantia**, v.66, n.1, p.11-16, 2007.

SOUSA, T. V.; CAIXETA, E. T.; ALKIMIM, E. R.; OLIVEIRA, A. C. B.; PEREIRA, A. A.; SAKIYAMA, N. S.; ZAMBOLIM, L.; RESENDE, M. D. V. Early selection enabled by the implementation of genomic selection in *Coffea arabica* breeding. **Frontiers in Plant Science**, v. 9, p. 1934, 2019.

SOUZA, A.B. Avaliação de cultivares de batata-doce quanto atributos agronômicos desejáveis. *Lavras. Ciência Agrotécnica*, v.24, n.4, p.841-845, 2000.

SZEGEDY, C. et al. Rethinking the inception architecture for computer vision. **IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**, p. 2818–2826, ISSN 1063-6919, 2016.

TEOW, C.C.; TRUONG, V.D.; MCFEETERS, R.F.; THOMPSON, R.L.; PECOT, K.V.; YENCHO, G.C. Antioxidant activities, phenolic and  $\beta$ -carotene contents of sweetpotato genotypes with varying flesh colors **Food Chem**, v.103, p.829-838, 2007.

TODD, S. M.; TRUONG, V. D.; PECOTA, K. V.; YENCHO, G. C. Combining ability of sweet potato germplasm for yield, dry matter content, and anthocyanin production. *Journal of the American Society for Horticultural Science*, v. 140, n. 3, p. 272-279, 2015.

UBBENS, J. R.; STAVNESS, I. Deep plant phenomics: a deep learning platform for complex plant phenotyping tasks. **Frontiers in plant science**, v. 8, p. 1190, 2017.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: **Proceedings of the XXIX Conference on Graphics, Patterns and Images**. p. 1–4, 2016.

VIANA, D. J. S.; ANDRADE JÚNIOR, V. C. DE; RIBEIRO, K. G.; PINTO, N. A.V. D.; NEIVA, I. P.; FIGUEIREDO, J. A.; LEMOS, V. T.; PEDROSA, C. E.; AZEVEDO, A. M. Potencial de silagens de ramas de batata-doce para alimentação animal. **Ciência Rural**, v.41, n.8, p.1466-1471, 2011.

VIZZOTTO, M.; PEREIRA E. DOS S.; DE CASTRO, L. A. S.; RAPHAELL, C. DE O.; KROLOW, A.C. Composição mineral em genótipos de batata-doce de polpas coloridas e adequação de consumo para grupos de risco. **Brazilian Journal of Food Technology**, v. 21, 2017.

WALTER, A.; LIEBISCH, F.; HUND, A. Plant phenotyping: from bean weighing to image analysis. **Plant methods**, v.11, n.1, p.1-11, 2015.

WANG, X.; TIAN, S.; LOU, H.; ZHAO, R. A reliable method for predicting bioethanol yield of different varieties of sweet potato by dry matter content. **Grain & Oil Science and Technology**, v.3, n.3, p.110-116, 2020.

WERA, B.; YALU, A.; RAMAKRISHNA, A.; DEROS, M. Genotypic Variability Estimates Of Agronomic Traits For Selection In A Sweetpotato (*Ipomoea batatas*) Polycross Population In Papua New Guinea. *J. Plant Breed. Genet.* v.2, n.3, p.131- 136, 2014.

WOOLFE, J. A. **Sweet potato: an untapped food resource**. Cambridge: Cambridge University Press, International Potato Centre, 2008.

## 4 ARTIGOS

### 4.1 CAPÍTULO I: Visão computacional no melhoramento genético para a produção de raízes de batata-doce

Artigo aceito no periódico Horticultura Brasileira.

#### Resumo

O melhoramento da batata-doce é um trabalho oneroso em decorrência do grande número de características analisadas, para a seleção dos melhores genótipos, fazendo-se necessária a adoção de novas tecnologias, como o uso de imagens, associadas ao processo de fenotipagem. Com isso, objetivou-se desenvolver uma metodologia para a fenotipagem da produção de raízes, para o melhoramento genético de progênies de meios irmãos de batata-doce, por meio da análise computacional de imagens e comparar seu desempenho com a metodologia tradicional de avaliação. Foram avaliadas 16 progênies de meios irmãos de batata-doce, em delineamento em blocos casualizados com quatro repetições. Avaliou-se, em nível de plantas, o peso por raiz. As imagens foram adquiridas em um “estúdio” feito com MDF com uma câmera digital modelo Canon PowerShot SX400 IS, sob iluminação artificial. As avaliações foram realizadas no software R. Para a predição do peso das raízes (em gramas), foi ajustado um modelo de regressão polinomial de segundo grau e foram obtidos os valores genéticos e ganhos esperados. A visão computacional permitiu a predição do peso das raízes, mantendo o ranqueamento dos genótipos e, conseqüentemente, a similaridade entre os ganhos esperados com a seleção. Assim, o uso de imagens é uma ferramenta eficiente aos programas de melhoramento genético da batata-doce, auxiliando no processo de fenotipagem da cultura.

**Palavras-chave:** *Ipomoea batatas* (L.) Lam. Análise de imagens. Fenotipagem de alta eficiência.

#### 4.1.1 Introdução

Entre as hortaliças mais importantes, no Brasil, destaca-se a batata-doce (*Ipomoea batatas* (L.) Lam.), podendo garantir a segurança alimentar de uma população pelo seu alto valor nutritivo, ser fonte significativa de carboidratos, vitaminas e minerais (VARGAS *et al.*, 2017). Pela sua rusticidade, fácil manejo e baixo custo de produção, a batata-doce é bastante utilizada por pequenos produtores, principalmente, por seus múltiplos usos tanto na alimentação humana quanto animal. Dessa forma, a hortaliça destaca-se pela sua grande importância socioeconômica, bem como alto potencial de rendimento e ampla variabilidade genotípica e fenotípica (OTOBONI *et al.*, 2020). Nesse sentido, esforços para o melhoramento genético da cultura são necessários, visando ao aumento de produtividade e qualidade das raízes.

O melhoramento genético de plantas é considerado uma das mais importantes estratégias na agricultura moderna, visando maximizar a produtividade e a qualidade dos alimentos por meio de indivíduos superiores aos pré-existentes (AMABILE, VILELA, PEIXOTO, 2018). As variedades melhoradas geralmente são mais produtivas, apresentam resistência a estresses bióticos e abióticos, maior qualidade nutracêutica e, conseqüentemente, reduzem o impacto ambiental e custo de produção (FANG; XIONG, 2015; RINCKER *et al.*, 2014; HUANG *et al.*, 2016).

O melhoramento genético da batata-doce é um trabalho oneroso, em decorrência do grande número de características analisadas, uma vez que inferências em campo, a fim de classificar os melhores genitores, devem basear-se em valores genotípicos e não apenas em valores fenotípicos (BORGES *et al.*, 2010). Assim, essas análises demandam uma grande quantidade de mão de obra, financiamento e disponibilidade de tempo, principalmente no trabalho com progênes de meios irmãos, uma vez que a análise dos caracteres deve ser realizada em nível de plantas.

Tendo em vista essa dificuldade de análise dos dados fenotípicos das raízes de batata-doce, faz-se necessário a utilização de ferramentas alternativas de fenotipagem. O processo de fenotipagem de plantas refere-se ao conjunto de metodologias usadas, para medir características de plantas em condições ambientais distintas com exatidão e precisão (MAHLEIN, 2016). Assim, o desenvolvimento de tecnologias de alto rendimento é de suma importância ao melhoramento genético de plantas, já que permitem a coleta de dados em grande escala com o mínimo de esforço (ZHOU *et al.*, 2018).

Nessa lógica, o uso de imagens apresenta grande potencial em aumentar o rendimento e a precisão das atividades de fenotipagem da batata-doce, conforme o verificado por Haque *et al.* (2021), na avaliação atributos qualitativos com menor custo, maior rapidez, eficiência e precisão. Uma vantagem dessa metodologia é que ela permite a reprodutibilidade dos resultados, após descarte do material vegetal, uma vez que o banco de imagens pode ser preservado. Porém, além de ser possível a fenotipagem de raízes de batata-doce para caracteres qualitativos (HAQUE *et al.*, 2021), é possível que informações sobre a produtividade de raízes também sejam obtidas pela análise de imagem. Caso possível, a fenotipagem de raízes de batata-doce para caracteres qualitativos e quantitativos simultaneamente pode ser de grande importância para o melhoramento genético da cultura.

Neste contexto, objetivou-se desenvolver uma metodologia para a fenotipagem da produção de raízes de progênes de meios irmãos de batata-doce, por meio da análise computacional de imagens e comparar seu desempenho com a metodologia tradicional de avaliação.

## **4.1.2 Material e Métodos**

### **4.1.2.1 Instalação e avaliação do experimento**

O experimento foi desenvolvido no Instituto de Ciências Agrárias da UFMG- Campus Montes Claros-MG (ICA/UFMG) (coordenadas: 16°40'58.16"S e 43°50'20.15"O) em que foram avaliadas 16 progênes de meios-irmãos de batata-doce (BELGARD (F2), CAMBRAIA (F4), LICURI (F5), UFVJM40 (F6), UFVJM01 (F7), ARRUBA (F8), UFVJM05 (F10), UFVJM15 (F13), UFVJM56 (F16), UFVJM31 (F20), UFVJM37 (F22), UFVJM54 (F24), UFVJM25 (F26), UFVJM29 (F27), TCARRO02 (F29), UFVJM09 (F25)), em cambissolo háplico sob condições de irrigação.

Os meios-irmãos de batata-doce foram obtidos pela coleta de sementes oriundas do banco de germoplasma, composto por acessos elite trazidos da Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM) e cultivado no ICA/UFMG. As sementes foram coletadas diariamente entre os meses de abril e outubro de 2018 e acondicionados em geladeira a 4°C. Posteriormente, as sementes foram submetidas à escarificação mecânica, com lixa para quebra de dormência (impermeabilidade tegumentar) e plantadas em bandejas de poliestireno de 72 células com substrato comercial. As bandejas foram mantidas em casa de vegetação e irrigadas diariamente, durante dois meses, quando as mudas se encontraram prontas para o plantio.

O plantio foi realizado em leiras, em delineamento de blocos ao acaso (DBC), com 16 progênies e quatro repetições, com leiras espaçadas entre si de 1m e espaçamento entre plantas de 0,4 m, em que foram avaliadas 10 plantas por parcela. Como as avaliações foram feitas em nível de planta, foi utilizado um espaçamento maior, a fim de facilitar a identificação de cada planta e facilitar a colheita. As adubações e tratos culturais foram realizados como o recomendado para a cultura no Novo Manual de Olericultura (FILGUEIRA, 2008). Foram utilizados 180 kg ha<sup>-1</sup> de fósforo e 30 kg ha<sup>-1</sup> de nitrogênio. Aos 30 dias do plantio das mudas, foi realizada uma adubação de cobertura com 30 kg ha<sup>-1</sup> de nitrogênio. A adubação de potássio não foi necessária de acordo com a análise química do solo.

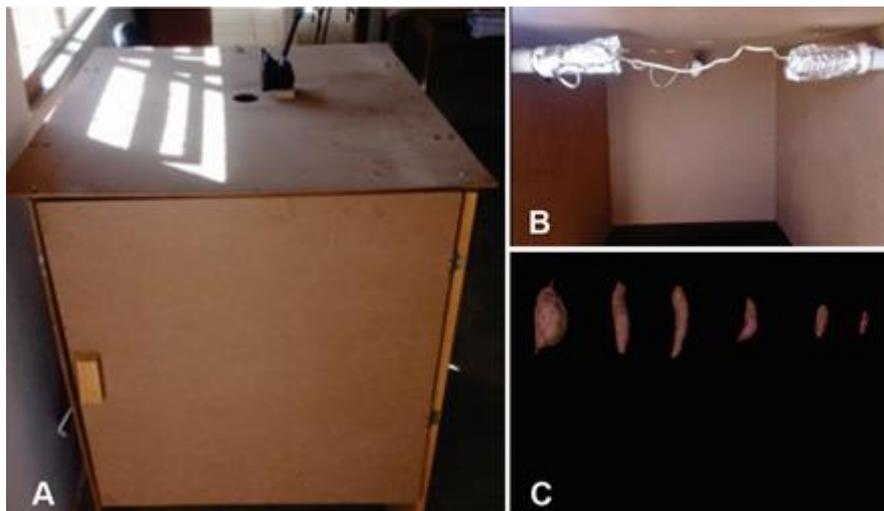
Para manter o solo com bom teor de umidade, inicialmente, a irrigação por aspersão foi aplicada todos os dias. Após o período crítico de estabelecimento da cultura, a irrigação foi aplicada duas vezes por semana.

A colheita manual foi realizada 165 dias após o plantio, e as análises foram realizadas, em nível de planta, nas quais inicialmente foi retirado o excesso de solo das raízes e obtido os seus pesos individuais.

#### 4.1.2.2 Aquisição e processamento de imagens

A aquisição de imagens foi realizada em um “estúdio” feito com mdf com dimensões no fundo de 0,50 x1,00 m e altura de 1,0 m (**Figura 1**). Foi utilizada uma câmera digital modelo Canon PowerShotSX400 IS, sob iluminação artificial, com lâmpada fluorescente. A câmera foi acoplada a um suporte, com o objetivo de padronizar a obtenção das imagens, de forma que todas as imagens foram obtidas da mesma altura (70 cm) e angulação (90°). As raízes foram colocadas sobre um fundo preto e espaçadas entre si, sem sobreposição.

Figura 1- Aquisição de imagens: A) “estúdio” para a aquisição das imagens; B) Iluminação artificial com lâmpada fluorescente; C) Imagem adquirida no estúdio

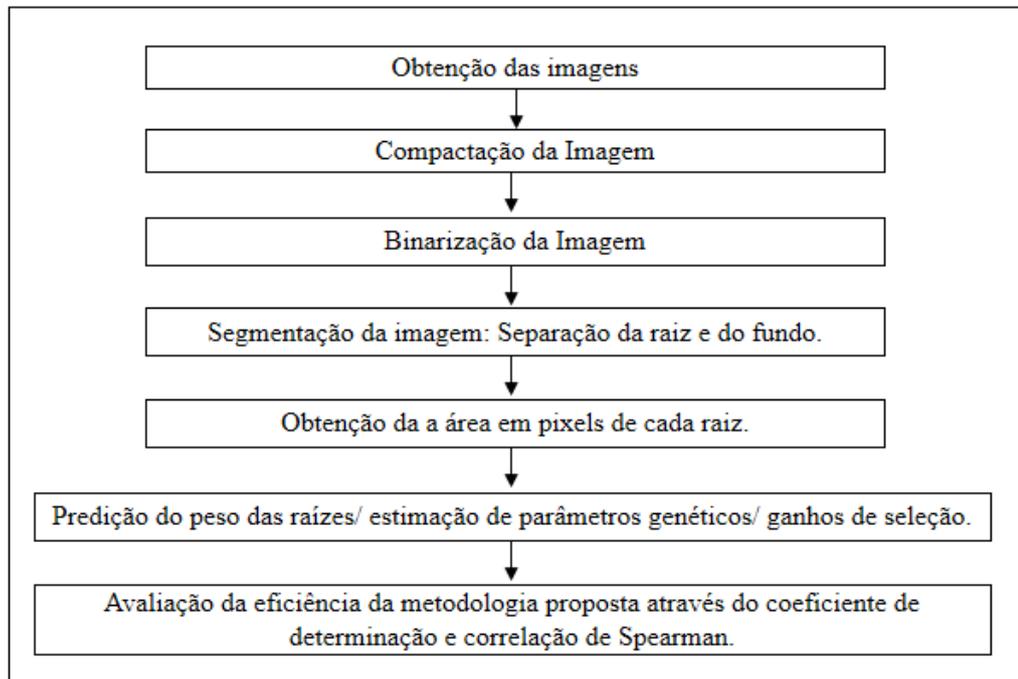


Fonte: Da autora,2022.

A análise das imagens seguiu as etapas apresentadas na **Figura 2**, realizada no *software* R (R CORE TEAM, 2019). As imagens originais (4608 x 3456 pixels) foram redimensionadas, a fim de possibilitar a análise computacional, em menor tempo. Foram utilizadas as funções *read\_image* e *resize\_image* do pacote *ExpImage*. As imagens foram compactadas para 1000 x 750 pixels de largura e altura. Após a

compactação, as imagens foram convertidas para a escala de cinza, por meio da função *gray\_scale*, considerando a banda de azul, obtendo-se uma matriz de 1000 colunas e 750 linhas com valores variando entre 0 e 1. Para a segmentação de cada imagem, a fim de separar a raiz do fundo, foi utilizado um limiar (*threshold*) de 0,35. Assim, todos os pixels com valor abaixo desse limiar foram considerados como fundo (preto), atribuindo-se o valor 0, aos demais pixels, correspondente às raízes, atribuiu-se o valor 1. Após a segmentação, para obter a área em pixels de cada uma das raízes, foi utilizada a função *measure\_image*.

Figura 2- Fluxograma das etapas realizadas a partir da obtenção das imagens



Fonte: Da autora, 2022.

#### 4.1.2.3 Predicão do peso das raízes e estimacão de parâmetros genéticos e ganhos de seleçã

Para a predicão do peso das raízes (em gramas), foi testado o ajuste do modelo de regressão polinomial de primeiro e segundo grau. Para a seleçã do melhor ajuste, considerou-se a significância dos coeficientes de regressão pelo teste t ( $p \leq 0,05$ ). Para isso, considerou-se, como variável resposta, o peso de cada uma das raízes e, como variável explicativa, o número de pixels. A análise de regressão também foi ajustada entre o peso real e predito, em nível de média de parcela e em nível de média de progênie, a fim de verificar a eficiência dessas predicões.

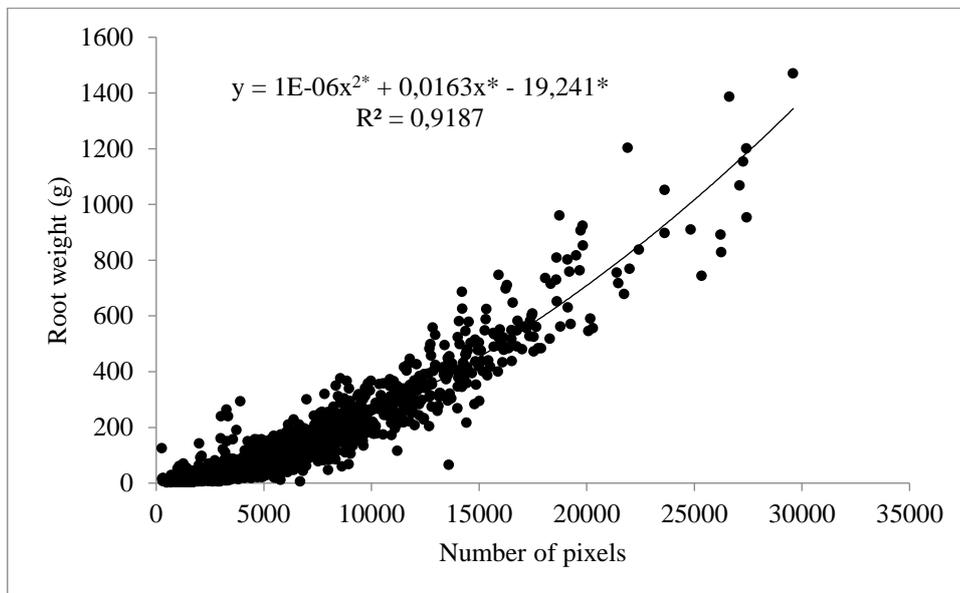
Para a análise genético-estatística dos pesos reais e preditos, foi utilizado o modelo  $\mathbf{y} = \mathbf{Xr} + \mathbf{Za} + \mathbf{Wp} + \mathbf{e}$ , em que:  $\mathbf{y}$  é o vetor de dados,  $\mathbf{y} \sim \mathbf{N}(\mathbf{Xr}, \mathbf{V})$ ,  $\mathbf{V} = \mathbf{ZA}\sigma_a^2\mathbf{Z}' + \mathbf{Wl}\sigma_c^2\mathbf{W}' + \mathbf{l}\sigma_e^2$ ,  $\sigma_a^2$  é a variância genética aditiva,  $\sigma_c^2$  é a variância ambiental entre parcelas e  $\sigma_e^2$  é a variância residual;  $\mathbf{r}$  é o vetor dos efeitos de repetição (assumidos como fixos) somados à média geral;  $\mathbf{a}$  é o vetor dos efeitos genéticos aditivos individuais (aleatórios), sendo  $\mathbf{a} \sim \mathbf{N}(\mathbf{0}, \mathbf{A}\sigma_a^2)$ ;  $\mathbf{p}$  é o vetor dos efeitos de parcelas (aleatórios), em que  $\mathbf{p} \sim \mathbf{N}(\mathbf{0}, \mathbf{l}\sigma_p^2)$ ; e  $\mathbf{e}$  é o vetor de erros (aleatórios), sendo  $\mathbf{e} \sim \mathbf{N}(\mathbf{0}, \mathbf{l}\sigma_e^2)$ . As letras maiúsculas representam as matrizes de incidência para os referidos efeitos.

A partir desse modelo, foram estimados os componentes de variância e preditos os valores genéticos pelo método dos modelos mistos REML-BLUP com o auxílio do *pacote sommer*. Posteriormente, com os valores genéticos para o peso observado e predito, foram estimados os ganhos diretos e indiretos esperados com a seleção das melhores progênes, bem como a correlação de Spearman entre os valores genéticos e a porcentagem de coincidência entre os genótipos selecionados. Para a seleção, foi considerada a intensidade de seleção de 30%.

#### 4.1.3 Resultados

Com o número de pixels correspondente às raízes, foram obtidas as estimativas de peso das raízes em nível de plantas de 16 progênes de meios-irmãos de batata-doce (**Figura 3**), sendo obtido alto coeficiente de determinação ( $R^2=0,9187$ ) entre o peso predito e o número de pixels. Esse resultado indica alta eficiência da metodologia desenvolvida em prever a produtividade de raízes a partir de imagens.

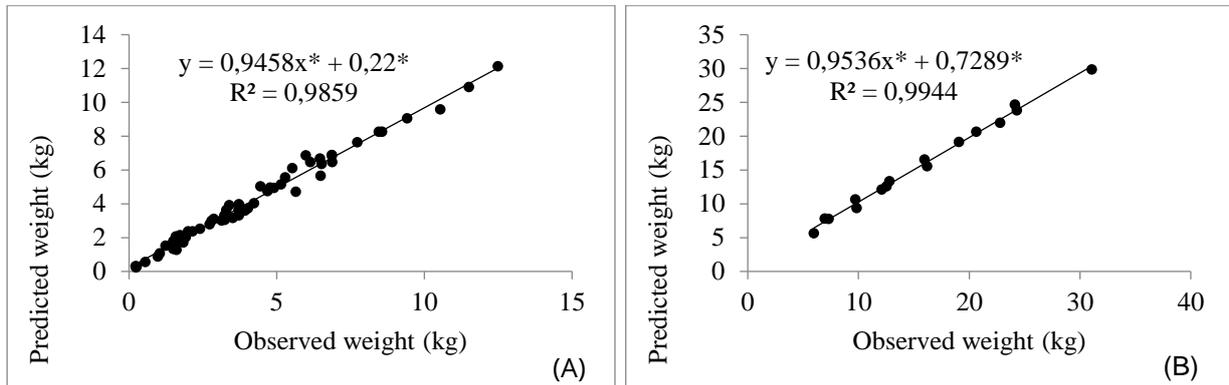
Figura 3- Regressão ajustada na predição do peso da raiz (g) a partir do número de pixels em imagens de 2233 raízes oriundas de 16 progênes de meios-irmãos de batata-doce. UFMG,2022



\* significativo ao nível de 5% pelo teste t. Fonte: Da autora, 2022.

Ao analisar a predição do peso em nível de parcela (**Figura 4-A**) e de família (**Figura 4-B**), foram observados altos coeficientes de determinação entre o peso predito e o peso observado. Em nível de família, foi observada melhor predição no peso, resultando em coeficiente de determinação de 0,9944 (**Figura 4-B**), ligeiramente superior ao obtido para a predição do peso em nível de parcela ( $R^2=0,9859$ ).

Figura 4- Dispersão gráfica entre os valores de peso total preditos e observados, em nível de parcela (A) e de família (B), em 16 progênes de meios-irmãos de batata-doce.



\* significativo ao nível de 5% pelo teste t. Fonte: Da autora, 2022.

Em relação aos parâmetros genéticos, foram observadas estimativas de variância genética aditiva ( $V_a$ ) superiores aos valores de peso observado (Peso obs), para raízes de batata-doce (6,45), em relação ao peso predito (Peso pred), sendo esse comportamento observado também para a variância fenotípica ( $V_f$ ) (7,61) (**Tabela 1**). Além disso, pode-se verificar, por meio dos coeficientes de variação genéticos ( $CV_g$ ), que existe grande variabilidade entre os genótipos. Foram obtidos valores de coeficiente de variação genética aditiva individual ( $CV_{gi}$ ), superiores aos coeficientes de variação residual ( $CV_e$ ) para o peso das raízes de batata-doce, observado em campo e o peso predito pela análise das imagens. Tanto para o peso observado quanto para o peso predito foram obtidos altos valores de herdabilidade ( $h^2_a$ ), para as progênes estudadas, sendo essas similares para as duas características avaliadas.

Tabela 1. Estimativas de parâmetros genéticos no estudo do peso observado e predito em 16 progênes de meios-irmãos de batata-doce. UFMG,2022.

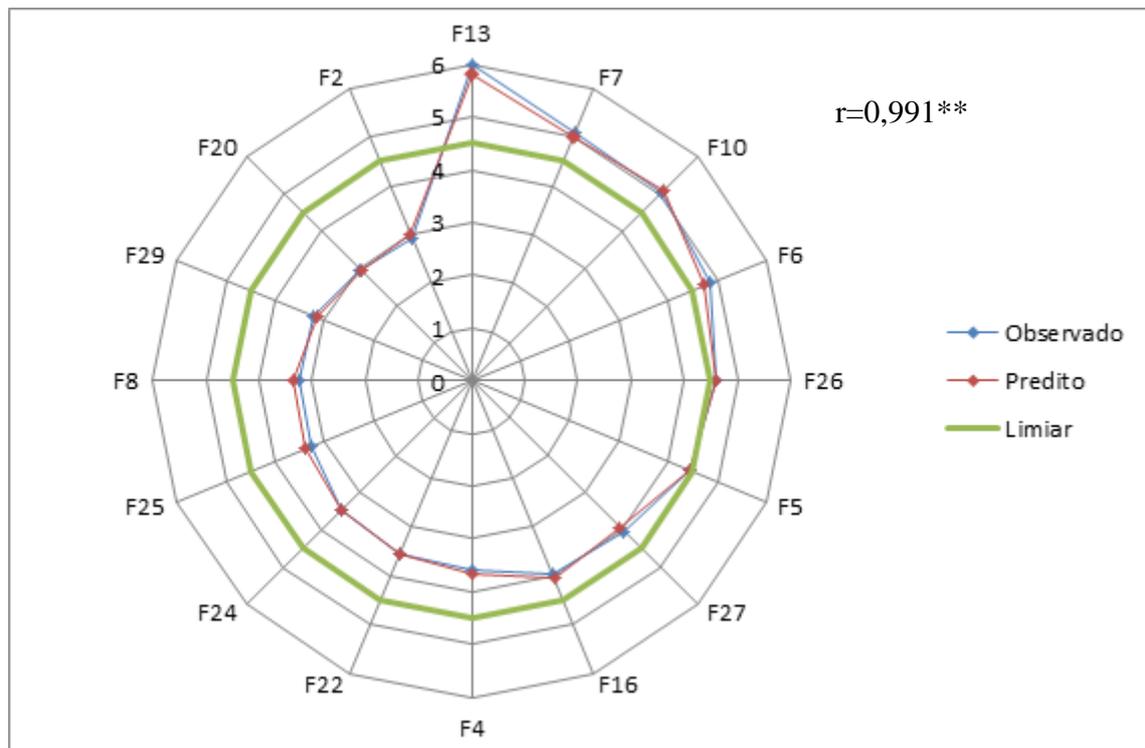
Parâmetros	Peso Obs	Peso pred
$V_a$	6.45	5.82
$V_e$	1.16	1.15
$V_f$	7.61	6.97
$h^2_a$	0.85	0.84
$h^2_{mp}$	0.52	0.51
Acprog	0.72	0.72
$h^2_{ad}$	0.81	0.79
$CV_{gi}\%$	63.31	60.12
$CV_{gp}\%$	31.65	30.06
$CV_e\%$	61.04	58.51
PEV	0.78	0.71
SEP	881.35	841.17

$V_a$ , variância genética aditiva;  $V_e$ , variância ambiental entre as parcelas;  $V_f$ , variância fenotípica;  $h^2_a$ , herdabilidade;  $h^2_{mp}$ , herdabilidade da média de progênes; ACprog, acurácia da seleção de progênes;

$h^2_{ad}$ , herdabilidade aditiva dentro de parcela;  $CV_{gi}$ , Coeficiente de variação genética aditiva individual;  $CV_{gp}$ , Coeficiente de variação genética entre progênies;  $CV_e$ , coeficiente de variação residual;  $PEV$ , = variância do erro de predição dos valores genotípicos de progênie;  $SEP$ , desvio-padrão do valor genotípico predito de progênie. Fonte: Da autora,2022.

A estimativa da correlação de Spearman foi positiva e significativa ( $r=0,991^{**}$ ) entre os valores genéticos considerando os pesos observados e preditos. Além disso, verifica-se a proximidade entre os valores genéticos, considerando os pesos preditos e observados, para todas as progênies (**Figura 5**). Esse resultado demonstra a equivalência da técnica tradicional com a seleção via análise de imagens. As progênies F26 (UFVJM25), F6 (UFVJM40), F10 (UFVJM05), F7 (UFVJM01) e F13 (UFVJM15) devem ser selecionadas, considerando a intensidade de seleção de 30% assim como tanto a produção predita como observada. Ou seja, houve 100% de coincidência entre as progênies selecionadas.

Figura 5- Dispersão em gráfico radar e correlação de Spearman ( $r$ ) dos valores genéticos aditivos de 16 progênies de meios-irmãos de batata-doce. UFMG,2022.



Fonte: Da autora,2022.

As variáveis, peso observado e peso predito apresentaram ganhos de seleção estimado similares, com estimativas em torno de 27,50 e 25,93%, respectivamente (**Tabela 2**). Pode ser observado ganho relativamente superior (1,6%) para o peso observado.

Tabela 2. Ganho de seleção direto (negrito) e indireto em progênes de meios-irmãos de batata-doce. UFMG,2022.

Variáveis selecionadas	Ganho de seleção (%)	
	Peso Obs	Peso Pred
Peso Obs	<b>27.50</b>	25.93
Peso Pred	27.50	<b>25.93</b>

Peso Obs. Peso observado; Peso Pred. Peso Predito. Fonte: Da autora,2022.

#### 4.1.4 Discussão

O grande aumento nas informações genômicas disponíveis na biologia vegetal levou à necessidade de fluxos de trabalho de fenotipagem de alto rendimento, para preencher a lacuna crescente entre genótipo e fenótipo (POUND *et al.*, 2017), uma vez que as ferramentas tradicionais de fenotipagem, geralmente, dependem da medição manual de várias características selecionadas de uma pequena amostra de plantas. Esse fato pode tornar o processo de fenotipagem limitado, retardando o conhecimento de como os fenótipos se correlacionam com fatores genéticos e ambientais, necessitando, assim, de uma análise mais abrangente (UBBENS, STAVNESS, 2017). Assim, o uso de imagens apresenta grande potencial em aumentar o rendimento e a precisão das atividades de fenotipagem em plantas (POUND *et al.*,2017).

Dessa forma, para as progênes de batata-doce estudada, a metodologia desenvolvida conseguiu obter de forma eficiente o peso das raízes com o número de pixels da imagem (**Figura 3**). Essa capacidade de predição do peso pela análise de imagens parte do princípio que uma imagem é composta por um conjunto de pixels que representam uma matriz de valores (x,y) (PELT, SETHIAN, 2018). A partir disso, a área ocupada por uma raiz refere-se a um número específico de pixels, sendo possível predizer o seu peso. Nessa perspectiva, pode-se inferir que quanto maiores as dimensões da raiz maior será o seu peso, corroborando com o observado em campo e refletindo no alto valor do coeficiente de determinação obtido.

Ao comparar o peso predito ao peso observado em campo, para as raízes de batata-doce, tanto em nível de parcelas quanto em nível de família (**Figura 4**), a metodologia desenvolvida apresentou alto desempenho. Obteve-se maior coeficiente de determinação, em nível de família, uma vez que, em nível de parcela, a variação total dos dados é relativamente maior que em nível de progênes, confirmando os coeficientes de determinação obtidos pela análise de imagens.

De forma geral, foram obtidas estimativas de parâmetros genéticos semelhantes, para o peso observado e o peso predito, evidenciando a eficiência da amostragem realizada em campo, bem como da utilização de imagens para predizer a produtividade de batata-doce. Esse fato pode ser confirmado pelos baixos valores de variância do erro de predição observados, para as duas características, bem como pela boa acurácia obtida para os atributos. Ao comparar os valores genéticos, para os pesos preditos e observados, percebe-se que os valores preditos são mais favoráveis à seleção (AZEVEDO *et al.*, 2017). Para pesos preditos, foi observada menor variância do erro de predição dos valores genotípicos de progênie (PEV), bem como menor desvio-padrão do valor genotípico predito (SEP), para cada progênie, se comparado ao peso observado (**Tabela 1**). Essa menor variação aos pesos preditos se dá pela

tendência que a metodologia segue de amenizar os efeitos de falsos outliers (erros de tabulação) no processo de predição.

Para as características avaliadas, as variâncias genóticas ( $V_a$ ) foram superiores à variância ambiental ( $V_e$ ), demonstrando que as progênes estudadas podem apresentar viabilidade para o melhoramento (BADU *et al.*, 2017; NARASIMHAMURTHY *et al.*, 2018). Isso é reforçado pela superioridade do coeficiente de variação genotípico em relação ao coeficiente de variação residual (**Tabela 1**). Além dos coeficientes de variância, informações de herdabilidade ( $h^2_a$ ), também, são de suma importância para a seleção eficiente de genitores (RAHAJENG *et al.*, 2020). Uma vez que a herdabilidade é a porção da variância genética em relação à variância fenotípica, esse parâmetro é utilizado para estimar a confiabilidade do valor fenotípico, como indicador do valor reprodutivo (RAMALHO *et al.*, 2012). Assim, altas herdabilidades, conforme obtidas para as progênes de batata-doce estudadas, indicam que o fenótipo é mais influenciado por fatores genéticos que ambientais (RAHAJENG *et al.*, 2020). Assim, a seleção se torna mais confiável associando-se altas herdabilidades ao  $CV_g$ , uma vez que indica que o caractere analisado apresenta maior efeito aditivo.

A equivalência da técnica tradicional com a seleção via análise de imagens comprovou a alta performance da metodologia para prever os pesos das raízes. Isso pode ser confirmado com a alta correlação obtida entre pesos preditos e observados (**Figura 5**). Além disso, a técnica permitiu identificar as progênes que obtiveram valores aditivos superior ao esperado para o conjunto de dados estudados.

Além da possibilidade de identificar progênes superiores, a metodologia foi eficaz, também, de acordo com os ganhos de seleção diretos e indiretos, para as duas características analisadas. Como as progênes selecionadas entre as duas metodologias foram as mesmas, os ganhos de seleção foram idênticos. Esse fato representa uma vantagem aos programas de melhoramento genético da cultura, uma vez que o seu objetivo principal é a obtenção de raízes dentro do peso adequado para fins comerciais. Portanto pode-se inferir que a metodologia desenvolvida possui alto potencial, para o uso na fenotipagem de alta eficiência pela análise de imagens, para a predição do peso de raízes de batata-doce, o que pode possibilitar o desenvolvimento de equipamentos, para a extração de informações em campo, no ato da colheita, além de armazenar a imagem, em um banco de dados, para consultas e comparações futuras, permitindo processar várias imagens, simultaneamente, além de fenotipar quanto a características qualitativas (HAQUE *et al.*, 2021) e estimar a produtividades de raízes.

#### 4.1.5 Conclusões

A visão computacional permitiu a predição do peso das raízes, mantendo o ranqueamento dos genótipos e, conseqüentemente, a similaridade entre os ganhos esperados com a seleção.

O uso de imagens é uma ferramenta eficiente aos programas de melhoramento genético da batata-doce, auxiliando no processo de fenotipagem da cultura para estimar a produtividade total de raízes.

#### 4.1.6 Agradecimentos

Agradecemos a CAPES (Conselho de Aperfeiçoamento de Pessoal de Nível Superior – código 001), à FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais) e ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) pelo apoio a este estudo.

#### 4.1.7 Referências

AMABILE, R. F.; VILELA, M. S.; PEIXOTO, J. R. Melhoramento de plantas: variabilidade genética, ferramentas e mercado. **Embrapa Cerrados-Livro técnico (INFOTECA-E)**, 2018.

AZEVEDO, A. M.; ANDRADE JÚNIOR, V. C. DE; PEDROSA, C. E.; DE OLIVEIRA, C. M.; DORNAS, M. F.; CRUZ, C. D.; VALADARES, N. R. Application of artificial neural networks in indirect selection: a case study on the breeding of lettuce. **Bragantia**, v.74, n.4, p.387-393. 2015b. DOI: <https://dx.doi.org/10.1590/1678-4499.0088>

AZEVEDO, A. M.; ANDRADE JÚNIOR, V. C.; SOUSA JÚNIOR, A. S.; SANTOS, A. A.; CRUZ, C. D.; PEREIRA, S. L.; OLIVEIRA, A. J. Eficiência da estimação da área foliar de couve por meio de redes neurais artificiais. **Horticultura Brasileira**, v. 35, n. 1, p. 14-19, 2017.

AZEVEDO, A.M.; ANDRADE JÚNIOR, V.C.; FERNANDES, J.S.C.; PEDROSA, C.E.; OLIVEIRA, C.M. Desempenho agrônomo e parâmetros genéticos em genótipos de batata-doce. **Horticultura Brasileira**. v. 33, p.084-090, 2015a. DOI: <http://dx.doi.org/10.1590/S0102-053620150000100014>

BADU, M.; ASHOK, P.; KIRANPATRO, T.S.K.K.; SASIKALA, K. Studies on Genetic Variability, Heritability and Genetic Advance for Growth, Yield and Quality Parameters among Orange Flesh Sweet Potato [*Ipomoea batatas* (L.) Lam.] Genotypes. **Int.J.Curr.Microbiol.App.Sci**, v.6, n.9, p.1894-1903, 2017.

BAEK, J.H.; LEE, E.; KIM, N.; KIM, S.L.; KIM, K.H. High Throughput Phenotyping for Various Traits on Soybean Seeds Using Image Analysis. **Sensors**, v. 20, n. 1, p. 248, 2020.

BORGES, V; FERREIRA, PV; SOARES, L; SANTOS, GM; SANTOS, AMM. Seleção de clones de batata-doce pelo procedimento REML/BLUP. **Acta Scientiarum. Agronomy**, v.32, p.643-649, 2010.

FANG, Y.; XIONG, L. General mechanisms of drought response and their application in drought resistance improvement in plants. **Cellular and Molecular Life Sciences**, v.72, p.673–689, 2015.

FILGUEIRA, F.A.R. **Novo manual de olericultura: agrotecnologia moderna na produção e comercialização de hortaliças**. Viçosa: UFV, 402p. 2008.

HAQUE, S.; LOBATON, E.; NELSON, N.; YENCHO, G. C.; PECOTA, K. V.; MIEROP, R.; KUDENOV, M. W.; BOYETTE, M.; WILLIAMS, C. M. Computer vision approach to characterize size and shape phenotypes of horticultural crops using high-throughput imagery. **Computers and Electronics in Agriculture**, v.182, 2021.

HUANG, S.; WEIGEL, D.; BEACHY, R.N.; LI, J. A proposed regulatory framework for genome-edited crops. **Nature Genetics**. V.48, n.2, 2016.

MAHLEIN, A. K. Plant disease detection by imaging sensors—parallels and specific demands for precision agriculture and plant phenotyping. **Plant disease**, v. 100, n. 2, p. 241-251, 2016.

NARASIMHAMURTHY, P.N.; PATEL, N.B.; PATEL, A.I.; KOTESWARA, RAO G. Genetic variability, heritability and genetic advance for growth, yield and quality parameters among sweet potato [*Ipomoea batatas* (L.) lam.] genotypes. **International Journal of Chemical Studies**, v.6, n.4, p.2410-2413, 2018.

OTOBONI, M. E. F.; OLIVEIRA, D. J. L. S. F. D.; VARGAS, P. F.; PAVAN, B. E.; ANDRADE, M. I. Genetic parameters and gain from selection in sweet potato genotypes with high beta-carotene content. **Crop Breeding and Applied Biotechnology**, v. 20, n. 3, 2020.

PELT, D. M.; SETHIAN, J. A. A mixed-scale dense convolutional neural network for image analysis. **Proceedings of the National Academy of Sciences**, v. 115, n. 2, p. 254-259, 2018.

POUND, M. P.; ATKINSON, J. A.; TOWNSEND, A. J.; WILSON, M. H.; GRIFFITHS, M.; JACKSON, A. S.; BULAT, A.; TZIMIROPOULOS, G.; WELLS, D.M.; MURCHIE, E.H.; PRIDMORE, T.P.; FRENCH, A. P. Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. **Gigascience**, v. 6, n. 10, p. gix083, 2017.

R CORE TEAM, 2019. R: A language and environment for statistical computing. **R Foundation for Statistical Computing, Vienna, Austria**. Disponível em: <http://www.r-project.org/index.html>.

RAHAJENG, W., RESTUONO, J., INDRIANI, F. C., PURWONO, P. Genetic Parameters of Agronomic Traits in Sweetpotato Accessions. **Biosaintifika: Journal of Biology & Biology Education**, v. 12, n. 2, p. 240-246, 2020.

RAMALHO, M.A.P.; ABREU, A.F.B.; SANTOS, J.B.; NUNES, J.A.R. **Aplicações da genética quantitativa no melhoramento de plantas autógamas**. UFLA, Lavras, 522p, 2012.

RESENDE, M.D.V de. Software Selegen-Reml / Blup. Colombo: **Embrapa Florestas**. 2002.

RINCKER, K.; NELSON, R.; SPECHT, J.; SLEPER, D.; CARY, T.; CIANZIO, S.R.; CASTEEL, S.; CONLEY, S.; CHEN, P.; DAVIS, V.; FOX, C.; GRAEF, G.; GODSEY, C.; HOLSHOUSER, D.; JIANG, G.L.; KANTARTZI, S.K.; KENWORTHY, W.; LEE, C.; MIAN, R.; MCHALE, L.; NAEVE, S.; ORF, J.; POYSA, V.; SCHAPPAUGH, W.; SHANNON, G.; UNIATOWSKI, R.; WANG, D.; DIERS, B. Genetic Improvement of U.S. Soybean in Maturity Groups II, III, and IV. **Crop Science**. V.54, p.1419–1432, 2014.

UBBENS, J. R.; STAVNESS, I. Deep plant phenomics: a deep learning platform for complex plant phenotyping tasks. **Frontiers in plant science**, v. 8, p. 1190, 2017.

VARGAS, P.F.; GODOY, D.R.Z.; ALMEIDA, L.C.F.; CASTOLDI, R. Agronomic characterization of sweet potato accessions. **Comunicata Scientiae**, v. 8, n. 1, p. 116-125, 2017.

ZHOU, N.; SIEGEL, Z.D.; ZARECOR, S.; LEE, N.; CAMPBELL, D.A.; ANDORF, C.M.; FRIEDBERG, I. Crowdsourcing image analysis for plant phenomics to generate ground truth data for machine learning. **PLoS computational biology**, v.14, n.7, e1006337, 2018.

## 4.2 CAPÍTULO II: Diversidade genética da cor da polpa de batata-doce por visão computacional e random Forest

Este artigo foi elaborado conforme normas do Programa de Pós-Graduação em Produção Vegetal.

### Resumo:

A adoção de novas tecnologias de fenotipagem de plantas é uma importante estratégia, para o melhoramento genético da batata-doce, possibilitando a avaliação de características de interesse em menor espaço de tempo e mão de obra, além de possibilitar a classificação de caracteres qualitativos, reduzindo a subjetividade que existe nas notas dadas pelos avaliadores. Nesse sentido, objetivou-se verificar a eficiência da análise computacional de imagens na avaliação da coloração de polpas de raízes de batata-doce e estudar a dissimilaridade genética existente. Foram avaliadas 16 progênies de meios-irmãos de batata-doce, em delineamento em blocos casualizados com quatro repetições. Foram realizados cortes de aproximadamente 1cm de espessura, em duas raízes de cada planta e fotografadas em um estúdio sob iluminação artificial. Para a classificação das polpas de batata-doce por análise de imagens, foram feitas paletas de cores para o treinamento de Florestas randômicas (*random forest*). Após o treinamento, foram obtidos valores referentes à classificação quanto à cor de cada um dos pixels correspondentes às raízes, a partir dos quais foram estimados parâmetros genéticos e ganhos de seleção. Foi verificada eficiência na classificação das polpas de raízes de batata-doce por meio da análise computacional de imagens. E verificou-se variabilidade genética entre os genótipos estudados quanto à coloração da polpa das raízes, em que os genótipos UFVJM-31 e UFVJM-40 apresentaram grande dissimilaridade se comparado aos demais.

**Palavras-Chave:** *Ipomoea batatas* (L.) Lam.. Fenotipagem de alta eficiência. Análise de imagens.

### 4.2.1 Introdução

Com o aumento populacional e as drásticas mudanças climáticas, faz-se necessária a melhoria das culturas para proporcionar a segurança alimentar e nutricional (VARSHNEY *et al.*, 2020). Nesse sentido, o melhoramento genético de plantas é uma estratégia importante na agricultura moderna, pois maximiza caracteres quantitativos e qualitativos, além de produzir indivíduos mais resistentes a estresses bióticos e abióticos (AMABILE, VILELA, PEIXOTO, 2018).

Considerar caracteres qualitativos é muito importante para o melhoramento genético. Na cultura da batata-doce, por exemplo, a coloração da polpa influencia diretamente no produto a ser comercializado, pois polpas de cores mais roxas e mais alaranjadas, além de serem preferidas por alguns consumidores, estão associadas a uma alta atividade antioxidante, possibilitando benefícios à saúde (TEOW *et al.*, 2007). Além disso, pode influenciar na finalidade da raiz, a qual pode ser destinada para consumo *in natura* ou como matéria-prima, para processos industriais, produzindo doces, farinhas, papel, cosméticos, amido e etanol (CARDOSO *et al.*, 2005; FRANCO, FRANCO, 2021).

A avaliação da cor da polpa, em raízes de batata-doce, além de ser importante como critério de seleção de genótipos, é importante também em estudos de dissimilaridade. A cultura possui mecanismo de autoincompatibilidade que conduz à polinização cruzada, resultando em grande variabilidade genotípica

e fenotípica e, conseqüentemente, na grande variabilidade de cor, sobretudo, da polpa das raízes (OTOBONI *et al.*, 2020). Essa variabilidade pode ser estudada por técnicas de análise multivariada que permitem analisar simultaneamente variáveis de interesse, a fim de caracterizar e selecionar genitores para a obtenção de alta variabilidade genética na população segregante (STRECK *et al.*, 2017, ANDRADE *et al.*, 2017).

O melhoramento genético para atributos de cor apresenta dificuldades, uma vez que a análise visual é influenciada pela subjetividade humana, além de dificultar a análise estatística paramétrica dos dados. Além disso, há grande variação das cores, ao longo da polpa, impossibilitando a classificação quanto a uma cor específica. Nesse sentido, faz-se necessária a utilização de ferramentas alternativas de fenotipagem de alta eficiência. Entre elas, pode-se destacar o uso de imagens associada à visão computacional, já que permite avaliar atributos qualitativos e de produtividade com menor custo, menor mão de obra, maior rapidez, eficiência e precisão (DE MESQUITA PINHEIRO *et al.*, 2022).

Algoritmos baseados na visão computacional, geralmente, fornecem recursos discriminativos associados a saídas extraídas, por meio de seu processo de treinamento, o que permite dissecar características complexas e determinar assinaturas visuais relacionadas a características em plantas (MOCHIDA *et al.*, 2019). Entre esses algoritmos, pode-se destacar a floresta randômica, que é um modelo de aprendizado de máquina que pode ser utilizado para desenvolver modelos de previsão em muitos ambientes de pesquisa. Além disso, as florestas randômicas permitem diminuir o número de variáveis necessárias para obter uma classificação, a fim de reduzir a carga de coleta de dados e melhorar a eficiência (SPEISER *et al.*, 2019). Nesse sentido, objetivou-se verificar a eficiência da análise computacional de imagens na avaliação da coloração de polpas de raízes de progênies de meios irmãos de batata-doce e estudar a dissimilaridade genética existente.

## **4.2.2 Material e Métodos**

### **4.2.2.1 Instalação e avaliação do experimento**

O experimento foi desenvolvido no Instituto de Ciências Agrárias da UFMG- Campus Montes Claros-MG (coordenadas: 16°40'58.16"S e 43°50'20.15"O) onde foram avaliadas 16 famílias de meios-irmãos de batata-doce (BELGARD (F2), CAMBRAIA (F4), LICURI (F5), UFVJM40 (F6), UFVJM01 (F7), ARRUBA (F8), UFVJM05 (F10), UFVJM15 (F13), UFVJM56 (F16), UFVJM31 (F20), UFVJM37 (F22), UFVJM54 (F24), UFVJM25 (F26), UFVJM29 (F27), TCARRO02 (F29), UFVJM09 (F25)), em cambissolo háplico sob condições de irrigação.

Os meios-irmãos de batata-doce foram obtidos com a coleta de sementes, oriundas do banco de germoplasma, composto por acessos elite trazidos da Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM) e cultivados no ICA/UFMG. As sementes foram coletadas diariamente entre os meses de abril e outubro de 2018 e acondicionados em geladeira a 4°C. Posteriormente, as sementes foram submetidas à escarificação mecânica com lixa para quebra de dormência (impermeabilidade tegumentar) e plantadas em bandejas de poliestireno de 72 células com substrato comercial. As bandejas foram mantidas em casa de vegetação e irrigadas diariamente, durante 2 meses, quando as mudas se encontraram prontas para o plantio.

O plantio foi realizado em leiras, em delineamento de blocos ao acaso (DBC) com 16 tratamentos (diferentes progênies) e quatro repetições, com leiras espaçadas entre si de 1m e espaçamento entre plantas de 0,4 m, totalizando 10 plantas por parcela. Como as avaliações foram feitas em nível de planta, foi utilizado um espaçamento maior a fim de facilitar a identificação de cada planta e facilitar a colheita. Os tratos culturais foram realizados como o recomendado para a cultura no Novo Manual de Olericultura (FILGUEIRA, 2008). Já, para a adubação, foram utilizados 180 kg ha<sup>-1</sup> de fósforo e 30 kg ha<sup>-1</sup> de nitrogênio. Aos 30 dias do plantio das mudas, foi realizada uma adubação de cobertura com 30 kg ha<sup>-1</sup> de nitrogênio. A adubação de potássio não foi necessária de acordo com a análise química do solo.

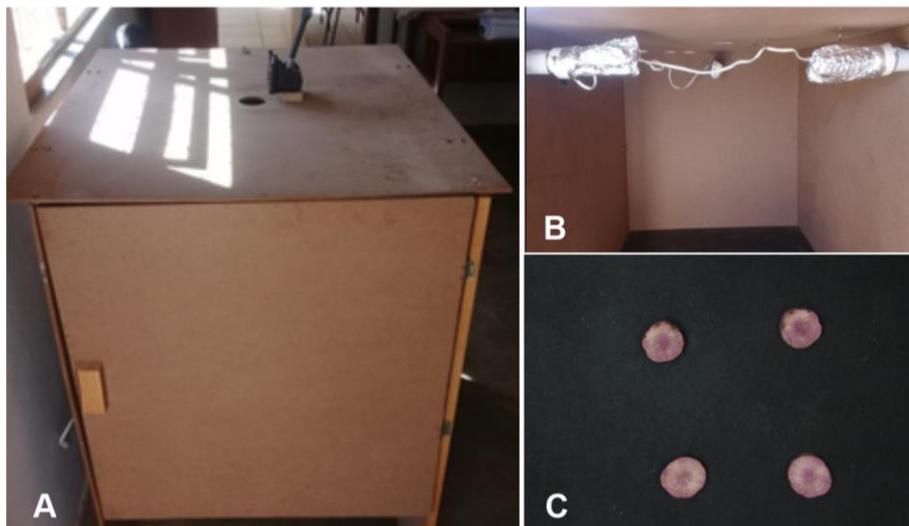
Para manter o solo com bom teor de umidade, inicialmente, a irrigação por aspersão foi aplicada todos os dias. Após o período crítico de estabelecimento da cultura, a irrigação foi aplicada duas vezes por semana.

A colheita manual foi realizada 165 dias após o plantio, e as análises foram realizadas, em nível de planta, em que, inicialmente, foi retirado o excesso de solo das raízes. Para a obtenção das imagens das polpas das raízes de batata-doce, foram realizados cortes de aproximadamente 1cm de espessura em duas raízes de cada planta e imediatamente foram obtidas as imagens.

#### 4.2.2.2 Aquisição e processamento de imagens

A aquisição de imagens foi realizada em um “estúdio” de mdf, com dimensões do fundo de 0,50 x1,00 m e altura de 1,0 m (**Figura 1**). Foi utilizada uma câmera digital modelo Canon PowerShotSX400 IS, sob iluminação artificial, com lâmpada fluorescente. A câmera foi acoplada a um suporte, com o objetivo de padronizar a obtenção das imagens, de forma que todas as imagens foram obtidas da mesma altura (70 cm) e angulação (90°). Os segmentos das raízes foram colocados sobre um fundo preto, espaçados entre si, sem sobreposição.

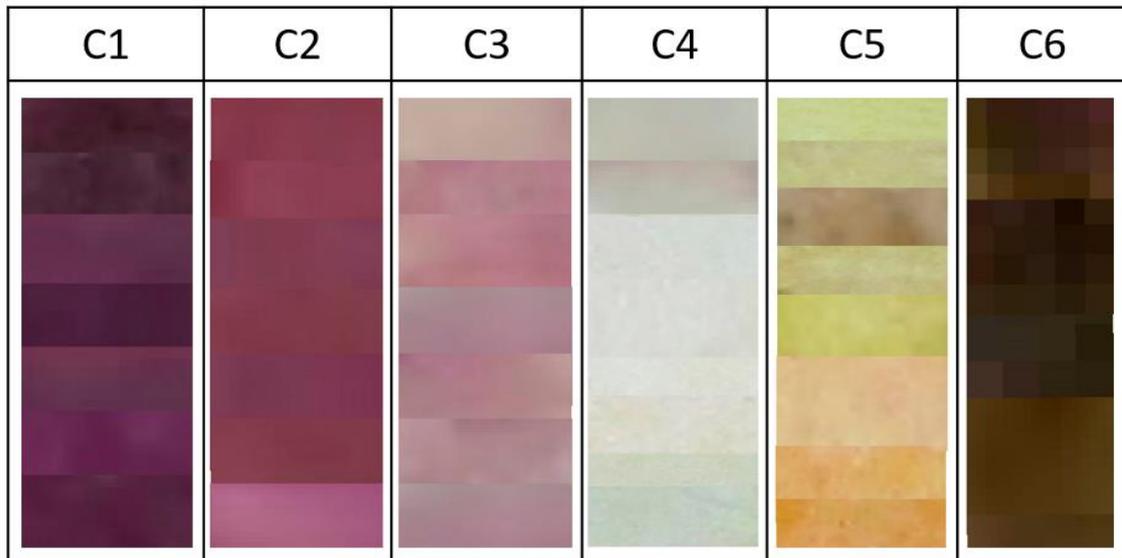
Figura 1- Aquisição de imagens: A) “estúdio” para aquisição das imagens; B) Iluminação artificial com lâmpada fluorescente; C) Imagem adquirida no estúdio.



Fonte: Da autora,2022.

Após a aquisição das imagens, foram feitas paletas de cores, para a classificação das polpas de batata-doce (**Figura 2**), as quais foram divididas em seis classes de cores (C1-C6). Essas paletas englobaram todas as variações de cores de polpa observadas, nas progênies estudadas, sendo, portanto bem representativas neste experimento. Essas paletas foram feitas no software GIMP e salvas em formato jpeg.

Figura 2- Paletas para a classificação das polpas de batata-doce quanto à cor

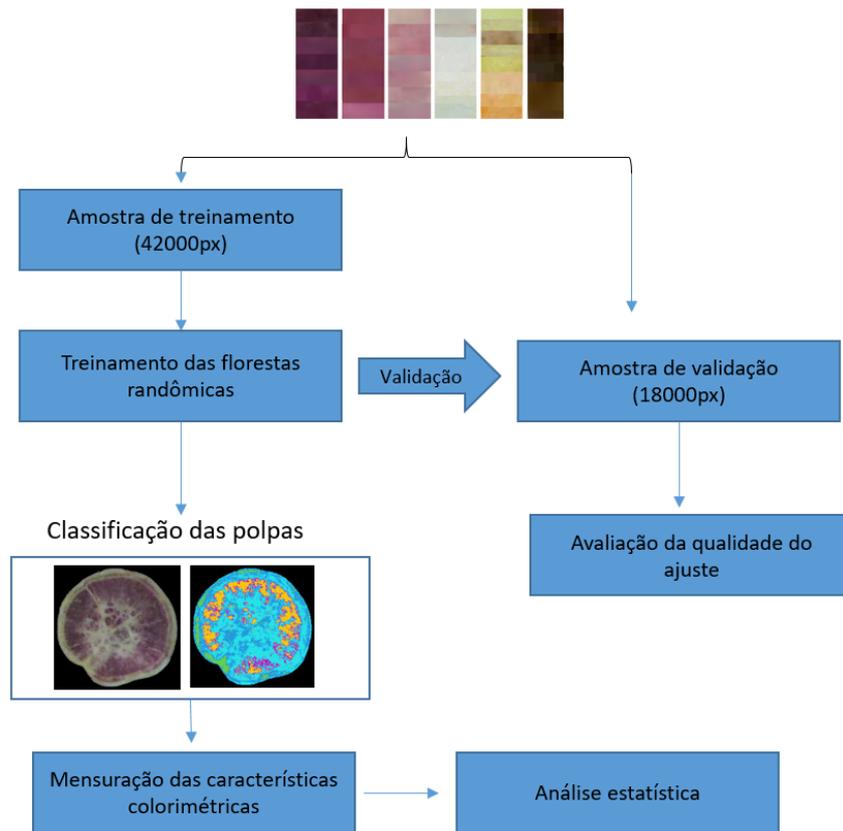


C1= Polpa Roxa escura; C2= Polpa Roxa clara; C3= polpas rosadas; C4= polpas brancas; C5= polpas amareladas; C6= Polpas marrons (danos). Fonte: Da autora,2022.

Cada paleta de cores foi feita com 200 pixels de largura e 500 pixels de altura, totalizando 100.000 pixels. A partir desses pixels, para cada paleta de cor, 42000 foram selecionados, aleatoriamente, com o auxílio do pacote *ExplImage* (AZEVEDO,2022), para compor a amostra de treinamento e 18000 para compor a amostra de validação (**Figura 3**). Para a classificação dos pixels, foi utilizada a metodologia de random forest, com o auxílio do pacote "*randomForest*" do R (LIAW, WIENER, 2022), utilizando sua configuração padrão. Para isso, consideraram-se, para cada pixel, os valores de R, G e B como variável de entrada e um número de 1 a 6 como variável de saída.

Após a etapa de treinamento, a eficiência do ajuste da *random Forest* foi verificada a partir da amostra de validação. Para isso, foram estimados a **Acurácia**: $(VP+VN)/(VP+VN+FP+FN)$ ; **F2**:  $2*(Precisão*Recall)/(Precisão + Recall)$ ; **Precisão**:  $VP/(VP+FP)$ ; **Recall**:  $VP/(VP+FN)$ ; e, **Especificidade**:  $VN/(FP+VN)$ . Em que VP= verdadeiro positivo; FN= falso negativo; FP= falso positivo e VN= verdadeiro negativo.

Figura 3: Fluxograma da metodologia de avaliação da coloração de polpas de raízes de batata-doce



Fonte: Da autora,2022.

Após o treinamento das florestas randômicas, para cada raiz foi obtida a porcentagem de pixels nas raízes com cada uma das seis colorações. Esses valores, estimados, em nível de planta, foram utilizados para o estudo de dissimilaridade genética. Para isso, utilizou-se o pacote `MultivariateAnalysis` (AZEVEDO,2021) do software R, a partir do qual foi estimada a matriz de dissimilaridade de Mahalanobis, feito o dendrograma e o estudo de variáveis canônicas.

#### 4.2.3 Resultados

Ao avaliar o ajuste da metodologia de florestas randômicas, na classificação de pixels, para as seis colorações, a partir da amostra de validação, foi feita uma tabela de confusão (**Tabela 1**), na qual se verificou que a maior parte dos pixels foram classificados corretamente.

Tabela 1: Tabela de confusão para a classificação de pixels correspondentes a diferentes tons na paleta de cores. UFMG,2022

Classificação real	Classificação predita					
	1	2	3	4	5	6
1	3025	3	0	0	0	0
2	1	2950	2	0	0	0
3	0	0	2991	1	0	0
4	0	0	0	3015	0	0
5	0	0	0	3	2970	0
6	0	0	0	0	7	3032

Fonte: Da autora, 2022.

Quanto aos avaliadores da qualidade do ajuste, foram obtidos excelentes resultados com estimativas superiores a 99% para a acurácia, F2 precisão, recall, especificidade e acurácia geral (**Tabela 2**).

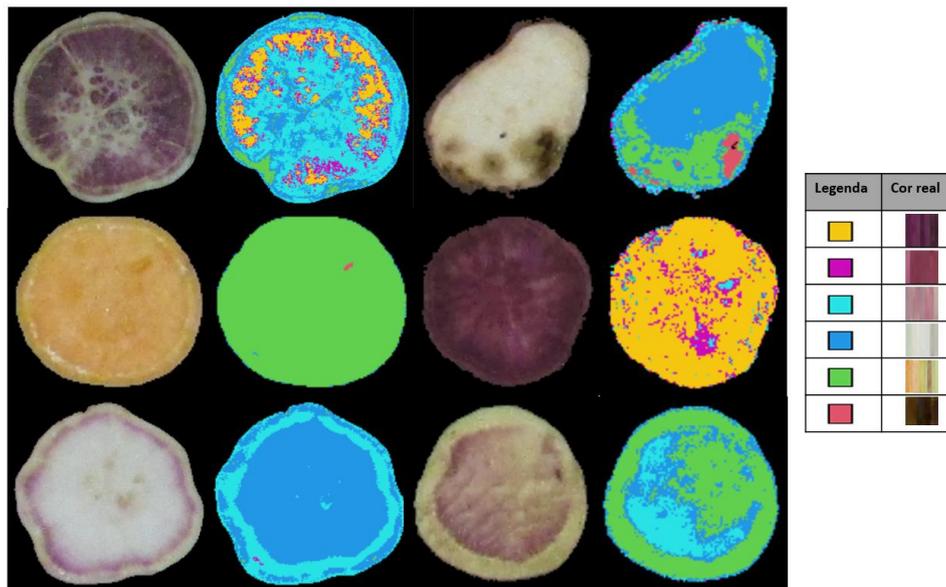
Tabela 2: Avaliadores da classificação do ajuste da metodologia de florestas randômicas, na classificação de pixels, em seis colorações de polpas de batata-doce. UFMG,2022.

Classe	Acurácia	F2	Precisão	Recall	Especificidade	Acurácia geral
1	0.999	0.999	1.000	0.999	1.000	0.999
2	0.999	0.999	0.999	0.999	1.000	
3	1.000	1.000	0.999	1.000	1.000	
4	1.000	1.000	0.999	1.000	1.000	
5	0.999	0.999	0.998	0.999	1.000	
6	0.999	0.998	1.000	0.998	1.000	

Fonte: Da autora,2022.

A eficiência na classificação dos pixels quanto à coloração em polpas de batata-doce pode ser visualizada na **Figura 4**, sendo possível observar a atribuição de cores pertencentes aos pixels de cada classe. Dessa forma, foi atribuída a coloração azul para pixels mais claros, verde para pixels amarelados e alaranjados, amarelo e rosa para pixels associados às classes C1, C2 e C6, ou seja, aos pixels mais escuros.

Figura 4. Demonstração da classificação de pixels quanto à coloração em polpas de batata-doce

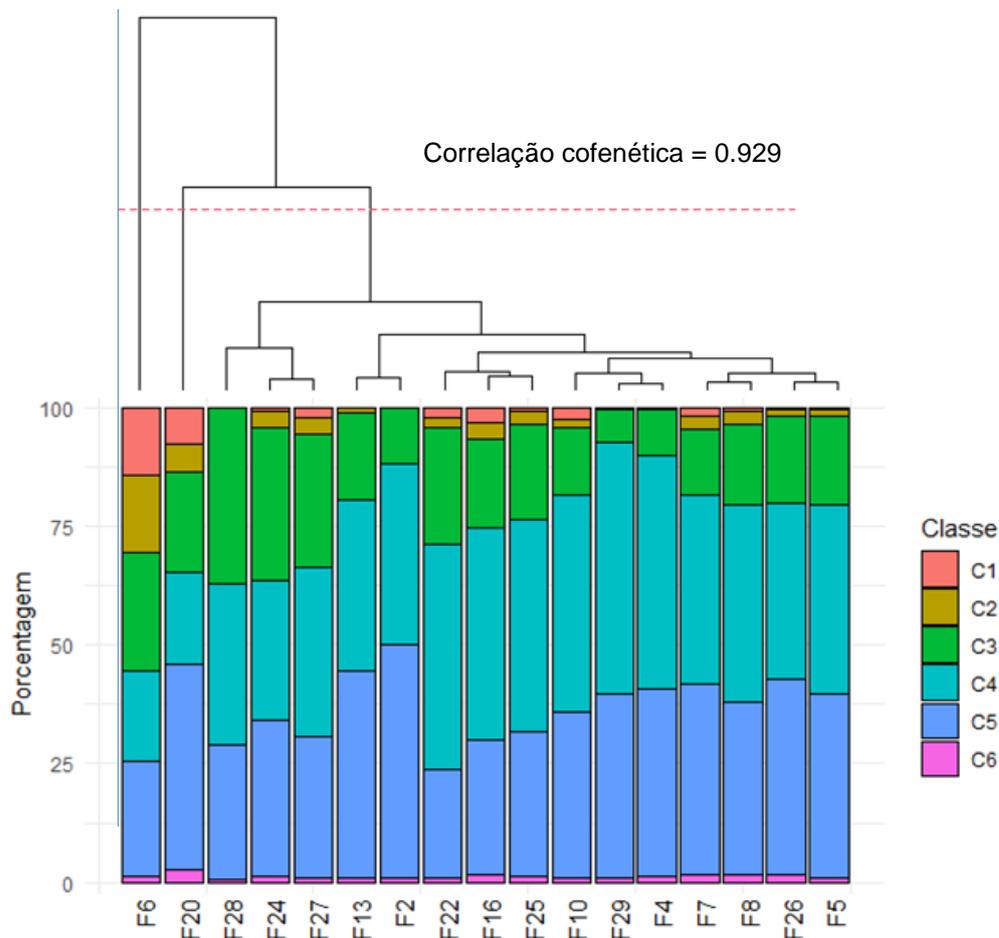


Fonte: Da autora, 2022.

Foi possível observar alta correlação cofenética (0,929) no dendrograma indicando alta eficiência em representar a matriz de dissimilaridade da população estudada (**Figura 5**). Assim, o dendrograma foi dividido em três grupos de dissimilaridade pelo critério de Mojena, sendo os dois primeiros compostos pelas progênies F20 e F6, enquanto as demais progênies foram agrupadas no grupo 3.

Nas progênies agrupadas nos grupos I e II (F20 (UFVJM-31) e F6 (UFVJM-40)), foi possível observar maiores proporções de pixels, nas classes de coloração C1 e C2, ou seja, são formados por raízes de coloração roxa e/ou bicolor. Já o grupo III é formado por progênies com maiores proporções de pixels mais claros pertencentes às classes 3, 4 e 5. É possível observar maiores proporções de pixels brancos, para as progênies F22 (UFVJM-37), F25 (UFVJM-09) e F4 (CAMBRAIA), que possuem polpa de coloração branca e/ou creme. Já as progênies F2 (BELGARD), F13 (UFVJM-15) e F7 (UFVJM-01) apresentam maiores proporções de pixels amarelados e representam as raízes de polpas amarelas e laranjadas (**Figura 5**).

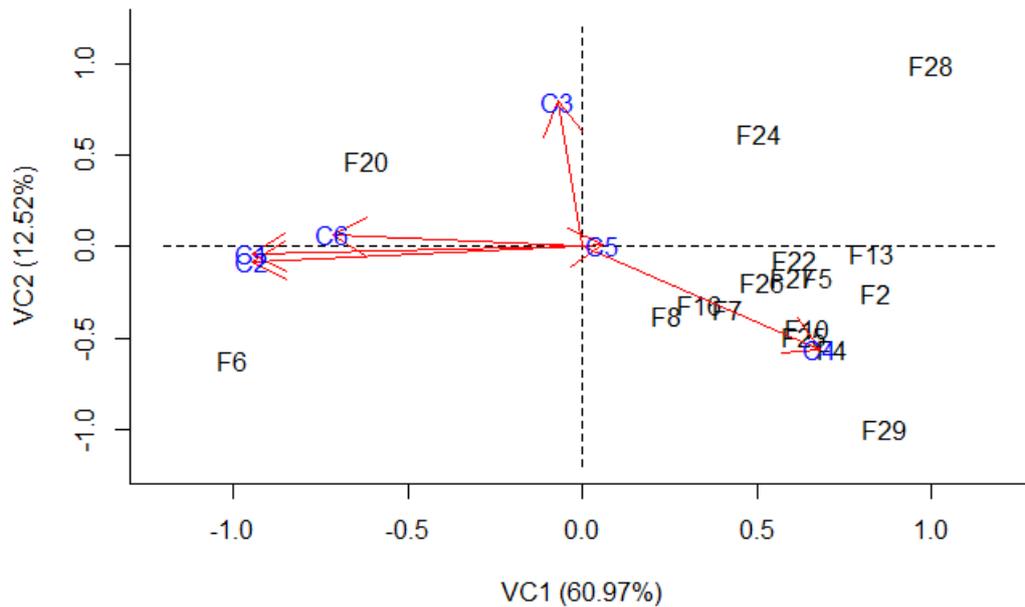
Figura 5. Dendrograma obtido pelo método UPGMA, para progênies de meios-irmãos de batata-doce, em função da proporção de pixels com cada uma das seis classificações de cor



Fonte: Da autora,2022.

Ao avaliar a dispersão das duas primeiras variáveis canônicas (VC1 e VC2), no estudo da coloração de raízes de progênies de meios-irmãos de batata-doce, verificou-se que 60,97% da variação foi explicada pela variável canônica 1 (VC1), e 12,57% foram explicados pela VC2. Além disso, verificou-se que as progênies F20 (UFVJM-31) e F6 (UFVJM-40) apresentam maiores proporções de pixels, nas classes de coloração mais escura (C1,C2 e C6), enquanto na família F24 (UFVJM 54) é possível observar a predominância de pixels da classe 3 (C3) (**Figura 6**). Isso indica que as plantas das progênies F20 e F6 apresentam polpa de coloração predominantemente roxa. Já as demais progênies apresentam polpas em tons amarelados e esbranquiçados, com a predominância de pixels mais claros.

Figura 6. Dispersão das duas primeiras variáveis canônicas no estudo da coloração de raízes de batata-doce



C1= Polpa Roxa escura; C2= Polpa Roxa clara; C3= polpas rosadas; C4= polpas brancas; C5= polpas amareladas; C6= Polpas marrons (danos). Fonte: Da autora,2022.

#### 4.2.4 Discussão

O melhoramento de plantas é de suma importância para a agricultura moderna. Neste contexto, a avaliação de caracteres quantitativos e qualitativos, simultaneamente, influenciam diretamente na escolha dos melhores genitores. Contudo métricas qualitativas (por exemplo, cor da polpa) são avaliadas em menores proporções e de forma subjetiva (HAQUE *et al.*, 2021), sendo mais difíceis de caracterizar manualmente em larga escala. Nesse sentido, a adoção de tecnologias para a fenotipagem de alto rendimento torna-se uma vertente (POUND *et al.*, 2017). Assim, no presente estudo, foi possível classificar polpas de raízes batata-doce pela análise de imagens e demonstrar sua funcionalidade no melhoramento.

Para as progênies de batata-doce estudadas, a metodologia de classificação de imagens desenvolvida obteve alta acurácia, uma vez que proporcionou a classificação das polpas das raízes de batata-doce por meio do número de pixels das imagens correspondentes a cada tom (**Tabela 1**). Essa eficiência na classificação de pixels, para as seis colorações de polpa analisadas, foi observada na acurácia geral da metodologia de florestas randômicas utilizadas (99,9%), uma vez que a metodologia apresentou alta especificidade e precisão para todas as classes analisadas (**Tabela 2**). Dessa forma, as florestas randômicas apresentam a capacidade de utilizar informações agregadas dos dados de treinamento para construir modelos que permitam a classificação dos caracteres de interesse (SPEISER *et al.*, 2019).

Para tornar possível a classificação da metodologia desenvolvida, foram atribuídas pseudocores para cada classe avaliada (**Figura 3**). Esse processo de classificação torna-se possível uma vez que cada pixel da imagem tem associado a si um valor de R, G e B. Esses valores referentes a cada pixel associam-se a cada classe (cor), pois apresentam uma matriz de valores diferentes, relacionados à intensidade da

cor e, conseqüentemente, uma resposta RGB diferente (PELT, SETHIAN, 2018). Possibilita a classificação de polpas de raízes de batata-doce quanto à sua coloração pela abordagem de visão computacional.

Ao analisar a dissimilaridade da população estudada, obteve-se correlação cofenética alta e positiva (0,929) (**Figura 4**), indicando confiabilidade dos resultados ao interpretar o dendrograma (CRUZ; REGAZZI; CARNEIRO, 2012). Uma vez que o coeficiente de correlação cofenética estima a representatividade dos dados da matriz de dissimilaridade no dendrograma, indica se os dados da matriz estão bem ajustados (MANTEL, 1967). O estudo de dissimilaridade é de suma importância ao melhoramento genético, uma vez que indica a variabilidade genética desejável para as características analisadas (CRUZ; REGAZZI; CARNEIRO, 2012). Para as progênies estudadas, foi possível obter três grupos de dissimilaridade, cujos grupos I e II foram formados pelas progênies F20 (UFVJM-31) e F6(UFVJM-40), compostos por raízes de coloração roxa ou bicolor (classes C1, C2 e C6), enquanto o grupo III agrupou progênies com raízes de polpas mais claras (classes C3, C4 e C5) (**Figura 4 e Figura 5**).

A coloração da polpa das raízes de batata-doce é influenciada diretamente pelos teores de carotenoides presentes nas raízes, sendo pigmentos naturais responsáveis pela coloração amarela, laranja ou roxa das raízes (SILVA, 2019). Nesse sentido, a escolha dos melhores genitores, para o melhoramento genético, com base nas características de cor, torna-se importante, uma vez que a polpa colorida está associada a uma alta atividade antioxidante e vem sendo considerada como uma escolha alimentar saudável e como fonte potencial de corantes naturais (TEOW *et al.*, 2007). As variedades de batata-doce de polpa alaranjada contêm quantidades consideráveis de  $\beta$ -caroteno, podendo ser uma alternativa para elevar os níveis de vitamina A (LAURIE *et al.*, 2013). Destaca-se a importância de considerar tanto as características de produtividade, como as características qualitativas simultaneamente.

Dessa forma, pode-se inferir que a metodologia desenvolvida apresentou eficiência na classificação da cor de polpas de batata-doce pela análise de imagens, permitindo a análise de caracteres qualitativos, para o melhoramento da cultura, reduzindo a subjetividade dos métodos tradicionais e permitindo a automação da tomada de decisão. A partir disso, muitas avaliações fenotípicas podem ser realizadas em um menor espaço de tempo (HAQUE *et al.*, 2021). Logo a associação entre análises de imagens, para características quantitativas e qualitativas, permite aos melhoristas avaliar a fenotipagem de forma mais eficaz, auxiliando na seleção de novas cultivares em menor espaço de tempo (HAQUE *et al.*, 2021).

#### **4.2.5 Conclusões**

A metodologia desenvolvida apresentou eficiência na classificação das polpas de raízes de batata-doce por meio da análise de imagens.

Há variabilidade genética entre os genótipos estudados quanto à coloração da polpa das raízes, em que os genótipos UFVJM-31 e UFVJM-40 apresentaram grande dissimilaridade se comparados aos demais.

A visão computacional é uma ferramenta eficiente, na análise de dados qualitativos, permitindo a redução da subjetividade humana e possibilitando a fenotipagem de uma grande quantidade de informações, podendo ser útil para diferentes culturas vegetais.

#### 4.2.6 Agradecimentos

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior- Brasil (CAPES) – Código de Financiamento 001, do Conselho Nacional de Desenvolvimento Científico (CNPQ) e a Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG).

#### 4.2.7 REFERÊNCIAS

AMABILE, R. F.; VILELA, M. S.; PEIXOTO, J. R. Melhoramento de plantas: variabilidade genética, ferramentas e mercado. **Embrapa Cerrados-Livro técnico (INFOTECA-E)**, 2018.

ANDRADE, E. K. V.; ANDRADE JÚNIOR, V. C.; LAIA, M. L.; FERNANDES, J. S. C.; OLIVEIRA, A. J. M.; AZEVEDO, A. M. Genetic dissimilarity among sweet potato genotypes using morphological and molecular descriptors. **Acta Scientiarum. Agronomy**, v. 39, n. 4, p. 447-455, 2017.

AZEVEDO, Alcinei Místico. 2021. "MultivariateAnalysis: Pacote para análise multivariada. Disponível em: < <https://cran.r-project.org/web/packages/MultivariateAnalysis/index.html>>.

AZEVEDO, Alcinei Místico. 2022. "ExplImage: Ferramenta para análise de imagens em experimentos. Disponível em: < <https://cran.r-project.org/web/packages/ExplImage/index.html>>.

BORGES, V.; FERREIRA, P.V.; SOARES, L.; SANTOS, G.M.; SANTOS, A.M.M. Seleção de clones de batata-doce pelo procedimento REML/BLUP. **Acta Scientiarum. Agronomy**, v.32, p.643-649, 2010.

CARDOSO, A.D.; VIANA, A.E.S.; RAMOS, P.A.S.; MATSUMOTO, S.N.; AMARAL, C.L.F.; SEDIYAMA, T.; MORAIS, O.M. Avaliação de clones de batata-doce em Vitória da Conquista. **Horticultura Brasileira**, v.23, p.911-914, 2005.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: Editora UFV. 2012, 514p.

DE MESQUITA PINHEIRO, R.; GADOTTI, G. I.; BERNARDY, R.; MONTEIRO, R. D. C. M.; MOREIRA, I. B. Processamento de imagens como ferramenta importante para inteligência artificial no setor de sementes. Image processing as to important tool for artificial intelligence in the seed sector. **Revista Agrária Acadêmica**, v.5, n.1, 2022.

FILGUEIRA, FAR. **Novo manual de olericultura: agrotecnologia moderna na produção e comercialização de hortaliças**. Viçosa: UFV, 2008,402p.

FRANCO, A. C.; FRANCO, L. S. Tecnologias na produção de etanol de batata-doce: estudo de processos entre Brasil e China. **Revista Brasileira de Tecnologia Agroindustrial**, v.15, n.2, 2021.

HAQUE, S.; LOBATON, E.; NELSON, N.; YENCHO, G.C.; PECOTA, K.V.; MIEROP, R.; KUDENOV, M. W.; BOYETTE, M.; WILLIAMS, C.M. Computer vision approach to characterize size and shape phenotypes of horticultural crops using high-throughput imagery. **Computers and Electronics in Agriculture**, v.182, 2021.

LAURIE, S. M.; CALITZ, F. J.; ADEBOLA, P. O.; LEZAR, A. Characterization and evaluation of South African sweet potato (*Ipomoea batatas* (L.) LAM) land races. **South African Journal of Botany**, v. 85, p. 10-16, 2013.

LIAW, A.; WIENER, M. "Classification and Regression by randomForest." **R News**, v.2, n.3, p.18-22, 2022.

MANTEL, N. The detection of disease clustering and a generalized regression approach. **Cancer Research**, v.27, n.2, p. 209-220, 1967.

MOCHIDA, K.; KODA, S.; INOUE, K.; HIRAYAMA, T.; TANAKA, S.; NISHII, R.; MELGANI, F. Computer vision-based phenotyping for improvement of plant productivity: a machine learning perspective. **GigaScience**, v.8, n.1, 2019.

OTOBONI, M.E.F.; OLIVEIRA, D.J.L.S.F.D.; VARGAS, P.F.; PAVAN, B.E.; ANDRADE, M.I. Genetic parameters and gain from selection in sweet-potato genotypes with high beta-carotene content. **Crop Breeding and Applied Biotechnology**, v.20, n.3, 2020.

PELT, D.M.; SETHIAN, J.A. A mixed-scale dense convolutional neural network for image analysis. **Proceedings of the National Academy of Sciences**, v.115, n.2, p.254-259, 2018.

POUND, M. P.; ATKINSON, J. A.; TOWNSEND, A. J.; WILSON, M. H.; GRIFFITHS, M.; JACKSON, A. S.; BULAT, A.; TZIMIROPOULOS, G.; WELLS, D.M.; MURCHIE, E.H.; PRIDMORE, T.P.; FRENCH, A. P. Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. **Gigascience**, v. 6, n. 10, p. gix083, 2017.

RAHAJENG, W.; RESTUONO, J.; INDRIANI, F.C.; PURWONO, P. 2020. Genetic Parameters of Agronomic Traits in Sweet-potato Accessions. **Biosaintifika: Journal of Biology & Biology Education**, v.12, n.2, p.240-246, 2020.

R CORE TEAM. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria, 2019.

SILVA, V. N. S. Época de colheita como modulador de compostos bioativos em cvs. de batata-doce minimamente processada. **Dissertação**. Universidade Federal Rural do Semi-Árido. Mossoró-RN. 2019,82f.

SPEISER, J. L.; MILLER, M. E.; TOOZE, J.; IP, E. A comparison of random forest variable selection methods for classification prediction modeling. **Expert systems with applications**, v.134, p.93-101, 2019.

STRECK, E. A.; AGUIAR, G. A.; MAGALHÃES JÚNIOR, A. M.; FACCHINELLO, P. H. K.; OLIVEIRA, A. C. Variabilidade fenotípica de genótipos de arroz irrigado via análise multivariada. **Revista Ciência Agronômica**, v. 48, n. 1, p. 101-109, 2017

TEOW, C.C.; TRUONG, V.D.;MCFEETERS, R.F.; THOMPSON, R.L.; PECOT, K.V.; YENCHO, G.C. Antioxidant activities, phenolic and  $\beta$ -carotene contents of sweetpotato genotypes with varying flesh colors **Food Chem**, v.103, p.829-838, 2007.

VARSHNEY, R. K., SINHA, P., SINGH, V. K., KUMAR, A., ZHANG, Q., & BENNETZEN, J. L. 5Gs for crop genetic improvement. **Current Opinion in Plant Biology**, v.56, p.190-196, 2020.

### 4.3 CAPÍTULO III: Redes neurais convolucionais no melhoramento qualitativo de raízes de batata-doce

Este artigo foi elaborado conforme normas do Programa de Pós-Graduação em Produção Vegetal.

#### Resumo

Objetivou-se verificar se redes neurais convolucionais podem ser úteis para a fenotipagem da batata-doce para caracteres qualitativos. Foram avaliadas 16 famílias de meios-irmãos de batata-doce no delineamento em blocos casualizados com quatro repetições. As imagens foram adquiridas em nível de planta, e o pacote *Explmage* do software R foi utilizado para a redução da resolução e individualização de uma raiz por imagem. As imagens foram agrupadas, de acordo com suas classificações quanto ao formato, cor da casca e danos causados por insetos. Foram destinadas ao treinamento das redes 600 raízes de cada classe, já as demais raízes foram utilizadas para verificar a qualidade do ajuste. Foi utilizada a linguagem python na plataforma do Google Colab. Utilizou-se a biblioteca Keras considerando as arquiteturas VGG-16, Inception-v3, ResNet-50, InceptionResNetV2 e EfficientNetB3. A arquitetura InceptionResNetV2 se destacou com alta precisão, ao classificar os indivíduos quanto ao formato, dano por inseto e coloração da casca. A análise de imagens associada ao aprendizado profundo pode ser utilizada no desenvolvimento de aplicativos, a fim de ser utilizado pelo produtor rural, bem como no melhoramento da batata-doce, reduzindo a subjetividade, mão de obra, tempo e recursos financeiros na fenotipagem.

**Palavras-chave:** *Ipomoea batatas* (L.) Lam.. Análise de imagens. Aprendizado profundo. Fenotipagem.

#### 4.3.1 Introdução

A batata-doce (*Ipomoea batatas* (L.) LAM) está entre as olerícolas mais consumidas no Brasil por ser uma rica fonte de carboidratos, fibras, vitaminas, minerais, antioxidantes e pelo baixo conteúdo glicêmico (RAHAJENG *et al.*, 2020). Além dos múltiplos usos como produção de bioetanol (WANG *et al.*, 2020), a batata-doce é uma hortaliça de papel importante no abastecimento de matérias-primas para a alimentação humana e animal (WERA *et al.*, 2014; PRADHAN *et al.*, 2015).

O principal desafio à comercialização da hortaliça encontra-se associado ao valor de mercado, já que esse é fortemente dependente das características qualitativas das raízes, como, por exemplo, o formato (HAQUE *et al.*, 2020). Dessa forma, os consumidores preferem os produtos que tenham um padrão comercial adequado e boa aparência. Assim, os produtos deformados, apesar de terem o mesmo valor nutricional que as raízes comerciais, são frequentemente descartados pelos consumidores (HAQUE *et al.*, 2020), tornando fonte de desperdício de alimentos.

Com o objetivo de aumentar a produtividade e qualidade do produto e amenizar o prejuízo do produtor rural, faz-se necessário investir em informações e novas tecnologias associadas ao melhoramento genético da hortaliça para fins comerciais (RAHAJENG *et al.*, 2020). Porém, na seleção dos melhores genótipos, é necessário avaliar muitas características quantitativas e qualitativas, sendo um processo oneroso e subjetivo, dificultando a análise do melhorista (SAMBASIVAM, OPIYO, 2021). Nesse sentido, a adoção de novas tecnologias associadas ao processo de fenotipagem representa um avanço, e, entre as possibilidades, tem-se a análise de imagens associada à inteligência computacional.

O uso de estratégias que permitem a aquisição e análise de dados de ambientes agrícolas pode ajudar a otimizar as práticas atuais, promovendo aumento de produtividade, melhores processos de controle de qualidade, flexibilização da gestão agrícola (VASCONEZ *et al.*, 2020). Além disso, essas novas tecnologias, que visam aumentar a acurácia e rapidez das mensurações fenotípicas, vêm sendo alvo de intensas pesquisas, nos últimos anos (MASSRUHÁ *et al.*, 2014), como no desenvolvimento de plataformas de fenotipagem (UBBENS, STAVNESS, 2017), sistemas automatizados de fenotipagem de alta eficiência (LEE *et al.*, 2018), bem como na caracterização de fenótipos (HAQUE *et al.*, 2021) em batata-doce usando imagens. E, para automatizar a interpretação dessas imagens, podem-se recorrer às redes neurais convolucionais (CNN's).

As redes neurais convolucionais (CNN) tornaram-se populares para a detecção de objetos, porque, além de conseguir classificá-los, também conseguem extrair descritores da imagem (SHAHEEN *et al.*, 2016). Além disso, podem atingir altos desempenhos, para diferentes problemas de classificação e detecção, alcançando tempos de inferência mais rápidos e taxas de detecção mais altas que os métodos tradicionais de visão computacional (LECUN *et al.*, 2015).

Dessa forma, o uso de imagens associadas às redes neurais de aprendizado profundo pode ser uma alternativa para a classificação eficiente das raízes em comerciais e não comerciais. Com isso, objetivou-se verificar se redes neurais convolucionais podem ser úteis para a fenotipagem da batata-doce para caracteres qualitativos.

#### **4.3.2 Material e Métodos**

##### **4.3.2.1 Instalação e avaliação do experimento**

O experimento foi desenvolvido no Instituto de Ciências Agrárias da UFMG- Campus Montes Claros-MG (coordenadas: 16°40'58.16"S e 43°50'20.15"O), em que foram avaliadas 16 famílias de batata-doce (BELGARD (F2), CAMBRAIA (F4), LICURI (F5), UFVJM40 (F6), UFVJM01 (F7), ARRUBA (F8), UFVJM05 (F10), UFVJM15 (F13), UFVJM56 (F16), UFVJM31 (F20), UFVJM37 (F22), UFVJM54 (F24), UFVJM25 (F26), UFVJM29 (F27), TCARRO02 (F29), UFVJM09 (F25)), em cambissolo háplico sob condições de irrigação.

Os meios-irmãos de batata-doce foram obtidos, por meio da coleta de sementes, oriundas do banco de germoplasma composto por acessos elite trazidos da UFVJM e cultivado no ICA/UFMG. As sementes foram coletadas diariamente entre os meses de abril e outubro de 2018 e acondicionadas em geladeira a 4°C. Posteriormente, as sementes foram submetidas à escarificação mecânica com lixa para quebra de dormência (impermeabilidade tegumentar) e plantadas em bandejas de poliestireno de 72 células com substrato comercial. As bandejas foram mantidas em casa de vegetação e irrigadas diariamente, durante dois meses, quando as mudas se encontraram prontas para o plantio.

O plantio foi realizado em leiras, em delineamento de blocos ao acaso (DBC), com 16 famílias (diferentes progênies) e quatro repetições, com leiras espaçadas entre si de 1m e espaçamento entre plantas de 0,4 m. Como as avaliações foram feitas, em nível de planta, foi utilizado um espaçamento maior, a fim de facilitar a identificação de cada planta e facilitar a colheita. As adubações e tratos culturais foram realizados como o recomendado para a cultura no Novo Manual de Olericultura (FILGUEIRA, 2008). Foram utilizados 180 kg ha<sup>-1</sup> de fósforo e 30 kg ha<sup>-1</sup> de nitrogênio. Aos 30 dias do plantio das mudas, foi realizada

uma adubação de cobertura com 30 kg ha<sup>-1</sup> de nitrogênio. A adubação de potássio não foi necessária de acordo com a análise química do solo.

Para manter o solo com bom teor de umidade, inicialmente, a irrigação por aspersão foi aplicada todos os dias. Após o período crítico de estabelecimento da cultura (dois meses após o transplante), a irrigação foi aplicada duas vezes por semana.

A colheita manual foi realizada 165 dias após o plantio, e as análises foram realizadas, em nível de planta, nas quais, inicialmente, foi retirado o excesso de solo das raízes para posterior obtenção das imagens. Foram analisadas as variáveis formato, cor da casca e danos causados por insetos. As avaliações foram realizadas, de acordo com os descritores e escalas de pontuação recomendados pelo Conselho Internacional de Recursos Genéticos Vegetais (IBPGR), elaborados por Huamán (1991) (**Tabela 1**).

Tabela 1. Escala de notas relacionada ao formato, cor predominante da casca e danos causados por insetos para progênies de meios-irmãos de batata-doce (*Ipomoea batatas* (L.) Lam).

Descritor	Notas
Formato	1- redondo; 2- elíptico arredondado; 3- elíptico; 4- oval; 5- oval invertido; 6- oblongo; 7- longo e oblongo; 8- longo e elíptico; 9- irregular (Adaptado)
Cor Predominante da casca	1- branco; 2- creme; 3- amarelo; 4- laranja; 5- marrom alaranjado 6- rosado; 7- avermelhado; 8- roxo avermelhado; 9- roxo escuro.
Danos causados por insetos	0-Ausente; 1-presente.

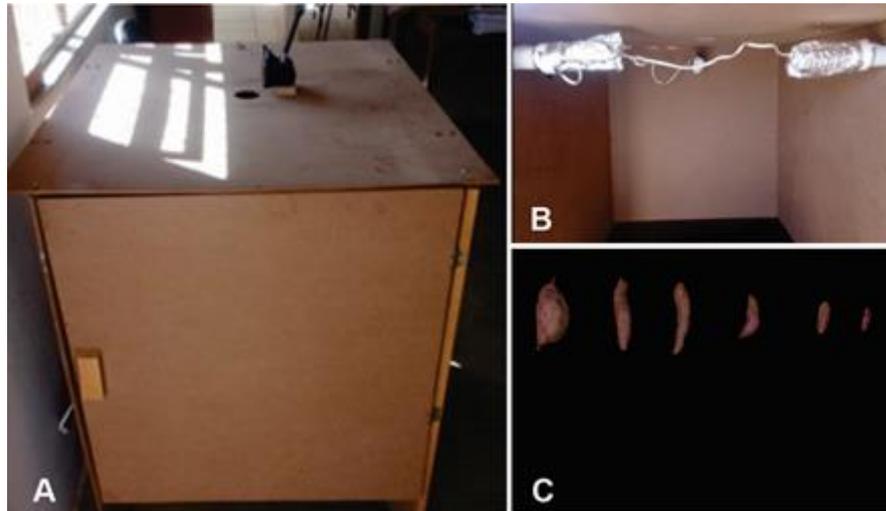
Fonte: Adaptado de Huamán (1991).

Para o formato as raízes foram classificadas em comercializáveis e não comercializáveis, sendo consideradas comercializáveis as que possuíam formato mais fusiforme. Já para a cor da casca as raízes foram divididas em raízes de coloração clara (branco, creme, amarelo, laranja) e raízes de coloração escura (rosado, avermelhado, roxo avermelhado, roxo escuro). Quanto aos danos causados por insetos as raízes foram classificadas quanto a presença e ausência de danos.

#### 4.3.2.2 Aquisição e processamento de imagens

A aquisição de imagens foi realizada em um “estúdio” feito com MDF com dimensões no fundo de 0,50 x 1,00 m e altura de 1,0 m (**Figura 1**). Foi utilizada uma câmera digital modelo Canon PowerShotSX400 IS, sob iluminação artificial, com lâmpada fluorescente. A câmera foi acoplada a um suporte, com o objetivo de padronizar a obtenção das imagens, de forma que todas as imagens foram obtidas da mesma altura (70 cm) e angulação (90°). As raízes foram colocadas sobre um fundo preto e espaçadas entre si, sem sobreposição.

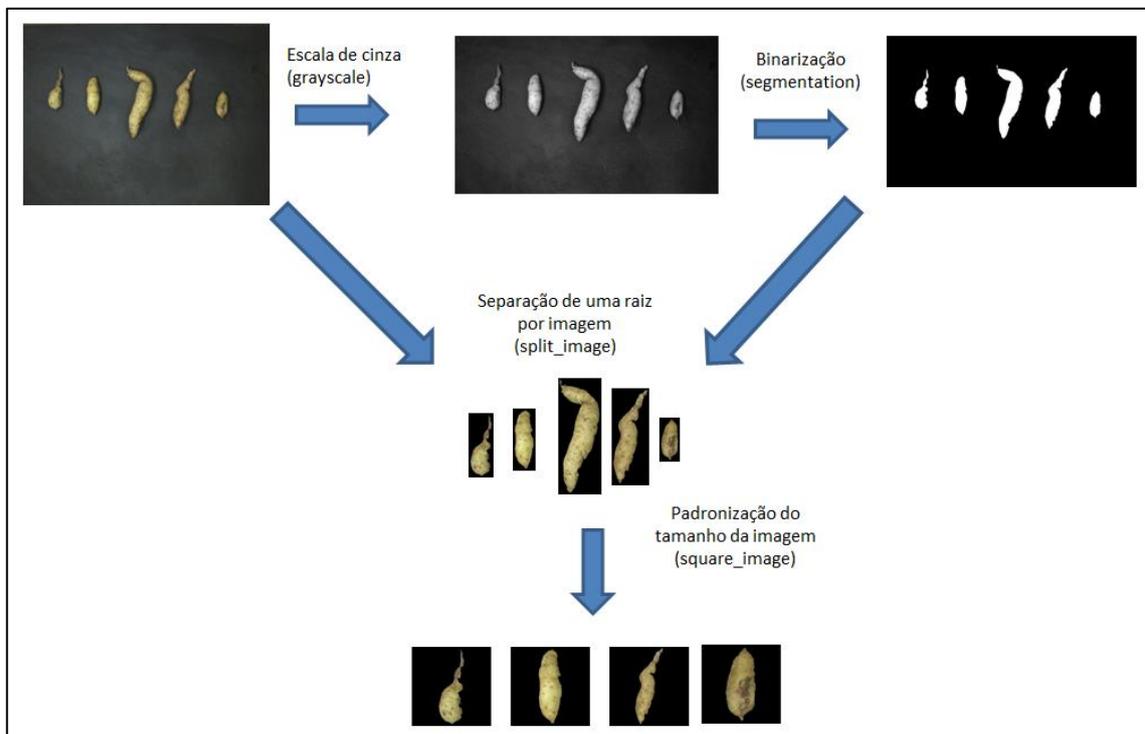
Figura 1 - Aquisição de imagens: A) “Estúdio” para aquisição das imagens; B) Iluminação artificial com lâmpada fluorescente; C) Imagem adquirida no estúdio.



Fonte: Da autora,2022.

O pacote ExplImage do software R foi utilizado para a redução da resolução e individualização de uma raiz por imagem (**Figura 2**). Essas imagens foram agrupadas de acordo com suas classificações quanto ao formato (comercializável e não comercializável), cor da casca (clara e escura) e ataque de insetos (danificada e não danificada).

Figura 2- Etapas para a individualização de uma raiz por imagem na fenotipagem de raízes de batata-doce por análise computacional de imagens



Fonte: Da autora,2022.

As imagens de raízes de batata-doce foram classificadas visualmente, considerando as variáveis cor da casca das raízes (raízes claras e raízes escuras), formato (comercializáveis e não comercializáveis) e danos causados por insetos (danificada e não danificadas) (**Tabela 2**). Assim, o conjunto de imagens originais para cada variável foram divididas em imagens destinadas para o ajuste das RCNNs e para a avaliação da qualidade do ajuste (Teste). Dessa forma, foram destinadas ao treinamento das redes 600 raízes de cada uma dessas classificações, já o restante foi utilizado para verificar a qualidade do ajuste (**Tabela 2**).

Tabela 2- Número de imagens de raízes de batata-doce utilizadas para a fenotipagem da cor, formato e danos por insetos (originais) e número de imagens destinadas ao ajuste das RCNNs (Ajuste) e avaliação da qualidade do ajuste (Teste)

Classes	Originais	Ajuste	Teste
----- Cor -----			
Clara	1160	600	560
Escura	1198	600	598
----- Formato -----			
Comercializáveis	1244	600	644
Não comercializáveis	744	600	144
----- Ataque insetos-----			
Danificada	744	600	144
Não danificada	1614	600	1014

Fonte: Da autora,2022.

A fim de ampliar o conjunto de dados no treinamento, cada uma das imagens foi replicada com quatro diferentes rotações (45°, 135°, 225° e 315°). Assim, para cada classificação, houve 3000 imagens no treinamento (600 + 4x600). Para o treinamento das redes, foi utilizada a linguagem python na plataforma do Google Colab. Utilizou-se a biblioteca Keras, considerando as arquiteturas VGG-16, Inception-v3, ResNet-50, InceptionResNetV2, EfficientNetB3. Considerou-se 100 o número máximo de iterações e early stopping com tolerância de cinco iterações.

Para avaliar os ajustes das redes convolucionais, foram construídas matrizes de confusão, tendo-se as classificações preditas pelas diferentes arquiteturas da rede em função das classificações visuais. Foram utilizadas as métricas Recall (Eq 1), Acurácia (Eq 2), Precisão (Eq 3), F-Measure (Eq 4) e Especificidade (Eq 5), para avaliar a eficiência da rede. Em que VP refere-se aos verdadeiros positivos, FN aos falsos negativos, FP aos falsos positivos e VN aos verdadeiros negativos.

$$Recall = \frac{VP}{VP+FN} \quad (Eq 1)$$

$$Acurácia = \frac{VP+VN}{VP+VN+FP+FN} \quad (Eq 2)$$

$$Precisão = \frac{VP}{VP+FP} \quad (Eq 3)$$

$$F - measure = \frac{2*Precisão*Recall}{Precisão+Recall} \quad (Eq 4)$$

$$\text{Especificidade} = \frac{VN}{FP+VN} \quad (\text{Eq 5})$$

### 4.3.3 Resultados

Para cada arquitetura, foi possível observar um número de épocas diferentes necessário para o treinamento (**Tabela 3**). Para o formato, danos causados por insetos e cor da casca, as arquiteturas Inception-v3 e InceptionResNetV2 apresentaram os menores tempos de classificação, bem como menores números de épocas (**Tabela 3**). Contudo, para todas as variáveis analisadas, houve maior taxa de classificações verdadeiras positivas (VP), para a arquitetura InceptionResNetV2. Já a arquitetura EfficientNetB3 foi a que apresentou menor eficiência, para a detecção do formato e danos causados por insetos, nas raízes de batata-doce, demandando maior número de épocas e, conseqüentemente, maior tempo para classificação (**Tabela 3**). Para a cor da casca, a arquitetura ResNet-50 apresentou menor eficiência.

Tabela 3: Número de épocas, tempo de treinamento e classificações realizadas por diferentes arquiteturas de RCNNs para raízes de batata-doce quanto ao formato, danos ocasionados por insetos e cor da casca. UFMG,2022.

Var	Arquitetura	Épocas	Tempo	VP	FN	FP	VN
Formato	VGG-16	70	0:29:58.5	517	127	132	382
	Inception-v3	25	0:12:00.3	543	101	158	356
	ResNet-50	76	0:31:48.1	526	118	160	354
	InceptionResNetV2	17	0:17:59.3	630	14	20	494
	EfficientNetB3	100	1:05:30.8	567	77	138	376
Danos causados por insetos	VGG-16	95	0:36:46.8	97	46	299	715
	Inception-v3	18	0:07:47.7	111	32	99	915
	ResNet-50	24	0:28:16.3	73	70	251	763
	InceptionResNetV2	11	0:06:27.7	140	3	38	976
	EfficientNetB3	100	0:45:22.6	87	56	262	752
Cor da casca	VGG-16	80	0:34:56.5	464	96	70	528
	Inception-v3	31	0:15:07.8	455	105	184	414
	ResNet-50	98	1:20:17.9	481	79	98	500
	InceptionResNetV2	32	0:34:53.3	540	20	2	596
	EfficientNetB3	98	1:06:43.3	505	55	54	544

VP: verdadeiros positivos; FN: falsos negativos; FP: falsos positivos; VN: verdadeiros negativos. Fonte: Da autora,2022.

A eficácia das arquiteturas das redes neurais convolucionais, ao classificar as raízes de batata-doce quanto ao formato considerado ideal, aos danos causados por insetos e quanto à cor da casca das raízes, foi avaliada, utilizando-se valores de precisão, recall, F-measure (F1), acurácia e especificidade (**Tabela 4**). A arquitetura InceptionResNetV2 foi a que obteve maior precisão, melhor acurácia e especificidade para todas as variáveis analisadas. Essa arquitetura possibilitou precisão próxima a 78,7%, para classificar

os danos causados por insetos, sendo essa métrica superior à precisão obtida pelas outras arquiteturas para essa mesma variável. Já para as demais arquiteturas, todas as métricas de avaliação foram inferiores a 91% para as três variáveis analisadas.

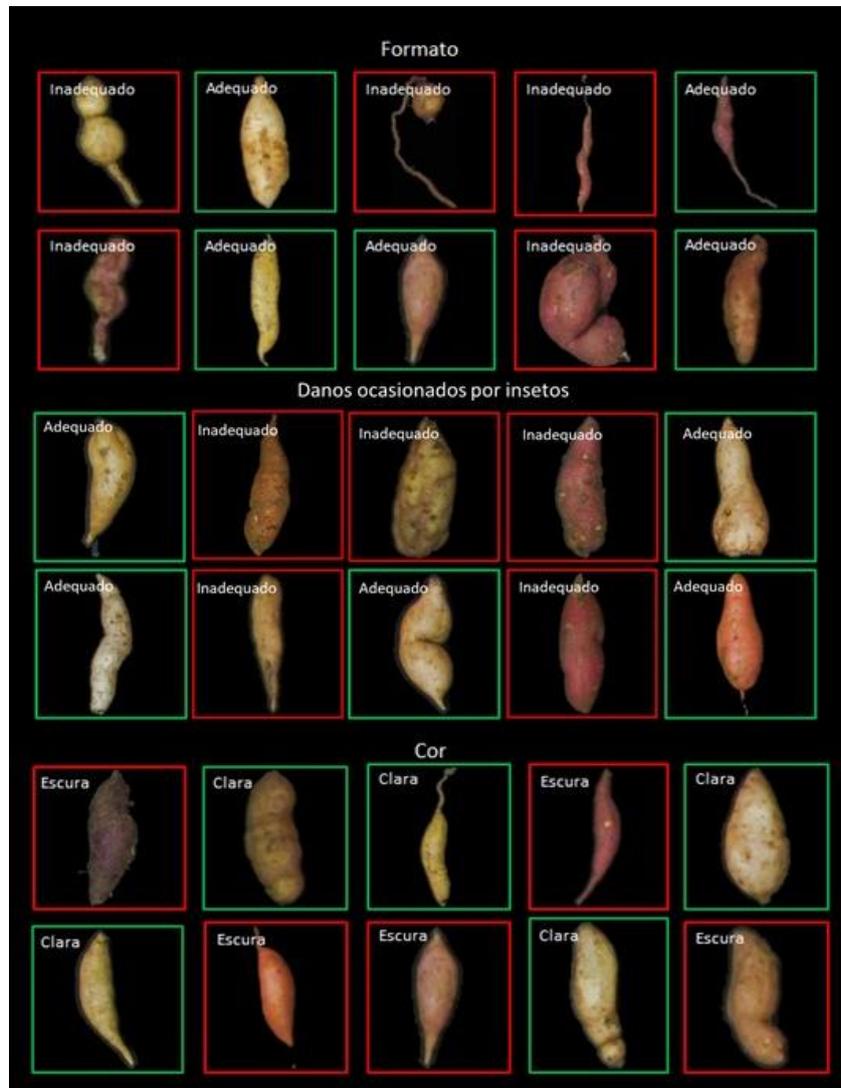
Tabela 4: Avaliadores da qualidade do ajuste de RCNNS com diferentes arquiteturas, na classificação de raízes de batata-doce, quanto ao formato, danos ocasionados por insetos e cor da casca. UFMG,2022.

Var	Arquitetura	Precisão	Recall	F1	Acurácia	Especificidade
<b>Formato</b>	<b>VGG-16</b>	0.797	0.803	0.800	0.776	0.743
	<b>Inception-v3</b>	0.775	0.843	0.807	0.776	0.693
	<b>ResNet-50</b>	0.767	0.817	0.791	0.760	0.689
	<b>InceptionResNetV2</b>	0.969	0.978	0.974	0.971	0.961
	<b>EfficientNetB3</b>	0.804	0.880	0.841	0.814	0.732
<b>Danos por insetos</b>	<b>VGG-16</b>	0.245	0.678	0.360	0.702	0.705
	<b>Inception-v3</b>	0.529	0.776	0.629	0.887	0.902
	<b>ResNet-50</b>	0.225	0.510	0.313	0.723	0.752
	<b>InceptionResNetV2</b>	0.787	0.979	0.872	0.965	0.963
	<b>EfficientNetB3</b>	0.249	0.608	0.354	0.725	0.742
<b>Cor da casca</b>	<b>VGG-16</b>	0.869	0.829	0.848	0.857	0.883
	<b>Inception-v3</b>	0.712	0.813	0.759	0.750	0.692
	<b>ResNet-50</b>	0.831	0.859	0.845	0.847	0.836
	<b>InceptionResNetV2</b>	0.996	0.964	0.980	0.981	0.997
	<b>EfficientNetB3</b>	0.903	0.902	0.903	0.906	0.910

Fonte: Da autora,2022.

Na **Figura 3**, a fim de visualizar a classificação de raízes selecionadas aleatoriamente das amostras teste pela InceptionResNetV2, foi possível perceber a precisão da rede, em classificar os indivíduos comercializáveis quanto ao formato, danos causados por insetos e coloração da casca. Foram considerados indivíduos indesejáveis (caixa vermelha) aqueles com formato irregular, com área danificada por insetos e de coloração escura. Dessa forma, para a população estudada, a classificação e, conseqüente fenotipagem das raízes de batata-doce quanto ao formato, danos causados por insetos e cor da casca, podem ser realizadas de forma eficiente e em menor tempo de resposta com a arquitetura InceptionResNetV2. Contudo vale ressaltar que apenas duas classes (adequado e inadequado) é importante, para a fase inicial do melhoramento genético, devendo essas classes serem destrinchadas em fases posteriores do melhoramento.

Figura 3- Classificação de raízes selecionadas aleatoriamente nas amostras pertencentes ao conjunto de imagens de teste por RCNN InceptionResNetV2, para a visualização da classificação de raízes de batata-doce quanto ao formato, danos ocasionados por insetos e cor da casca. UFMG,2022.



Fonte: Da autora,2022.

#### 4.3.4 Discussão

As redes neurais convolucionais (CNN) são tendência no processamento de informações imagéticas, por sua adaptabilidade e eficiência na detecção de objetos (VASCONEZ *et al.*,2020). O processo de aprendizado das redes se dá, por meio de processos iterativos de ajustes aplicados aos pesos sinápticos (treinamento), em que o aprendizado só ocorre, quando a rede neural atinge uma solução generalizada para um determinado problema (ZHANG *et al.*, 2020). Contudo, quanto maior o número de iterações utilizado no treinamento, maior tende a ser a memorização dos dados pelas redes, tendo, como consequência, o caráter não generalista do sistema (overfitting). Assim, é de suma importância a definição de um número ótimo de iterações, para os conjuntos de dados analisados, o que pode ser feito utilizando uma estratégia chamada earllystopping. Para cada tipo de arquitetura, bem como para cada variável utilizada, foi possível observar um número de épocas diferentes, em que as arquiteturas Inception-v3 e InceptionResNetV2 apresentaram os menores números de épocas e, conseqüentemente, os menores

tempos de classificação (**Tabela 3**). Isso indica a confiabilidade e eficiência do método, além de possibilitar a economia de tempo na obtenção de resultados (NI *et al.*, 2020). Em contrapartida, as arquiteturas EfficientNetB3 e ResNet-50 foram as que apresentaram menor eficiência na classificação do conjunto de dados estudado e tiveram também o maior número de épocas e maior tempo para a obtenção dos resultados.

A seleção da arquitetura que possibilita a análise de forma mais eficiente permite qualificar características em grandes conjuntos de dados e pouca mão de obra. Isso pode ajudar os melhoristas, na avaliação da interação genótipos x ambientes de forma mais eficaz, levando à identificação de potenciais novas cultivares em menor espaço de tempo (Haque *et al.*, 2020). Dessa forma, quanto maior a taxa de verdadeiro positivo identificado por cada arquitetura, para as variáveis formato da raiz, danos causados por insetos e cor da casca, maior foi a precisão e, conseqüentemente, recall, acurácia e F-Measure. Isso é corroborado, no presente estudo, pela arquitetura InceptionResNetV2, uma vez que foram obtidos maior precisão, melhor acurácia e especificidade para todas as variáveis analisadas, sendo a precisão, para essa rede, superior a 91% para a variável formato e coloração da casca da raiz, e superior a 78% para danos causados por insetos (**Tabela 4**). Ao utilizar imagens de batata-doce, para treinar um classificador de redes neurais, para classificar o formato da raiz de batata-doce, Haque *et al.* (2020) obtiveram precisão inferior às obtidas no presente estudo. Essa diferença de precisão pode estar associada às classes utilizadas por cada pesquisador para considerar a raiz com formato ideal ou não. Essas altas precisões das arquiteturas, utilizadas na classificação das raízes, podem levar a melhorar ainda mais o processo de tomada de decisão nas práticas agrícolas (VASCONEZ *et al.*, 2020).

A eficácia em classificar as raízes de batata-doce por análise computacional, na classificação de objetos, utilizando a arquitetura InceptionResNetV2, foi visualizada na **Figura 3**. Para isso, foram aplicadas caixas delimitadoras, que permitiram demonstrar a categorização dos dados. Por essa abordagem, a rede neural pode adquirir a capacidade de classificação das raízes de batata-doce para cada característica analisada. A alta eficiência obtida pelo sistema desenvolvido pode ser justificada também pela ausência de sobreposição dos objetos, permitindo maior precisão na identificação (TODA *et al.*, 2020). Além disso, outro fator que pode ter influenciado nos valores das métricas é a nitidez de coloração das raízes, uma vez que regiões defeituosas podem ter um padrão de cor diferente do resto da raiz, melhorando a precisão da classificação, fornecendo novas informações sobre a qualidade das raízes (HAQUE *et al.*, 2020). Assim, quanto mais fácil a distinção da matriz RGB da imagem quanto aos seus objetos e suas partes, maior a chance de sucesso na detecção das classes avaliadas (JEONG *et al.*, 2020).

Por meio dos resultados obtidos, pode-se inferir que o sistema desenvolvido se mostra eficiente na classificação de raízes de batata-doce, tornando a interpretação mais rápida, precisa e menos subjetiva. Esse é um dos principais diferenciais da técnica desenvolvida, uma vez que a maioria das pesquisas, investigando características de tamanho e forma de culturas hortícolas, são realizadas em escala de laboratório, não sendo adaptadas à produção para grande escala (HOWARTH E SEARCY, 1991; CLEMENT *et al.*, 2013). Logo o sistema desenvolvido tem grande potencial, para ser adaptado e utilizado, para coletar e analisar dados pelo pequeno produtor, bem como em largas escalas comerciais. Auxilia não apenas no processo de fenotipagem da cultura, mas também na separação de raízes consideradas comerciais ou não. Para o melhoramento genético, a quantificação de raízes pertencentes a cada classe

pode ser utilizada, como critério de seleção, seja individual ou simultâneo, agregando informações importantes para o melhor progresso genético da cultura.

Dessa forma, essa abordagem torna possível quantificar a perda pela deformação de formato e danos causados por insetos. Contudo um grande desafio, na implantação desse método, é o poder computacional necessário, para o processamento de centenas de milhares raízes, o que requer computadores de elevado custo. Uma maneira de mitigar esse desafio é o uso de servidores em nuvem (HAQUE *et al.*,2020). Além disso, a metodologia desenvolvida propicia oportunidades, para investigar outras culturas hortícolas, indicando a possibilidade de desenvolvimento de equipamentos de esteira com câmeras para a fenotipagem em alta escala, seja para fins comerciais ou para o melhoramento genético.

#### 4.3.5 Conclusões

A arquitetura InceptionResNetV2 teve melhor desempenho, ao classificar os indivíduos quanto ao formato, dano causado por insetos e coloração da casca, obtendo altas estimativas para os parâmetros utilizados na avaliação da qualidade do ajuste.

A análise de imagens associada ao aprendizado profundo é uma alternativa com potencial, para o melhoramento qualitativo de raízes de batata-doce, permitindo a redução da subjetividade das análises, bem como a redução no tempo de fenotipagem da cultura.

A eficiência das metodologias utilizadas indica o potencial do desenvolvimento de aplicativos e aparelhos, para a classificação de raízes de batata-doce, a fim de auxiliar o produtor rural, comerciantes e melhoristas.

#### 4.3.6 Agradecimentos

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior- Brasil (CAPES) – Código de Financiamento 001, do Conselho Nacional de Desenvolvimento Científico (CNPQ) e a Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG).

#### 4.3.7 Referências

CLEMENT, J.; NOVAS, N.; GAZQUEZ, J.-A.; MANZANO-AGUGLIARO, F. An active contour computer algorithm for the classification of cucumbers. **Comput. Electron. Agric.** v.92, p.75–81.2013.

FILGUEIRA, FAR. **Novo manual de olericultura: agrotecnologia moderna na produção e comercialização de hortaliças**. Viçosa: UFV, 2008,402p.

HAQUE, S.; LOBATON, E.; NELSON, N.; YENCHO, G.C.; PECOTA, K.V.; MIEROP, R.; KUDENOV, M. W.; BOYETTE, M.; WILLIAMS, C.M. Computer vision approach to characterize size and shape phenotypes of horticultural crops using high-throughput imagery. **Computers and Electronics in Agriculture**, v.182, 2021.

Howarth, M.S.; Searcy, S.W. Fresh market carrot inspection by machine vision. In: Optics in Agriculture, vol. 1379, International Society for Optics and Photonics, p.141–150, 1991.

HUAMÁN, Z. **Descriptors for sweet potato**. Rome: International board for plant Genétic Resources, 1991. 133p.

JEONG, YS, LEE, HR, BAEK, JH, KIM, KH, CHUNG, YS, & LEE, CW. Deep learning-based rice seed segmentation for phynotyping. **Journal of the Korea Industrial Information Systems Research**, v.25, n.5, p.23-29, 2020.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning.**Nature**, v. 521, n. 7553, p. 436-444, 2015.

LEE, U.; CHANG, S.; PUTRA, G. A.; KIM, H.; KIM, D. H. An automated, high-throughput plant phenotyping system using machine learning-based plant segmentation and image analysis. **PloS one**, v. 13, n. 4, p. e0196615, 2018.

MASSRUHÁ, S. M. F. S.; LEITE, M. A. A.; LUCHIARI JUNIOR, A.; ROMANI, L. A. S. **Tecnologias da Informação e Comunicação e suas relações com a agricultura**. Brasília, DF: Embrapa, p. 85–100, 2014.

NI, X.; LI, C.; JIANG, H.; & TAKEDA, F. Deep learning image segmentation and extraction of blueberry fruit traits associated with harvestability and yield. **Horticulture Research**, v7, n.1, p.1-14, 2020.

PRADHAN, D.M.P.; MUKHERJEE, A.; GEORGE, J.; CHAKRABARTIS, K.; VIMALA B.; NASKAR, S.K.; SAHOOB, K.; SAMAL, S. High starch, beta carotene and anthocyanin rich sweet potato: ascent to future food and nutrition security in coastal and backward areas. **International Journal of Tropical Agriculture**, v. 33, n. 2 (Part I), p.397-400, 2015.

RAHAJENG, W.; RESTUONO, J.; INDRIANI, F.C.; PURWONO, P. 2020. Genetic Parameters of Agronomic Traits in Sweet-potato Accessions. **Biosaintifika: Journal of Biology & Biology Education**, v.12, n.2, p.240-246, 2020.

SAMBASIVAM, G., OPIYO, GD. Uma aplicação preditiva de aprendizado de máquina na agricultura: detecção e classificação de doenças da mandioca com conjunto de dados desequilibrado usando redes neurais convolucionais. **Jornal de Informática Egípcia**, v.22, n.1, p.27-34, 2021.

SHAHEEN, F.; VERMA, B. E ASAFUDDOULA, M. IMPACT OF AUTOMATIC FEATURE EXTRACTION IN DEEP LEARNING ARCHITECTURE. In Digital Image Computing: Techniques and Applications (DICTA), **International Conference on**, pp. 1–8. IEEE, 2016.

SILVA, M. J.; SCHIMIGUEL, J. Identificação de Doenças em Plantas por meio de Processamento de Imagens: Redes Neurais Convolucionais como Auxílio à Agricultura. **Revista de Ubiquidade**, v.3, n.1, p.91-111, 2020.

TODA, Y.; OKURA, F.; ITO, J.; OKADA, S.; KINOSHITA, T.; TSUJI, H.; SAISHO, D. Training instance segmentation neural network with synthetic datasets for crop seed phenotyping. **CommunBiol**, v.3, n.173, 2020.

UBBENS, J. R.; STAVNESS, I. Deep plant phenomics: a deep learning platform for complex plant phenotyping tasks. **Frontiers in plant science**, v. 8, p. 1190, 2017.

VASCONEZ, J. P.; DELPIANO, J.; VOUGIOUKAS, S.; CHEEIN, F. A. Comparison of convolutional neural networks in fruit detection and counting: A comprehensive evaluation. **Computers and Electronics in Agriculture**, v. 173, p. 105348, 2020.

WANG, X.; TIAN, S.; LOU, H.; ZHAO, R. A reliable method for predicting bioethanol yield of different varieties of sweet potato by dry matter content. **Grain & Oil Science and Technology**, v.3, n.3, p.110-116, 2020.

WERA, B.; YALU, A.; RAMAKRISHNA, A.; DEROS, M. Genotypic Variability Estimates Of Agronomic Traits For Selection In A Sweetpotato (*Ipomoea batatas*) Polycross Population In Papua New Guinea. **J. Plant Breed. Genet.** v.2, n.3, p.131- 136, 2014.

ZHANG, J.; YAN, Y.; CHENG, Z .; WANG, W. Rede de pirâmide de atenção leve para detecção de objetos e segmentação de instâncias. **Appl. Sci.** v.10, p.883, 2020.

## 5 CONSIDERAÇÕES FINAIS

A utilização da análise de imagens associada à visão computacional, bem como ao aprendizado profundo, é uma ferramenta eficiente, em programas de melhoramento genético de batata-doce, auxiliando na fenotipagem da cultura, bem como na tomada de decisões. Além disso, pode ser utilizada no desenvolvimento de aplicativos e aparelhos que auxiliam o produtor na classificação de raízes de batata-doce. Isso possibilita a utilização da hortaliça, para os seus mais diversos fins, reduzindo perdas e desperdício no processo produtivo da cultura.

A visão computacional permitiu a predição do peso das raízes, mantendo o ranqueamento dos genótipos, quando comparados à metodologia de avaliação tradicional, levando conseqüentemente à similaridade entre os ganhos esperados com a seleção.

A metodologia desenvolvida apresentou eficiência na classificação das polpas de raízes de batata-doce pela análise de imagens, indicando a variabilidade genética entre os genótipos estudados quanto à coloração da polpa das raízes, em que os genótipos UFVJM-31 e UFVJM-40 apresentaram grande dissimilaridade se comparados aos demais.

A análise de imagens associada ao aprendizado profundo com a arquitetura InceptionResNetV2 é uma alternativa eficiente, no melhoramento qualitativo de raízes de batata-doce, permitindo a redução da subjetividade humana das análises, bem como a redução no tempo de fenotipagem da cultura.