



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA QUÍMICA
ENG07053 - TRABALHO DE DIPLOMAÇÃO EM ENGENHARIA
QUÍMICA



Classificação de Batatas-doces através de Técnicas de Inteligência Artificial

Autora: Bruna Pinto

Orientador: Prof. Dr. Jorge Otávio Trierweiler

Coorientador: Rafael Henrique Martello

Porto Alegre, maio de 2021

Autora: Bruna Pinto

Classificação de Batatas-doces através de Técnicas de Inteligência Artificial

Trabalho de Conclusão de Curso apresentado à COMGRAD/ENQ da Universidade Federal do Rio Grande do Sul como parte dos requisitos para a obtenção do título de Bacharel em Engenharia Química

Coorientador: Rafael Henrique Martello

Banca Examinadora:

Prof. Dr. Edson Cordeiro Do Valle, UFRGS

Msc. Eng. Lucas Ranzan, UFRGS

Porto Alegre

2021

AGRADECIMENTOS

Primeiramente, agradeço à minha família por todo o suporte e incentivo que me ofereceram em toda minha trajetória, sem medir esforços. Obrigada por serem meu porto seguro, por acreditarem nos meus sonhos e torcerem por mim.

Agradeço ao meu orientador Prof. Dr. Jorge Otávio Trierweiler pela oportunidade e proposição do tema, me incentivando sempre a ir em busca do conhecimento.

Ao meu coorientador Rafael Henrique Martello pela incansável ajuda que me foi fornecida durante este período, pelas aulas sobre redes neurais e por sempre buscar, junto comigo, os melhores resultados para este trabalho.

RESUMO

Aplicações modernas que utilizam de métodos de Inteligência Artificial (IA) para resolver problemas em diferentes áreas de atuação tem se tornando cada vez mais comuns. Um uso muito popular desta tecnologia é no reconhecimento e classificação de imagens, onde técnicas de IA baseadas em Redes Neurais Convolucionais (CNN) podem apresentar ótimos resultados. Essas redes se baseiam na operação de convolução, sendo compostas por uma camada de entrada, uma camada de saída e diversas camadas ocultas, algumas das quais são convolucionais. Essa técnica tangibiliza o conceito de classificação de imagens que será utilizada neste trabalho com batata-doce brasileira roxa. A batata-doce é uma cultura extremamente popular e cultivada no mundo todo. No entanto, sua comercialização é diretamente afetada pela sua aparência ao consumidor final, muitas vezes gerando desperdício pela falta de classificação qualitativa desses produtos pelos produtores. Estudos realizados pela Organização das Nações Unidas para a Alimentação e a Agricultura (FAO), estimam que 1,3 bilhões de toneladas de comida são desperdiçadas por ano, representando aproximadamente 30% de toda a comida produzida no planeta. Além disso, a crescente busca por uma economia circular e desenvolvimento sustentável vem sendo cada mais procurada, onde o que não é consumido como alimento para o ser humano possa servir como matéria-prima para outros fins, mantendo os materiais em ciclo produtivo. O objetivo deste trabalho é classificar imagens de batata-doce brasileira roxa de forma qualitativa em relação a probabilidade de ser consumida, baseada em técnicas de visão computacional. Além disso, também são abordados os principais pontos e conceitos relacionados à identificação de imagens, etapas do processamento de imagens e rede neural. Para obter uma alta acurácia, avaliamos quatro modelos de redes neurais convolucionais treinadas com diferentes parâmetros de ajuste fino e aprendizagem de transferência. Os resultados experimentais mostram uma eficácia promissora dos modelos pré-treinados, resultando em uma acurácia de 90,14% com o modelo Xception. Logo, a aplicação de rede neural convolucional pode ser uma alternativa interessante para a classificação de imagens, o que permite contribuir para apoiar agricultores e comerciantes a reduzirem o desperdício de batatas-doces. Essa ferramenta integra um projeto maior voltado ao desenvolvimento de “Biorrefinarias Descentralizadas e Autônomas: um Modelo de Economia Circular, Verde e Digital”.

Palavras-chave: *batata-doce roxa, redes neurais, inteligência artificial, classificação de imagens*

ABSTRACT

Modern applications that use Artificial Intelligence (AI) methods to solve problems in different areas of activity have become increasingly common. A very popular use of this technology is in the recognition and classification of images, where AI techniques based on Convolutional Neural Networks (CNN) can present excellent results. These networks are based on the convolution operation, being composed of an input layer, an output layer and several hidden layers, some of which are convolutional. This technique gives tangibility to the concept of image classification that will be used in this work with Brazilian sweet potato. Sweet potato is an extremely popular crop and is grown around the world. However, their commercialization is directly affected by their appearance to the final consumer, often generating waste due to the lack of qualitative classification of these products by the producers. Studies carried out by the Food and Agriculture Organization of the United Nations (FAO), estimate that 1.3 billion tons of food are wasted per year, representing approximately 30% of all food produced on the planet. In addition, the growing search for a circular economy and sustainable development has been increasingly sought after, where what is not consumed as food for humans can serve as raw material for other purposes, keeping the materials in a productive cycle. The objective of this work is to classify images of Brazilian sweet potato in a qualitative way in relation to the probability of being consumed, based on computer vision techniques. In addition, the main points and concepts related to image identification, image processing steps and neural networks are also addressed. To obtain high accuracy, we evaluated four models of convolutional neural networks trained with different parameters of fine-tuning and transfer learning. The experimental results show a promising efficacy of the pre-trained models, resulting in an accuracy of 90.14% with the Xception model. Therefore, the application of convolutional neural network can be an interesting alternative for the classification of images, which allows to contribute to support farmers and traders to reduce the waste of sweet potatoes. This tool is part of a larger project aimed at the development of "Decentralized and Autonomous Biorefineries: a Circular, Green and Digital Economy Model".

Keywords: purple sweet potato, neural networks, artificial intelligence, image classification

LISTA DE FIGURAS

Figura 1: Esquema de neurônios compondo uma Rede Neural Artificial Multicamadas	3
Figura 2: Modelo não-linear dos neurônios utilizados na rede neural artificial	4
Figura 3: Produção total de batata-doce no Brasil, em toneladas.	9
Figura 4: Exemplos do conjunto de imagens das duas espécies de batata-doce utilizadas, por classe	14
Figura 5: Arquitetura da Rede Neural Convolutcional	15
Figura 6: Representação de um <i>maxpooling</i>	17
Figura 7: Resumo do modelo XCEPTION.....	19
Figura 8: Resumo do “Meu modelo”	20
Figura 9: Gráficos da Acurácia de treinamento e teste, pelo nº de épocas	21
Figura 10: Gráficos da perda de validação, pelo nº de épocas.....	22
Figura 11: Matriz de confusão para os dados de teste do modelo Xception.....	22

LISTA DE TABELAS

Tabela 1: Cultivares de batata-doce de acordo com a EMBRAPA.	8
Tabela 2: Produção total e produtividade de batata-doce nas regiões brasileiras (2019).....	10
Tabela 3: Listas de hardware e software usados neste estudo	13
Tabela 4: Métricas de Desempenho usadas para avaliar os modelos de aprendizagem profunda	21

LISTA DE ABREVIATURAS E SIGLAS

IA	Inteligência Artificial
RGB	Red, Green, Blue
KNN	K-Nearest Neighbor
ANN	Artificial Neural Network
SVM	Support Vector Machine
DL	Deep Learning
CNN	Convolutional Neural Network
ReLU	Rectified Linear Unit

SUMÁRIO

1	Introdução	1
2	Fundamentação Teórica sobre Redes Neurais	3
2.1	Redes Neurais	3
2.1.1	Modelos para Classificação de Imagens	4
2.1.2	Redes Neurais Convolucionais	5
2.2	Critérios de Qualidade de Modelo	5
3	Revisão Bibliográfica	7
3.1	Biorrefinarias - Modelos de Produção	7
3.2	Produção de Batata Doce	8
3.3	Técnicas de Classificação de Imagens	11
4	Materiais e Métodos	13
4.1	Software e Hardware	13
4.2	Aquisição de Imagens	13
4.3	Pré-Processamento das Imagens	14
4.4	Segmentação das Imagens	14
4.5	Transferência de Aprendizado	15
4.6	Extração de recursos	16
4.7	Classificação	17
4.8	Medidas de desempenho	17
5	Resultados	19
5.1	Modelo	19
5.2	Desempenho	20
6	Conclusões e Trabalhos Futuros	24
	REFERÊNCIAS	25

1 Introdução

A batata-doce (*Ipomoea batatas L.*) pertence à família da Convolvulaceae e tem sua origem da América Central e do Sul, onde é vista como um potencial ativo de alto valor agregado. Por ser uma espécie rústica que se adapta facilmente em diversos climas e solos, além de possuir uma grande resistência a pragas e doenças, a batata-doce é considerada mundialmente a hortaliça com baixo custo de produção (Marques, 2015). Ela é cultivada em mais de 100 países e o Brasil está na 16ª posição do ranking mundial, sendo a China a líder com produção anual de cerca de 52 milhões de toneladas (FAOSTAT, 2019). Em relação as suas aplicações, a batata-doce é caracterizada como uma cultura alimentar básica no desenvolvimento de países e pode servir também como ração animal e matéria prima para diversos produtos industriais (Unajan et al., 2017). Além desses usos, estudos mostram que a batata-doce possui elevada capacidade de conversão de biomassa em matéria prima para produção de etanol (Lourenço et al., 2018). Nesse sentido, sua produção para este fim torna-se uma opção de renda, além da venda como alimento, para pequenos e grandes agricultores.

De acordo com as tendências apontadas pela ONU, a população mundial poderá chegar a 8,6 bilhões de pessoas em 2030 (ONU, 2017). Com esta perspectiva de crescimento populacional, uma grande preocupação vem à tona, a geração de resíduos alimentares e por consequência, o desperdício, que é um dos maiores problemas enfrentados na comercialização de produtos. Segundo a Organização das Nações Unidas para a Alimentação e a Agricultura (FAO), 1,3 bilhões de toneladas de comida são desperdiçadas por ano, representando aproximadamente 30% de toda a comida produzida no planeta. Sendo assim, esse resíduo alimentar é tradicionalmente destinado para a produção de ração, uso em fertilizantes ou aterrados com outros resíduos (Bueno, 2019).

No Brasil, pesquisas realizadas pela Embrapa Agroindústria de Alimentos indicaram que as perdas no segmento de hortaliças atingem, em média, 35%. Dentre as principais causas destaca-se o manuseio inadequado no campo, classificação não padronizada, veículos supercarregados, excesso de “toque” nos produtos por parte dos consumidores, acúmulo de produtos nas gôndolas de exposição no varejo, além de outros aspectos. Ainda, para corroborar com os dados anteriores, um trabalho realizado pela FAO (2017) verificou que produtos selecionados e classificados apresentam melhor aceitação pelos consumidores na hora da compra. Diante desse contexto, a crescente demanda por proteção ambiental juntamente com o desejo de conservação de recursos, está gerando uma valorização mais eficiente desses resíduos para a produção de ingredientes alimentares de maior valor agregado, produtos químicos e biocombustíveis.

Como solução para esse tipo de questão, há uma motivação crescente para o uso de técnicas de visão computacional, tornando possível a criação de um sistema para identificar e classificar os produtos de forma automática e ágil. Isso permite reduzir o desperdício de batatas comercializadas cujas características físicas não estão alinhadas com o gosto dos consumidores. Com um crescente aumento de produção do agronegócio, torna-se essencial introduzir técnicas da indústria 4.0 para promover a competitividade do cenário agroindustrial nacional, possibilitando a automatização da linha de produção. Sendo assim, todas as batatas-doces classificadas com baixa probabilidade de serem consumidas, poderão ser destinadas

para a produção de biocombustíveis antes mesmo de serem dispostas nos supermercados ou serem descartadas.

Nesse contexto, destaca-se o conceito de biorrefinarias de resíduos de batata-doce como uma alternativa de viabilidade econômica e sustentável, alinhada com a proposta de economia circular. As biorrefinarias são uma possibilidade de agregar valor à biomassa através de sua transformação em produtos químicos e combustíveis. Para o sucesso econômico desse modelo descentralizado de produção, há a necessidade da integração entre processos inovadores e eficientes, com logística e acompanhamento remoto diferenciado, de modo a potencializar todos os integrantes da cadeia produtiva (Trierweiler et al., 2020).

Um exemplo de aplicação é a destinação para a produção de etanol. Estudos mostram que o pico na produção de etanol é de 25 dias após a colheita da batata-doce (Schweinberger, 2016). Nesse sentido, é imprescindível que uma classificação prévia seja realizada para que se garanta a eficiência do processo, sem comprometer sua aplicação para o setor alimentício para aquelas que estão de acordo com os padrões aceitos pelo consumidor. Com isso, torna-se viável obter o total aproveitamento da batata-doce, sendo a proposta destinar aquelas não selecionadas para o consumo à produção de etanol, álcool em gel e bebida destilada shochu por meio de Biorrefinarias Descentralizadas e Autônomas apresentado em mais detalhes no trabalho de Trierweiler et al. (2020). Dessa forma, é possível utilizar a classe de baixa probabilidade de consumo para ser diretamente destinada à indústria de produção de etanol.

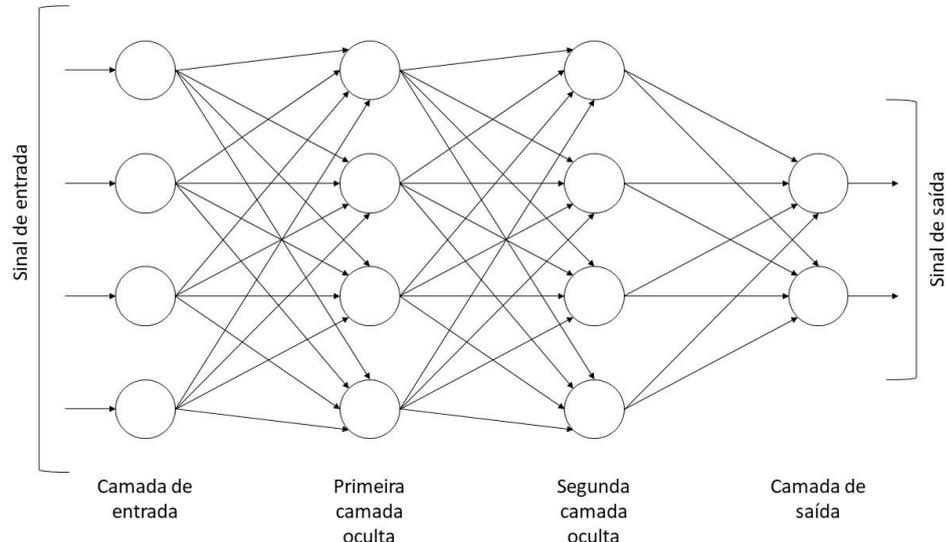
A aplicação da batata-doce não selecionada na produção de etanol seria uma das diferentes formas que poderíamos contribuir com o desenvolvimento sustentável e aproveitamento total desta hortaliça. Assim sendo, o presente trabalho tem o objetivo de desenvolver métodos de utilização de inteligência artificial no processo de classificação de imagens de batata-doce brasileira roxa. Isso possibilitará efeitos positivos na diminuição de resíduos e desperdícios, contribuindo para o meio ambiente e retorno econômico para os produtores rurais.

2 Fundamentação Teórica sobre Redes Neurais

2.1 Redes Neurais

Uma rede neural é uma estrutura de computação influenciada pelo processamento neural biológico. Ela realiza cálculos úteis por meio de um processo de aprendizagem e têm sido largamente aplicadas em tarefas relacionadas com o reconhecimento de padrões, processamento de sinais e identificação e controles de sistemas (Bastos, 2007). Quando estamos trabalhando com redes neurais, também estamos utilizando tensores. Os tensores são uma estrutura de dados com, pelo menos, mais de três dimensões, que podem assumir N dimensões. Quando utilizamos imagens em redes neurais, estamos introduzindo números dentro de um operador que gera novos números. Esse é o processo que acontece dentro dos neurônios. Cada neurônio é equivalente à uma operação numérica, que se associa em camadas que farão as operações matemáticas com os tensores, aplicando funções e gerando novos resultados. As redes neurais que serão abordadas neste trabalho são baseadas no código disponível em Keras, API de redes neurais de alto nível com capacidade de execução no TensorFlow (<https://keras.io/api/applications/>).

Figura 1: Esquema de neurônios compondo uma Rede Neural Artificial Multicamadas



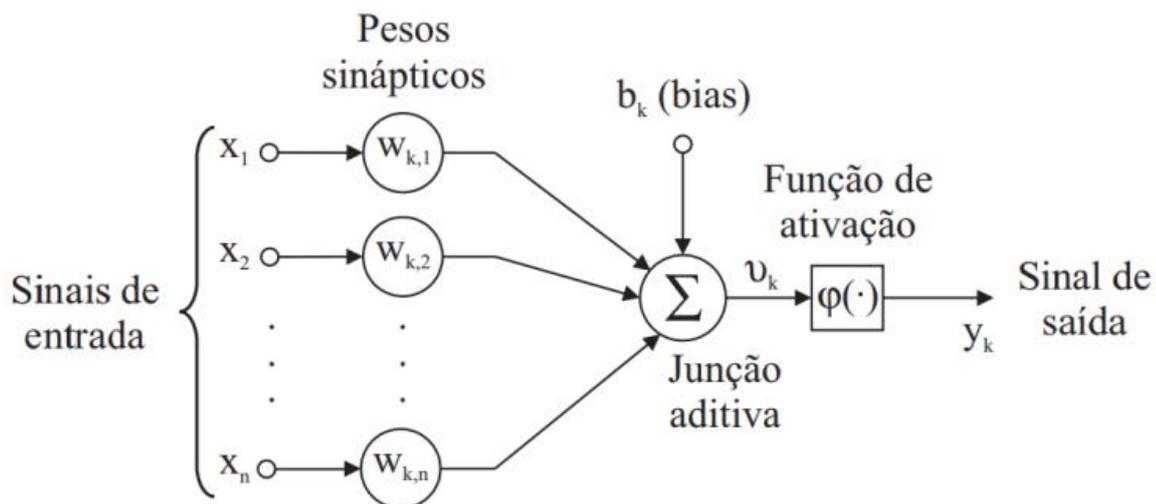
Fonte: Adaptado de De Oliveira et al. (2010)

A Figura 1 apresenta um esquema de neurônios compondo uma rede neural artificial do tipo multicamadas. A camada de entrada é responsável por receber os dados e distribuir as informações para a(s) camada(s) oculta(s) da rede. Como neste trabalho estamos trabalhando com imagens, esta camada receberá os pixels. As camadas internas e a camada de saída, são responsáveis por realizarem operações matemáticas, sendo que, neste trabalho, esta última obterá o resultado da classificação. Os neurônios são representados pelos círculos na imagem,

sendo responsáveis por receberem as entradas de cada camada e multiplicarem por um determinado peso sináptico, que são as linhas que chegam em cada círculo. O processamento em cada neurônio se dá pelo que chamamos de função de ativação. A escolha das funções de ativação de uma rede neural são uma consideração importante uma vez que definem como devem ser seus dados de entrada. A função de ativação faz esse processo para cada um dos neurônios e no final gera uma saída que significa o percentual de chance que a entrada é aquilo que deseja identificar.

A Figura 2 representa um neurônio, onde há primeiramente os sinais de entradas, que serão imagens de batatas-doces representadas como vetores entre 0 e 1. A partir do momento que entram em um neurônio, elas recebem um peso, denominado de pesos sinápticos, que pode ser definido pelo usuário ou pela própria rede. Neste trabalho utilizou-se uma RNA do tipo perceptron de múltiplas camadas (MLP - *Multilayer Perceptron*). Esse tipo de rede o sinal de entrada se propaga para frente (*feedforward*), camada por camada, e após cada resultado, é feito um processo de retro propagação (*backpropagation*) para a correção do erro, ou seja, serão feitos diversos ajustes iterativos dos pesos da rede para minimizar as diferenças entre a saída atual e a saída desejada. Este processo é repetido durante várias iterações até a finalização do treinamento. No final do treinamento, o resultado será a probabilidade que determinada imagem de ser aquilo que queremos identificar.

Figura 2: Modelo não-linear dos neurônios utilizados na rede neural artificial



Fonte: Zanetti et al., 2008

2.1.1 Modelos para Classificação de Imagens

Nos últimos anos, inúmeras técnicas de aprendizado de máquina vem sendo exploradas para classificação de imagens. Para realizar este processo, as redes neurais aplicadas são as mais diversas, cada qual com suas vantagens e desvantagens. Para mais informações sobre a comparação dos diferentes métodos e aplicações de redes neurais, consultar o artigo nas feito

por Hameed et al., (2018) nas referências. Este artigo apresenta uma revisão abrangente de técnicas de classificação de frutas e vegetais estudada por diversos pesquisadores. Para o presente trabalho, utilizamos a rede neural convolucional nas simulações do algoritmo, e esta será melhor explanada no subcapítulo a seguir.

2.1.2 Redes Neurais Convolucionais

Esse tipo de rede é utilizado principalmente em visão computacional para o caso que se deseja processar vídeos e imagens e pode ser aplicada em carros autônomos, robôs, reconhecimento facial, entre outras. A visão computacional é a tecnologia que concede as máquinas a capacidade de enxergar, é responsável por estudos e desenvolvimento de algoritmos que fornecem a capacidade para o computador processar e entender imagens. As redes neurais convolucionais (CNN) se baseiam na operação de convolução. O número de parâmetros da rede não é vinculado ao número do tamanho da entrada da imagem, resolvendo o problema que as redes neurais tradicionais possuem de ter muitos parâmetros, consumindo uma grande memória do processador. Além disso, algo que se tornou essencial em muitas aplicações é que a rede convolucional interpreta a imagem de forma intuitiva. As características que são aprendidas podem ser interpretadas visualmente e são as mais utilizadas quando se tenta produzir uma interpretação da metodologia baseada em Deep Learning (DL). Assim, este estudo tem como objetivo o desenvolvimento de um algoritmo de classificação automática de batata-doce, fazendo com que o computador possa distinguir entre imagens com alta, média e baixa probabilidade de serem aceitas pelo consumidor final.

O DL é uma técnica de aprendizado de máquina baseado no ajuste dos modelos a partir de um conjunto de dados de calibração. Uma imagem, em termos computacionais, nada mais é que uma matriz de pixels, ou seja, são números organizados em uma matriz que definem a imagem. Além das diversas vantagens dessa técnica, a mais importante no processamento de imagem é a automatização computacional. Já em relação as desvantagens, a principal delas é o tempo de treinamento geralmente mais longo. Além disso, ao usar modelos pré-treinados em conjuntos de dados que são pequenos ou significativamente diferentes, podem ocorrer problemas de otimização devido à complexidade dos modelos, bem como restrições de hardware (Kamilaris e Prenafeta-Boldú, 2018).

Neste trabalho, foram testados modelos utilizando um conjunto de imagens e uma rede neural convolucional. Em geral, as camadas de convoluções operam em tensores 3D, chamados mapas de recursos, com dois eixos espaciais de altura e largura, além de um eixo de canal chamado de profundidade. Para uma imagem colorida RGB, a dimensão do eixo de profundidade é 3, pois a imagem possui três canais: vermelho, verde e azul (Prabhu, 2018). O processo da CNN começa com convolução e agrupamento, decompondo a imagem em características e analisando-as de forma independente.

2.2 Critérios de Qualidade de Modelo

Para a avaliação dos resultados da metodologia proposta, várias métricas são utilizadas para avaliar o desempenho do método, sendo popularmente usadas na comparação de

modelos de classificação. Cada uma delas busca avaliar um aspecto diferente do modelo e permitem com que possamos avaliar os resultados de maneira mais simples na forma de um único número. A função perda indica quanto a resposta da rede neural está distante da resposta desejada. Além disso, também é utilizado a acurácia, que é a média entre o número de classes do verdadeiro positivo dividido pelos verdadeiros positivos e os falsos positivos (Selvaraj et al., 2019). E por fim, as matrizes de confusão são mecanismos abreviados que ilustram como as amostras pertencentes a um único tópico, agrupamento ou classe (linhas na matriz) são atribuídas à pluralidade de tópicos, agrupamentos ou classes possíveis.

3 Revisão Bibliográfica

3.1 Biorrefinarias - Modelos de Produção

Como consequência da produção e processamento de alimentos, quantidades significativas de subprodutos e resíduos são gerados diariamente. Isso faz com que, cada vez mais, se busque encontrar alternativas apropriadas para essas práticas, visando a utilização de técnicas mais sustentáveis e econômicas, minimizando a geração de resíduos e efeitos ambientais. Neste contexto, vamos apresentar no presente capítulo os modelos de produção centralizada e descentralizada, para o qual esta última será dado maior foco.

A centralização da produção em plantas industriais é caracterizada por ter ao seu entorno diversas monoculturas do produto de interesse, acarretando um elevado gasto com transporte desses locais até a indústria onde será efetivamente processado. Além disso, essa prática pode implicar no desgaste de recursos humanos, naturais, pouca interação com a pecuária, destruição da diversidade ecológica e das economias locais, gerando um baixo poder de distribuição de renda (Ortega et al., 2008).

O ponto chave para esta questão é que a biomassa e os resíduos industriais e agrícolas estão disponíveis de forma descentralizada. Desse modo, surge o conceito de biorrefinarias descentralizadas com o objetivo de utilizar e potencializar as biomassas disponíveis regionalmente. Dentre os vários benefícios, essa prática também otimiza recursos existentes, promove o reaproveitamento de resíduos do processo nas propriedades rurais próximas, reduz custos com logística e gera menores impactos ambientais (Borges, 2010). Nesse caso, a unidade produtiva de biomassa está perto das localidades onde é produzida a biomassa, nas quais diversos cultivos podem ser aproveitados ao mesmo tempo, transportando para as indústrias somente o produto acabado ou semiacabado. Esse processamento inicial de biomassa em pequena escala perto do local de colheita promoverá vários benefícios em termos de transporte mínimo, novas formas de integração na utilização de energia e organização do trabalho (Sanders et al. 2007).

O conceito de biorrefinaria não é uma ideia nova, mas nos últimos anos tem ganhado relevância mundial, pois agrega valor à biomassa através de sua transformação não somente em combustíveis, mas também em produtos químicos e outros tipos de materiais. Esse contexto vai ao encontro da busca crescente por uma economia circular. Este desafio chave significa que vários produtos e serviços, bem como modelos de negócios, precisam ser considerados e projetados em conjunto, levando em conta a sustentabilidade da cadeia como um todo (Natural Resources Institute Finland, 2016). Para as futuras biorrefinarias, a correta seleção dos recursos de biomassa é extremamente crítica para que essa alternativa possa operar em simbiose, ao invés de competir, com o setor de alimentos e outros mercados e contribuir para uma completa utilização do produto. Desta forma, para o caso do uso de batata-doce como biomassa para a produção de etanol, a seleção automatizada irá proporcionar a seleção do que é destinado a alimentos e o que é destinado à indústria.

Segundo estudos realizados por Schweinberger (2016), o modelo de biorrefinarias descentralizadas para a produção de etanol através da batata-doce, onde o etanol é processado em pequenas unidades, obteve resultados promissores. Para a análise de uma

microusina, foi adotada uma capacidade produtiva de 1000 L/dia de etanol hidratado. Em relação a matéria-prima, Schweinberger destaca que é possível obter um aproveitamento total da batata-doce, desde as raízes até as folhas. A grande vantagem é justamente a flexibilização na destinação desta hortaliça, onde as batatas com as melhores apresentações podem ser vendidas in natura e as demais podem ser direcionadas à produção de etanol ou outras destinações.

3.2 Produção de Batata Doce

O cultivo da batata-doce é considerado relativamente simples. Ela é uma hortaliça de raiz tuberosa rica em amido e, geralmente, pouco ramificada. Dentre suas características estão a alta rusticidade e adaptação em diferentes tipos de solo e clima, além de possuir baixo custo de produção (Senff, 2018). De acordo com a EMBRAPA (2021), a Tabela 1 mostra as características das cultivares de batata-doce disponibilizadas no mercado brasileiro. Neste trabalho, focaremos na batata-doce brasileira roxa por ser a mais comumente encontrada nos supermercados e feiras do Rio Grande do Sul.

Tabela 1: Cultivares de batata-doce de acordo com a EMBRAPA.

Tipo	Cor da película externa	Formato das raízes
Beauregard	Alaranjada	Obovada
Brazilândia Branca	Branca	Alongado e uniforme
Brazilândia Rosada	Rosa	Alongado, cheio e uniforme
Brazilândia Roxa	Roxa	Alongado e uniforme
Brazilândia Amélia	Rosa	Fusiforme
BRS Cuia	Branca	Longo e cheio
BRS Rubissol	Roxa	Longo e cheio
Coquinho	Branca	Arredondado e desuniforme
Princesa	Branca a creme	Alongado e uniforme

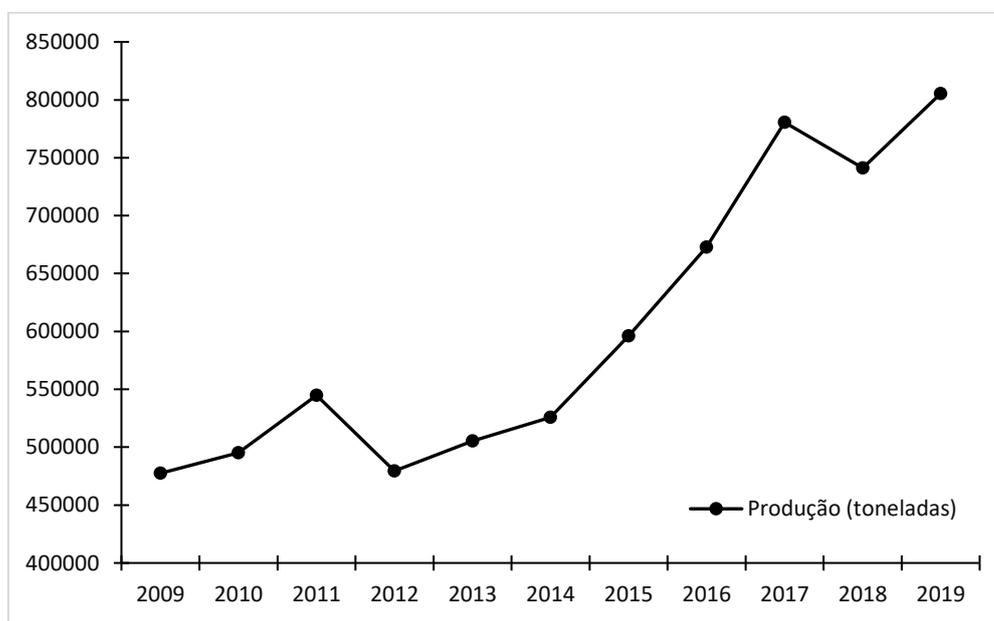
Fonte: EMBRAPA

A batata-doce é cultivada em mais de 100 países, sendo que aproximadamente 90% da produção é obtida na Ásia, 5% na África e 5% no restante do mundo. No que se refere ao volume de produção mundial, a batata-doce ocupa o sétimo lugar, mas é a décima quinta em custo de produção, o que indica ser universalmente uma cultura de baixo custo de produção

(Freitas et al., 2007). De acordo com estudos feitos pela Confederação da Agricultura e Pecuária do Brasil (CNA), o agronegócio brasileiro representou 21,4% do PIB nacional em 2019. Esse setor representa uma parcela importante da economia e desenvolvimento do Brasil, justificando o investimento e o incentivo às pesquisas relacionadas.

Em função da amplitude de ambientes em que a batata-doce pode ser plantada, sua produção está presente de norte a sul do Brasil e seu cultivo é viável ao longo de todo o ano na maior parte do país. Os últimos dados da Produção Agrícola Municipal (PAM) do IBGE indicam que, em 2019, a produção total foi de 805,4 mil toneladas de batata-doce no Brasil. Na Figura 3, podemos observar que a partir de 2012 houve um grande aumento na produção.

Figura 3: Produção total de batata-doce no Brasil, em toneladas.



Fonte: Base de dados PAM, IBGE 2020.

Esse aumento da produção total reflete a alta demanda por raízes de batata-doce no mercado nacional e sua respectiva relevância no presente estudo. A Tabela 2 mostra a produção total e a produtividade da batata doce nas cinco regiões brasileiras, no ano de 2019. As principais regiões produtoras de batata-doce no Brasil são Norte (317,3 mil toneladas), Sul (252,9 mil toneladas) e Sudeste (213,9 mil toneladas). O estado que apresenta a maior produção nacional é o Rio Grande do Sul, com 175 mil toneladas, responsável por aproximadamente 31% da produção do país.

Tabela 2: Produção total e produtividade de batata-doce nas regiões brasileiras (2019)

Região	Produção (toneladas)	Produtividade (ton/ha)
Sul	252.936	16,3
Sudeste	213.972	16,6
Centro-oeste	10.598	20,7
Norte	10.641	14,9
Nordeste	317.265	11,5

Fonte: Base de dados PAM, IBGE 2020.

Apesar de ter ocorrido um aumento na produção de batata-doce nos últimos anos, a falta de qualidade não acompanhou este aumento de produção devido uma combinação de vários fatores, incluindo mutação, vírus e outros patógenos, em doenças da batata-doce ou práticas de produção inadequadas (Misra e Jeeva, 2012). Além disso, os agricultores experimentam um problema na classificação das variedades de plantas devido à falta de acesso de especialistas agrícolas. Conseqüentemente, resultando em baixo rendimento (Patil e Kumar, 2017).

O conhecimento das características do processamento e os padrões de qualidade estabelecidos pelo mercado consumidor é fundamental para o desenvolvimento das indústrias deste segmento, a qual se inicia na escolha da matéria-prima e termina com o transporte do produto acabado. Dessa forma, a classificação desta mercadoria é significativa para garantir a qualidade dos produtos antes de chegar à comercialização (Arruda, 2004). De acordo com a EMBRAPA, no Brasil não há uma norma oficial para classificação, mas as normas extraoficiais utilizadas nos principais mercados consumidores são bastante exigentes. As batatas-doces devem ser bem conformadas, lisas, de formato alongado e uniforme, com película da cor específica de cada variedade e isentas de pragas e doenças (Kalkmann, 2011). Portanto, uma classificação prévia antes de sua comercialização promoverá uma melhor aceitação do consumidor e conseqüentemente menor desperdício, pois não apodrecerão e nem serão descartadas.

Segundo o pesquisador Marcos Fonseca, da Embrapa Agroindústria de Alimentos, há medidas simples que podem ser implantadas para minimizar as perdas e o desperdício. Fonseca ainda fala que há dois tipos de perda, a qualitativa, que se refere ao alimento feio, rejeitado, embora ótimo para consumo. E a quantitativa, que é aquela que acontece nas centrais de abastecimento, onde o que não é vendido acaba ficando por lá mesmo (Bello, 2016). Neste trabalho, vamos nos ater ao primeiro aspecto, a qualitativa, o qual é atrelado ao segundo aspecto.

3.3 Técnicas de Classificação de Imagens

Avanços recentes em técnicas de visão computacional têm possibilitado uma ampla gama de aplicações em diversos segmentos. Uma dessas áreas de aplicação é a classificação de imagens, que vem ganhando destaque em muitas aplicações da indústria agrícola (Kuang et al., 2018). Tais técnicas automatizadas proporcionam uma maior agilidade e acurácia quando aplicadas em larga escala, substituindo a seleção manual e outros métodos tradicionais. A classificação manual apresenta muitas restrições relacionadas ao ser humano como, por exemplo, um indivíduo precisa estar familiarizado com as muitas características do produto de interesse e ser consistente e suas avaliações, sendo esta operação bastante demorada e onerosa. Nos últimos anos, inúmeras técnicas de aprendizado de máquina vem sendo exploradas para classificação de frutas e vegetais, como por exemplo, máquina de vetores de suporte (SVM), K-vizinho mais próximo (KNN), árvores de decisão, redes neurais artificiais (ANN) e redes neurais convolucionais (CNN) (Hameed et al., 2018). Para esses métodos automatizados, as imagens geralmente devem passar por quatro etapas principais, incluindo pré-processamento, segmentação, extração de recursos e classificação (Iqbal et al., 2018).

Em 2018, Kamilaris e Prenafeta-Boldi (2018) publicaram uma análise bibliográfica sobre 40 artigos de pesquisa que aplicavam técnicas de *Deep Learning* (DL) para abordar vários problemas agrícolas. Dentre os métodos tradicionais existentes, a autora destaca que o mais comum é o baseado em satélite, usando imagens multiespectrais e hiper espectrais. Já as menos utilizadas são as técnicas baseadas em radar de abertura sintética (SAR), câmeras térmicas e infravermelho próximo (NIR), enquanto imagens ópticas e de raios-X estão sendo aplicadas na classificação de frutas e alimentos embalados. As técnicas que utilizam redes neurais do tipo redes neurais convolucionais (CNN) são utilizadas em 42% dos artigos pesquisados. Como conclusão dos estudos apresentados, as técnicas de DL superaram os métodos tradicionais de processamento de imagem.

Dentro do paradigma da aprendizagem profundo, Mercurio e Hernandez (2019) propuseram um método para se classificar batatas-doces produzidas nas Filipinas em várias categorias com base na textura de cozimento e funcionalidade de ingredientes. Foi utilizado um conjunto de 3.000 imagens de cinco espécies de batata-doce, sendo que cada espécie continha 600 imagens. Os conjuntos de treinamento consistiam em 70% de imagens (420 por variedade), 20% de imagens para validação e 10% (60 por variedade) para imagens de teste. Os resultados do estudo mostraram que o modelo pré-treinado de identificação da variedade de batata-doce com base na CNN, por meio do ResNet34 que é o modelo pré-treinado, atingiu 96,33% de acurácia na classificação. Os resultados indicados demonstram uma ótima possibilidade de classificar batatas-doces através do uso de redes neurais convolucionais (Mercurio e Hernandez, 2019).

Em Selvaraj et al. (2019), técnicas de visão computacional e aprendizado de máquina foram aplicadas para a detecção de doenças diferentes em banana. Os autores propuseram um sistema de detecção de doenças e pragas da banana baseado em inteligência artificial usando uma rede neural convolucional (CNN) para apoiar produtores de banana. Foi utilizado uma base de dados de 18.000 imagens pré-selecionadas da doença da banana e sintomas de pragas / danos foram coletados, sendo classificadas em 18 classes. Seis modelos diferentes foram testados com redes neurais pré-treinadas, a ResNet50 e a InceptionV2, que obtiveram um desempenho melhor em comparação com MobileNetV1. Essas arquiteturas representam

os resultados de ponta de doenças da banana e detecção de pragas com uma acurácia de mais de 90% na maioria dos modelos testados (Selvaraj et al., 2019).

Nasirahmadi e Ashtiani (2017) avaliaram técnicas de visão computacional para a classificação de 20 variedades de amêndoas, totalizando 2.000 imagens. Este estudo avaliou o desempenho de dois conjuntos de classificadores: o classificador baseado em distância ou K-vizinho mais próximo (KNN - *K-Nearest Neighbor*) e o classificador baseado em aprendizagem (SVM - *Support Vector Machine*), e foram comparados usando os resultados da classificação dos modelos. KNN pontua as características de uma categoria com base na categoria de vizinhos k e é comumente aplicado para avaliar e comparar o desempenho de outros classificadores devido à sua simplicidade e eficácia de classificação. O SVM basicamente constrói um hiperplano em um espaço dimensional elevado que separa os dados em duas categorias para classificação. Ao comparar os classificadores utilizados, observou-se que os classificadores SVM tiveram um melhor desempenho, com acurácia de 79% a 91%, do que o classificador KNN, com acurácia de 67% a 79% (Nasirahmadi e Ashtiani, 2017).

Kusrini et al. (2020), propuseram um sistema de identificação automatizado de pragas em campos de produção de manga através de técnicas de visão computacional e aprendizagem profunda. O modelo de rede pré-treinada utilizada foi a rede VGG-16. Como o conjunto de dados era limitado contendo 510 imagens originais, os autores também testaram outro conjunto que propuseram um processo de aumento de imagens para validar a ideia de que se obteria um melhor desempenho. A combinação desses conjuntos de dados inclui a classificação dos dados originais sem processo de aumento, resultando em 67% de acurácia, enquanto a avaliação no processo com dados aumentados resultou em 76% de acurácia.

Após o levantamento de uma vasta gama de artigos que possuíam o objetivo de classificar imagens, foi definido que a rede neural convolucional (CNN - *Convolutional Neural Network*) seria utilizada nos nossos experimentos. Segundo a literatura, a CNN apresenta uma elevada capacidade e eficácia de classificação de objetos e imagens quando comparada às demais redes. Entre os artigos estudados, esse tipo de rede obteve resultados muito confiáveis, alta acurácia, baixa taxa de erro e excelente eficiência de treinamento (Mercurio e Hernandez, 2019).

4 Materiais e Métodos

4.1 Software e Hardware

A Tabela 3 apresenta a lista de hardware e software empregado na realização deste trabalho. Para implementação de algoritmo e scripts de dados, foi utilizado a plataforma do Google Colaboratory (Colab), que é um serviço de nuvem gratuito hospedado pelo Google para incentivar a pesquisa de Aprendizado de Máquina e Inteligência Artificial. O ambiente de execução foi a GPU ou unidade de processamento visual, que é o nome dado a um tipo de microprocessador especializado em processar gráficos em computadores pessoais. É uma unidade de computação que oferece uma grande vantagem em comparação com uma unidade de processamento central (CPU) (Cengil et al., 2017). Todas as implementações neste trabalho foram realizadas em linguagem Python e fez uso de Keras para a adaptação das redes neurais.

Tabela 3: Listas de hardware e software usados neste estudo

Modelo	Tamanho
Memória RAM (Colab)	12 Gb
GPU	NVIDIA Tesla K80
Biblioteca Deep learning	Tensorflow 2.4.1
Linguagem de programação	Python 3.7.10

4.2 Aquisição de Imagens

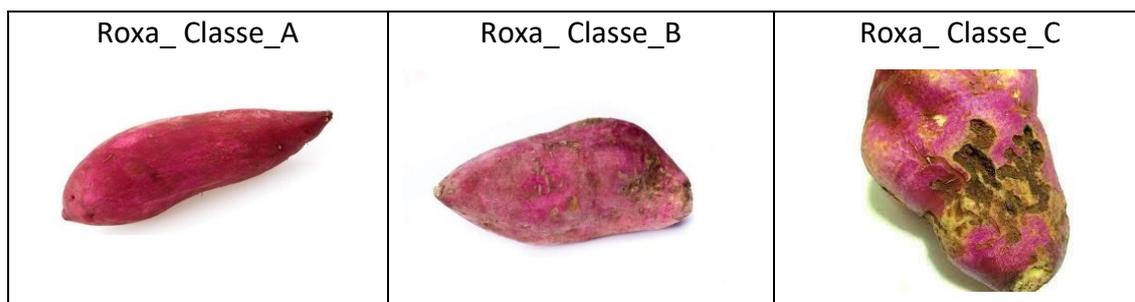
As imagens foram capturadas por uma câmera digital em mercados da região e algumas imagens da internet, utilizando diferentes angulações. A abordagem proposta foi testada usando um conjunto de dados com 356 imagens de batata-doce brazlândia roxa. Para cada uma das imagens, foi rotulado três subclasses para treinar o modelo de aprendizado profundo de maneira supervisionada, sendo elas:

- Alta probabilidade de compra (Classe “A”);
- Média probabilidade de compra (Classe “B”);
- Baixa probabilidade de compra (Classe “C”);

As imagens classificadas como “A”, são aquelas batatas-doces que possuem um formato alongado e uniforme, com coloração homogênea e casca sem machucados. Já as de classe “B”, possuem uma forma larga e cheia (não alongada) ou que possuem alguns machucados na casa, mas que estão próprias para consumo. E por fim, as de classe “C” são aquelas que

possuem um formato muito distinto do aceitável, não uniforme e com a casca machucada ou visualmente imprópria para consumo. Abaixo, a Figura 4 ilustra a base de dados utilizada para cada uma das classes, totalizando 3 subclasses.

Figura 4: Exemplos do conjunto de imagens das duas espécies de batata-doce utilizadas, por classe



4.3 Pré-Processamento das Imagens

A etapa de pré-processamento é importante para melhorar a qualidade da imagem de entrada e facilitar o treinamento da rede. A abordagem proposta usa um conjunto de 356 imagens de diferentes fontes, e por isso, várias delas possuem ruídos, incluindo fundo de baixa qualidade, diferentes tamanhos e contrastes, o que afeta a acurácia da segmentação. Para padronizar as entradas, as imagens selecionadas passaram por processamento de redimensionamento e posterior normalização.

Em relação aos parâmetros de entrada, definimos a largura e a altura da imagem de entrada em 600 e 400 pixels, respectivamente, através de um processo de redimensionamento. Também utilizamos a técnica de normalização para que todos os dados tenham a mesma escala. Sabemos que quando estamos trabalhando com imagens, os valores de cada pixel podem variar de 0 a 255. Sendo assim, para normalizarmos a escala, dividimos todos os valores por 255, significando o fator de multiplicação para cada pixel da imagem. Além disso, utilizamos um ajuste fino de 100% para as camadas de rede com os pesos obtidos do ImageNet, que é um dataset do Google com 1.4 milhões de imagens e mais de 1000 classes. E por fim, para mitigar o impacto das várias sobrecargas de extração de dados, foi utilizado também o recurso do “Autotune”, que tem a função de acelerar o desempenho do código (Roman, 2020).

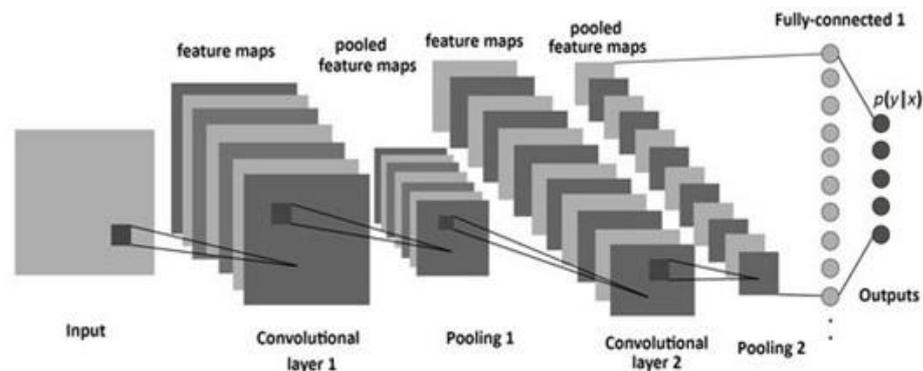
4.4 Segmentação das Imagens

De acordo com Nasirahmadi e Ashtiani (2017), a maioria dos 40 artigos de pesquisa estudados dividiu o seu conjunto de dados entre dados de treinamento e teste usando uma proporção de 80–20 ou 90–10, respectivamente. Tendo como base essa referência, dividimos o total de imagens coletadas em 80% para o treinamento (285 imagens) e 20% para o teste (71 imagens), que também serviu como validação do modelo.

4.5 Transferência de Aprendizado

Em nossos experimentos, treinamos nosso modelo com uma rede pronta, utilizando um processo chamado Transfer Learning ou Transferência de Aprendizado. Com Transfer Learning, utilizamos o método baseado em rede neural convolucional (CNN) já treinada com imagens rotuladas para aprender como identificá-las e classificá-las de acordo com os padrões visuais. No nosso caso, usamos implementações de código aberto de quatro modelos de aprendizagem profunda reconhecidos em cima da base de dados do ImageNet: ResNet-50, Xception, MobileNet e InceptionV3. Além destes, também foi proposto um modelo mais simples chamado de “Meu Modelo”. A Figura 5 mostra um esboço da arquitetura de uma rede neural convolucional utilizada para a classificação das imagens.

Figura 5: Arquitetura da Rede Neural Convolucional



Fonte: Albelwi, S., & Mahmood, A., 2017

Primeiramente, colocamos como input a nossa base de dados. A CNN irá transformar a camada de entrada através da multiplicação por uma matriz de convolução e vai extrair as características mais importantes da imagem de forma numérica. Dessa forma, ela reduzirá a dimensionalidade e subseqüentemente, aplicará funções de convolução para tornar o resultado mais padronizado, que consiste em duas etapas de convolução e duas etapas de Pooling. Em cada uma das camadas, utilizaremos a função Unidade Linear Retificada (ReLU - *Rectified Linear Unit*), que é a função de ativação mais amplamente utilizada ao projetar redes neurais atualmente. Ela é utilizada como se um “filtro” passasse sobre a imagem, varrendo alguns pixels por vez e criando um mapa de características que prevê a classe à qual cada recurso pertence. A próxima etapa é a camada de pooling (*downsampling*), que reduz a quantidade de informação em cada característica reconhecida na camada convolucional e mantém as informações mais importantes. Na última cada, optamos por utilização a função SoftMax para padronizar e amenizar a distribuição na saída. Essa função força a saída de uma rede neural a representar a probabilidade de os dados serem de uma das classes definidas. Sem ela as saídas dos neurônios são simplesmente valores numéricos onde o maior indica a classe vencedora. Após a função de ativação, os resultados passam por uma função objetivo ou de perda, que é atualizada através de um processo de otimização. As funções de otimização

são associadas à minimização da função perda, que neste caso está associada a acurácia de classificação das imagens.

O tamanho do lote (batch), cuja definição é o número de imagens de treinamento que passará pela rede neural por vez, foi definido igual a 9. Outro parâmetro é o número de épocas, que representam o número de vezes que o algoritmo passará a informação pela rede para ela aprender as características, definido como 50. Para a avaliação dos resultados, testamos a utilização de dois otimizadores, Adam e SGD, com taxa de aprendizagem, cuja definição é a rapidez com que uma rede aprende, de 0,001. Esse valor foi escolhido diante dos valores taxas de aprendizagem mais utilizados, que são de 0,001 e 0,005 até 0,01 (Kamilaris e Prenafeta-Boldú, 2018).

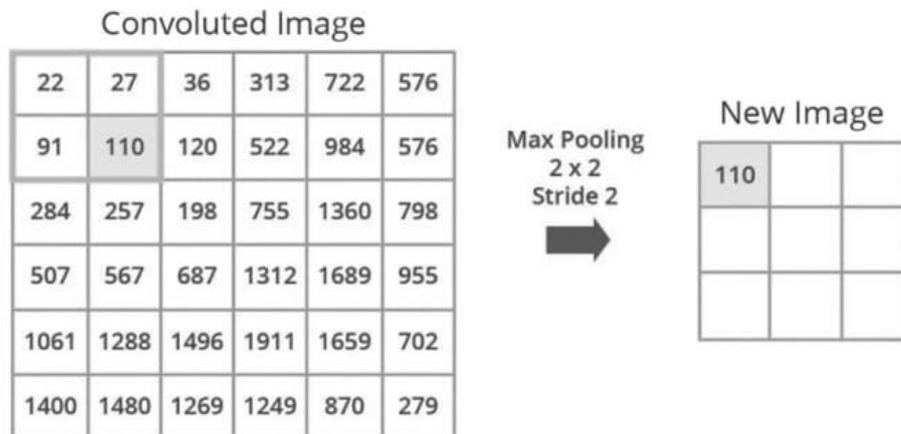
Nesse processo de treinamento da rede neural, há dois problemas que podem acontecer, que são o *overfitting* e o *underfitting*. O *overfitting* é o estado quando um modelo de aprendizagem de máquina se adapta demais aos dados imputados de treinamento, ou seja, a rede perde seu potencial de generalização. Já o *underfitting* ocorre quando o modelo de aprendizagem não se adapta o suficiente aos dados de treinamento. Uma maneira de controlar e evitar que esses problemas ocorram é sempre realizar uma análise de teste para verificar se os dados estão dentro do esperado.

4.6 Extração de recursos

A extração de recursos desempenha um papel importante no campo da visão computacional e do aprendizado de máquina para a descrição de um objeto na imagem de entrada. Através dos pixels de cada imagem, essa técnica identifica cores diferentes e as convertem em valores numéricos, para posterior procura de padrões, como cores e texturas similares. (Marengoni e Stringhini, 2010). Cada imagem tem sua própria forma, movimento, tamanho e cor. Portanto, por meio da extração de recursos, o objeto extraído é classificado em sua classe de relevância (Iqbal et al., 2018). Ainda segundo Iqbal, o ponto crítico na fase de extração de recursos é a seleção dos recursos mais proeminentes, pois os recursos irrelevantes e redundantes afetam a acurácia e o tempo de execução do sistema. Neste trabalho, foram extraídas características físicas das imagens para que fosse possível sua posterior classificação, tais como forma, tamanho, cor e textura.

Para a extração das principais características, é comum reduzir a dimensão espacial das imagens ao longo das camadas da rede. Uma das camadas presentes dentro da CNN é a operação de maxpooling, sendo a mais comumente empregada (De Oliveira et al., 2010). Essa operação tem a função de diminuir o número de parâmetros para reduzir o tamanho das imagens através da extração das principais características gerais, comuns entre o conjunto de imagens. A Figura 6 representa um exemplo de um mapa de características com tamanho 6x6 e o resultado de uma operação de maxpooling utilizando uma janela 2x2.

Figura 6: Representação de um *maxpooling*



Fonte: Balachandran, 2020

4.7 Classificação

Após a etapa de extração de características, temos a etapa de classificação da imagem. Primeiro, passa por uma camada que se chama Flatten, onde ela recebe a saída das camadas anteriores, e faz um processo de achatar a imagem e transformação em um único vetor para introduzir como entrada na próxima camada. O próximo estágio é uma camada totalmente conectada, onde ela aplica pesos para prever a classificação correta. E por último, aplicamos uma camada de Softmax. Essa camada fornece as probabilidades finais para cada classificação.

Na tarefa de classificação, o método avalia a família da batata-doce roxa e a sua aceitação no mercado consumidor. Nossa metodologia avalia o desempenho de quatro redes neurais profundas bem conhecidas da literatura já mencionadas no item 4.5. Inicialmente, a transferência de aprendizagem da rede utiliza os pesos obtidos no ImageNet e a parte de classificação são utilizados pesos inicializados aleatoriamente.

4.8 Medidas de desempenho

Diversas medidas de desempenho são utilizadas na literatura para avaliação de algoritmos. Como indicadores para a análise dos resultados, avaliamos os valores de acurácia do treinamento e validação da acurácia para as imagens de teste, o tempo de treinamento dos modelos de aprendizagem profunda na tarefa de classificação e finalmente, o tempo que o modelo leva para prever a classe.

A acurácia é definida como a proporção de amostras que são classificadas corretamente de acordo com seus rótulos de classe. Após o treinamento da rede neural, utilizamos o conjunto de amostras de teste para validação do modelo. Ele é usado para ajustar os hiper parâmetros da rede e comparar como as mudanças neles afetam a acurácia preditiva do modelo. É útil porque permite a comparação repetida desses diferentes parâmetros e arquiteturas com os mesmos dados e pesos de rede, para observar como essas alterações

afetam o poder preditivo da rede em cima dos dados de teste. O tempo que o modelo leva para prever é o fator mais importante para cenários de tempo real, portanto vários pesquisadores calculam o tempo para verificar a eficiência de seus métodos e será também avaliado no presente trabalho. Além disso, também calculamos uma matriz de confusão (CM), para comparar rótulos de categorias previstos com o rótulo verdadeiro.

5 Resultados

Neste Capítulo são apresentados e discutidos os resultados obtidos pela metodologia abordada no Capítulo 4. Trata-se da análise do desempenho dos modelos calibrados, bem como os componentes da ferramenta de rede neural artificial, quais sejam, determinações de neurônios das camadas, entre outros elementos necessários. Por fim, examinou-se os resultados obtidos pela ferramenta e o efeito no processo de classificação das batatas-doces.

5.1 Modelo

Para a determinação do melhor modelo, foram testadas várias configurações de rede neurais para a etapa de classificação e dois otimizadores, SGD e Adam, sendo que o otimizador SGD obteve melhores resultados. A Figura 7 representa o resumo de um dos modelos utilizados, o Xception, onde a explicação mais detalhada pode ser encontrada em <https://keras.io/api/applications/>. Nesse caso, estamos trabalhando com um número total de 85.611.563 parâmetros. Isso quer dizer que cada época executada, são atualizados todos os parâmetros.

Figura 7: Resumo do modelo XCEPTION

Layer (type)	Output Shape	Param #	Connected to
input_2 (InputLayer)	[(None, 400, 600, 3)] 0		
block1_conv1 (Conv2D)	(None, 199, 299, 32) 864		input_2[0][0]
block1_conv1_bn (BatchNormaliza	(None, 199, 299, 32) 128		block1_conv1[0][0]
(...)			
conv2d_3 (Conv2D)	(None, 13, 19, 1024) 745472		add_10[0][0]
block13_pool (MaxPooling2D)	(None, 13, 19, 1024) 0		block13_sepconv2_bn[0][0]
batch_normalization_3 (BatchNor	(None, 13, 19, 1024) 4096		conv2d_3[0][0]
add_11 (Add)	(None, 13, 19, 1024) 0		block13_pool[0][0] batch_normalization_3[0][0]
block14_sepconv1 (SeparableConv	(None, 13, 19, 1536) 1582080		add_11[0][0]
block14_sepconv1_bn (BatchNorma	(None, 13, 19, 1536) 6144		block14_sepconv1[0][0]
block14_sepconv1_act (Activatio	(None, 13, 19, 1536) 0		block14_sepconv1_bn[0][0]
block14_sepconv2 (SeparableConv	(None, 13, 19, 2048) 3159552		block14_sepconv1_act[0][0]
block14_sepconv2_bn (BatchNorma	(None, 13, 19, 2048) 8192		block14_sepconv2[0][0]
block14_sepconv2_act (Activatio	(None, 13, 19, 2048) 0		block14_sepconv2_bn[0][0]
flatten_1 (Flatten)	(None, 505856)	0	block14_sepconv2_act[0][0]
dense_2 (Dense)	(None, 128)	64749696	flatten_1[0][0]
dense_3 (Dense)	(None, 3)	387	dense_2[0][0]
Total params: 85,611,563			
Trainable params: 85,557,035			
Non-trainable params: 54,528			

Fonte: Autora

A título de comparação, o “Meu Modelo” é um modelo proposto bem mais simples que os demais, representado na Figura 8. Nele, estamos trabalhando com um número total de 277.827 parâmetros, ou seja, 308 vezes menos parâmetros que a Xception, demonstrando um ganho na velocidade da classificação e memória de máquina. Apesar de ter demonstrado uma acurácia inferior e uma perda maior em relação aos dados de teste, ainda assim pode ser considerado como um modelo promissor de previsão de classificação de batatas-doces roxa.

Figura 8: Resumo do “Meu modelo”

```

Model: "sequential"
-----
Layer (type)                Output Shape                Param #
-----
conv2d_98 (Conv2D)          (None, 398, 598, 32)      896
max_pooling2d_4 (MaxPooling2 (None, 199, 299, 32)      0
conv2d_99 (Conv2D)          (None, 197, 297, 64)     18496
max_pooling2d_5 (MaxPooling2 (None, 98, 148, 64)      0
conv2d_100 (Conv2D)         (None, 96, 146, 128)     73856
max_pooling2d_6 (MaxPooling2 (None, 48, 73, 128)      0
conv2d_101 (Conv2D)         (None, 46, 71, 128)     147584
max_pooling2d_7 (MaxPooling2 (None, 23, 35, 128)      0
conv2d_102 (Conv2D)         (None, 21, 33, 32)       36896
dropout (Dropout)           (None, 21, 33, 32)       0
global_average_pooling2d (Gl (None, 32)                0
dense_8 (Dense)              (None, 3)                 99
-----
Total params: 277,827
Trainable params: 277,827
Non-trainable params: 0

```

Fonte: Autora

5.2 Desempenho

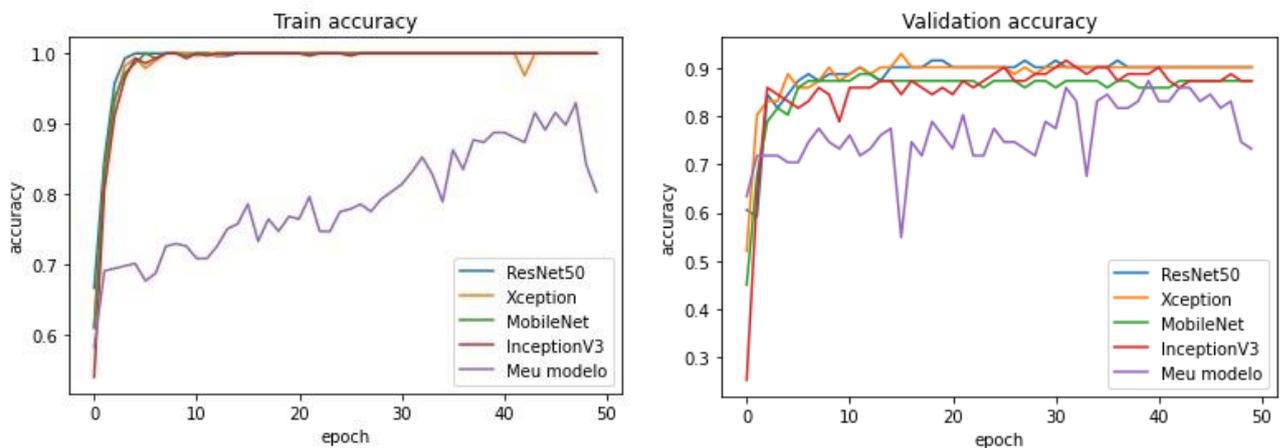
A Tabela 4 mostra os resultados da acurácia e validação obtidos pelos modelos de aprendizagem profunda. A tabela também mostra o tempo de treinamento de cada época, em segundos, o tamanho ocupado por cada rede, e o número de parâmetros. Nossa abordagem alcançou o melhor resultado com a Xception, obtendo uma validação de acurácia de 90,14% para as imagens testadas. Apesar desse resultado ainda ser preliminar, o mais importante é errar menos em relação ao que é enviado para o consumidor final, que será mostrado na matriz de confusão.

Tabela 4: Métricas de Desempenho usadas para avaliar os modelos de aprendizagem profunda

Modelo	Tamanho	Nº de parâmetros	Tempo de cada época	Acurácia de Teste
Resnet50	98 MB	88.337.795	21 s	90,14%
Xception	88 MB	85.611.563	46 s	90,14%
Mobile Net	16 MB	31.540.931	13 s	86,32%
InceptionV3	92 MB	70.825.001	16 s	87,32%
Meu modelo	1,06 MB	277.827	3 s	73,24%

Na Figura 9, podemos observar os gráficos da acurácia das imagens de treinamento e de teste, que vai aumentando com o passar das épocas. Para o gráfico de acurácia de treinamento, com exceção do “Meu Modelo”, a partir de 5 épocas é atingido 100% de acurácia para as quatro redes pré-treinadas utilizadas: Resnet50, Xception, InceptionV3 e MobileNet. O gráfico de acurácia de validação, utilizando as imagens de grupo de teste, podemos observar que, para aquelas mesmas quatro redes, com 20 épocas já se atinge um patamar de acurácia. Logo, a partir dos resultados atingidos vemos que os percentuais de acurácia de validação podem ser melhorados aumentando a base de dados, já que estes percentuais ficaram abaixo da acurácia de treinamento. Apesar do pequeno conjunto de imagens, os resultados parciais mostram um ótimo indicativo da abordagem empregada.

Figura 9: Gráficos da Acurácia de treinamento e teste, pelo nº de épocas



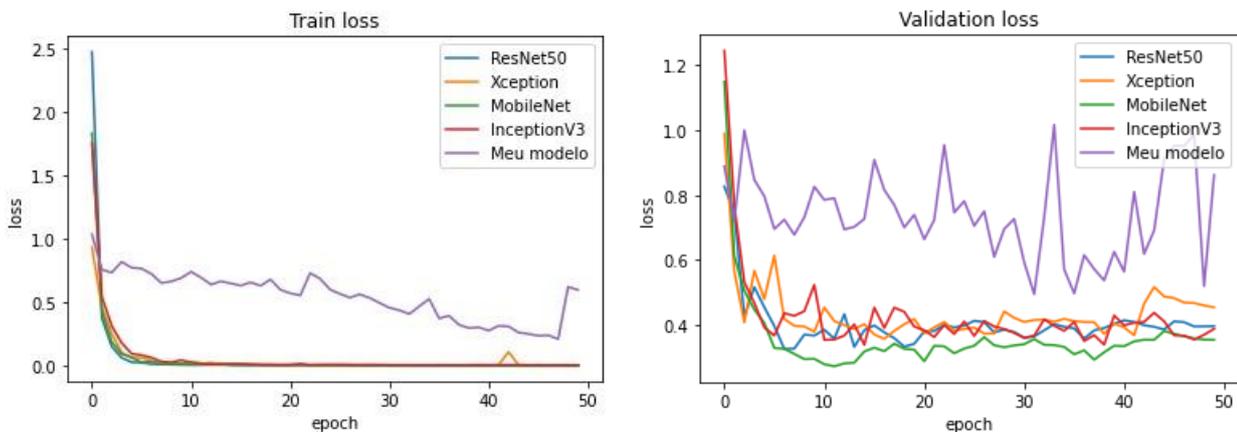
Fonte: Autora

Na literatura, Selvaraj et al. (2019) obteve resultados excelentes usando o modelo Inception para detecção de doenças e pragas da banana com uma acurácia de mais de 90%.

No presente trabalho, obtivemos este mesmo percentual com o modelo Xception e Resnet50, demonstrando o excelente potencial da técnica utilizada.

Abaixo, é mostrado na Figura 10 a perda na validação dos modelos das imagens de teste. O gráfico mostra alguns ruídos, principalmente no “Meu modelo”, onde os valores de perda são mais elevados. Apesar disso, a partir da quinta época, todos os modelos apresentam perdas inferiores a 1.

Figura 10: Gráficos da perda de validação, pelo nº de épocas



Fonte: Autora

A Figura 11, mostra a matriz de confusão do modelo Xception para as 71 imagens testadas, através de um layout de tabela específico que permite a visualização do desempenho de um algoritmo.

Figura 11: Matriz de confusão para os dados de teste do modelo Xception

		PREDIÇÃO		
		Classe A	Classe B	Classe C
REAL	Classe A	25 (35,21%)	0 (0,0%)	0 (0,0%)
	Classe B	4 (5,63%)	23 (32,39%)	1 (1,41%)
	Classe C	1 (1,41%)	1 (1,41%)	16 (22,54%)

Fonte: Autora

O importante neste modelo é evitar os falsos positivos, pois estaríamos enviando batatas-doces que provavelmente o consumidor não iria comprar. O falso positivo ocorre quando o modelo prevê que a classe da batata é a A (mais provável de ser comprada), mas a classe verdadeira é a B ou a C (menos provável de ser comprada). Ou seja, é quando o modelo exalta a classe da batata. Como podemos observar, o modelo possui um erro de 7,04% para este tipo de situação, pois como a amostragem ainda é muito pequena, o modelo fica mais suscetível a este tipo de erro. Isso poderá ser melhorado com o aumento da nossa base de dados, pois poderá ser possível obter uma acurácia de validação mais próxima de 100%. Já o falso negativo seria quando prevemos uma classe inferior à classe verdadeira. Neste caso, não seria um problema, pois de qualquer forma, a batata poderia ser utilizada para outra finalidade, como já abordado ao longo do trabalho.

O tempo que o modelo leva para prever a classe de uma imagem é de 1,29 milissegundos. Como um exemplo de aplicação prática, o agricultor classificaria de forma extremamente rápida e teria a seu favor o uma excelente ferramenta de tomada de decisão. Ou seja, quando o mercado de consumo de batatas-doce estiver absorvendo alta quantidade, ele poderia enviar ao mercado as batatas de classe A e B. Caso contrário, um mercado em baixa, ele poderia selecionar apenas as de classe A e as restantes seriam destinadas à produção etanol, álcool em gel e bebida destilada shochu através do conceito de Biorrefinarias Descentralizadas e Autônomas. Isso permitirá o aproveitamento total e fechamento do ciclo de produção.

6 Conclusões e Trabalhos Futuros

Técnicas utilizando a inteligência artificial vem sendo cada vez mais estudadas e praticadas através de novos algoritmos e softwares. O campo de desenvolvimento e aplicação dessa técnica tem crescido aceleradamente e é uma das competências mais procuradas na atualidade. Seu impacto na sociedade será, sem dúvida, enorme. Neste projeto, propusemos o desenvolvimento de uma metodologia capaz de classificar imagens de batata-doce brasileira roxa através de técnicas de visão computacional e aprendizado de máquina. As soluções propostas traduzem a simbiose indústria-setor alimentício, avançando a ciência e a tecnologia por trás da classificação automática aplicada ao cenário industrial e promovendo a completa utilização dessa hortaliça.

Os resultados experimentais mostram uma eficácia promissora dos modelos pré-treinados, resultando em uma acurácia de 90,14% com o modelo Xception. Com esta solução, promovemos uma ferramenta eficiente de tomada de decisão da produção de batata-doce. A partir do momento que esta classificação deixa de ser feita manualmente e torna-se automática através de classes bem definidas, o agricultor poderá dar outros fins àquelas que seriam provavelmente rejeitadas ao consumidor final. Isso faz com que seja incentivado o empoderamento do agricultor familiar, em que ele poderá gerar outras rendas a partir da destinação, como por exemplo, à produção de etanol. Apesar dos resultados ainda serem preliminares com um baixo número de amostragem de imagens, o método desenvolvido apresenta ser uma alternativa promissora para a classificação de imagens de batata-doce.

Devido algumas limitações do estudo atual em relação ao conjunto de imagens, a acurácia da classificação do modelo de CNN acaba sendo afetada. Como parte de trabalhos futuros, vê-se a necessidade de firmar parcerias com indústrias ou agricultores que possam fornecer uma gama maior de imagens de batatas-doces. A partir do modelo criado, é possível colocá-lo em um aplicativo, como por exemplo no TensorFlow Lite para Android, para o reconhecimento das classes da batata-doce através do uso de um dispositivo eletrônico que possua câmera que poderia ser acoplado na linha de produção. Além disso, também poderá ser criado modelos para classificar todas as variedades de batatas-doces cultivadas no Brasil. Da mesma forma, avaliar novas abordagens de aprendizagem de máquina, visando melhorar os processos, parâmetros e resultados de classificação, é essencial para obter-se uma maior acurácia.

Como objetivo final, o aperfeiçoamento do método poderá proporcionar a implementação desta solução através de um protótipo operacional que classificará as imagens em tempo real, com sensores na linha de produção. Isso permitirá maior eficiência do processo de classificação e uso efetivo dos produtos. O desenvolvimento do sistema é primordial para possibilitar a automatização do processo de classificação de batata-doce. Com avanços dessa natureza alinhados com os princípios da Indústria 4.0, poderemos garantir a competitividade do Brasil no cenário econômico mundial. Existem várias possibilidades de utilização dos recursos da biomassa, mas o conhecimento de biorrefinarias podem trazer novas competências e oportunidades de trabalho tanto em países industrializados quanto em desenvolvimento, mas os processos precisam ser projetados de forma sustentável.

REFERÊNCIAS

- ARRUDA, C. R. Análise Das Etapas Do Processamento De Batata Chips. 2004.
- BASTOS, E. N. F. Uma Rede Neural Auto-Organizável Construtiva para Aprendizado Perpétuo de Padrões Espaço-Temporais. p. 137, 2007.
- BALACHANDRAN, Sandeep. Machine Learning - Max & Average Pooling. 2020.
- BELLO, Liliane. Perdas e desperdício de alimentos reduzem sustentabilidade. Embrapa Agrobiologia, 2016.
- BORGES, F. C. Proposta de Um Modelo Conceitual de Biorrefinaria Com Estrutura Descentralizada. p. 105, 2010.
- BROWNLEE, Jason. Deep Learning for Computer Vision: Image Classification, Object Detection and Face Recognition in Python, v. 1.8, p. 82, 2020.
- BUENO, P. H. T. Panorama geral das perdas e desperdício de alimentos e soluções para o acesso à alimentação. 2019.
- CENGIL, E.; ÇINAR, A.; GÜLER, Z. A GPU-based convolutional neural network approach for image classification. **IDAP 2017 - International Artificial Intelligence and Data Processing Symposium**, n. February, 2017.
- DE OLIVEIRA, A. C. S.; DE SOUZA, A. A.; LACERDA, W. S.; et al. Application of artificial neural networks in the forecast of alcohol production. **Ciencia e Agrotecnologia**, v. 34, n. 2, p. 279–284, 2010.
- EMBRAPA. Cultivares de Batata-doce da Embrapa. Disponível em <<https://www.embrapa.br/cultivar/batata-doce>>. Acesso em 27 de mar. De 2021.
- FAO. Food wastage footprint: summary report. Food and Agriculture Organization. 2017.
- FAOSTAT. Food and Agriculture Organization of the United Nations - Statistics Division. Disponível em: <http://www.fao.org/faostat/en/#rankings/countries_by_commodity>. Acesso em 17 abr. 2021.
- FREITAS, G. M.; BUENO, G. S.; FERREIRA, J. D. F.; et al. ELABORAÇÃO E ANÁLISE SENSORIAL DE DOCE DE BATATA-DOCE ELABORATION AND SENSORIAL ANALYSIS OF CANDY OF SWEET. 2007.

HAMEED, K.; CHAI, D.; RASSAU, A. A comprehensive review of fruit and vegetable classification techniques. **Image and Vision Computing**, v. 80, p. 24–44, 2018.

IQBAL, Z.; KHAN, M. A.; SHARIF, M.; et al. An automated detection and classification of citrus plant diseases using image processing techniques: A review. **Computers and Electronics in Agriculture**, v. 153, n. July, p. 12–32, 2018.

KALKMANN, D. C. Produtividade, qualidade de raiz, resistência aos insetos de solo e aos nematóides-das-galhas, e estimativas de parâmetros genéticos em clones de batata-doce cultivados no distrito federal. 2011.

KAMILARIS, A.; PRENAFETA-BOLDÚ, F. X. Deep learning in agriculture: A survey. **Computers and Electronics in Agriculture**, v. 147, n. February, p. 70–90, 2018.

KUANG, H.; LIU, C.; CHAN, L. L. H.; et al. Multi-class fruit detection based on image region selection and improved object proposals. **Neurocomputing**, v. 283, p. 241–255, 2018.

KUSRINI, K.; SUPUTA, S.; SETYANTO, A.; et al. Data augmentation for automated pest classification in Mango farms. **Computers and Electronics in Agriculture**, v. 179, p. 105842, 2020.

LOURENÇO, A. M.; TAVARES, A. T.; FERREIRA, T. A.; et al. Potencial De Clones Experimentais De Batata-Doce Para Produção De Etanol. **Nativa**, v. 6, n. 4, p. 352, 2018.

MARENGONI, M.; STRINGHINI, S. Tutorial: Introdução à Visão Computacional usando OpenCV. **Revista de Informática Teórica e Aplicada**, v. 16, n. 1, p. 125–160, 2010.

MARQUES, K. R. Reação genótipos de batata-doce à isolados locais de *Plenodomus destruens* Harter, agente causal do mal-do-pé. p. 54, 2015.

MERCURIO, D. I.; HERNANDEZ, A. A. Classification of sweet potato variety using convolutional neural network. **2019 IEEE 9th International Conference on System Engineering and Technology, ICSET 2019 - Proceeding**, n. October, p. 120–125, 2019.

MISRA, V. H. R. S.; JEEVA, M. L. Sweet Potato Diseases : Diagnosis and Management. **Fruit, vegetable and cereal science and biotechnology**, v. 6, n. 1, p. 65–78, 2012.

NASIRAHMADI, A.; MIRAEI ASHTIANI, S. H. Bag-of-Feature model for sweet and bitter almond classification. **Biosystems Engineering**, v. 156, p. 51–60, 2017.

NATURAL RESOURCES INSTITUTE FINALND. **Biorefineries in decentralized**

environment. [s.l: s.n.].

ONU NEWS. Organização das Nações Unidas – Notícias. Disponível em: <<https://news.un.org/pt/story/2017/06/1589091-populacao-mundial-atingiu-76-bilhoes-de-habitantes>>. Acesso em 15 abr. 2021.

ORTEGA, E.; WATANABE, M.; CAVALETT, O. A produção de etanol em micro e mini-destilarias. **Biomassa para Energia**, p. 475–492, 2008.

PATIL, J. K.; KUMAR, R. Advances in Image Processing for Detection of Plant Disease. The SIJ Transactions on Computer Science Engineering & its Applications (CSEA), v. 05, n. 02, p. 08–10, 2017.

PIB do Agronegócio. Confederação da Agricultura e Pecuária do Brasil (CNA). Março de 2020. Disponível em: www.cnabrazil.org.br.

PRABHU. Understanding of Convolutional Neural Network (CNN) — Deep Learning. Mar, 2018.

ROMAN, V. Convolutional Neural Networks in Practice CIFAR-10 Dataset. p. 1–20, 2020.

SANDERS, J., Scott, E., Weusthuis, R., & Mooibroek, H. 2007. Bio-refinery as the bio-inspired process to bulk chemicals. *Macromolecular bioscience*, 7(2), 105-117.

SCHWEINBERGER, C. M. Inovação e otimização no processo de produção de etanol a partir de batata doce. p. 173, 2016.

SELVARAJ, M. G.; VERGARA, A.; RUIZ, H.; et al. AI-powered banana diseases and pest detection. **Plant Methods**, v. 15, n. 1, p. 1–11, 2019.

SEFFF, S. E. Integração Genótipos x Ambientes em Componentes Morfológicos e Produtivos de Batata-Doce (*Ipomoea batatas* L.), Universidade Federal de Santa Catarina, 2018.

TRIERWEILER, J. O.; TRIERWEILER, L. F., Farenzena, M. Biorrefinarias Descentralizadas e Autônomas: um Modelo de Economia Circular, Verde e Digital. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2020.

UNAJAN, M. C.; TABADA, W. M.; GERARDO, B. D.; et al. Sweet Potato (*Ipomoea batatas*) Variety Recognizer Using Image Processing and Artificial Neural Network. 2017.

ZANETTI, S. S.; SOUSA, E. F.; CARVALHO, D. F. DE; et al. Estimação da

evapotranspiração de referência no estado do Rio de Janeiro usando redes neurais artificiais. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 12, n. 2, p. 174–180, 2008.