PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS ESCOLA DE CIÊNCIAS EXATAS E DA COMPUTAÇÃO GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO



## DETERMINAÇÃO DE CONCENTRAÇÃO DE AMIDO EM CURCUMA LONGA UTILIZANDO DEEP LEARNING

WALACY XAVIER RODRIGUES

GOIÂNIA 2020

## WALACY XAVIER RODRIGUES

## DETERMINAÇÃO DE CONCENTRAÇÃO DE AMIDO EM CURCUMA LONGA UTILIZANDO DEEP LEARNING

Trabalho de Conclusão de Curso apresentado à Escola de Ciências Exatas e da Computação, da Pontifícia Universidade Católica de Goiás, como parte dos requisitos para a obtenção do título de Bacharel em Ciência da Computação. Orientador(a):

Prof. Dr. Arlindo Rodrigues Galvão Filho

Banca examinadora:

Prof. Dr. Clarimar José Coelho

Msc. Isaac Yves Lopes de Macêdo

Prof. Dr. Rafael Viana de Carvalho

GOIÂNIA 2020

## WALACY XAVIER RODRIGUES

## DETERMINAÇÃO DE CONCENTRAÇÃO DE AMIDO EM CURCUMA LONGA UTILIZANDO DEEP LEARNING

Trabalho de Conclusão de Curso aprovado em sua forma final pela Escola de Ciências Exatas e da Computação, da Pontifícia Universidade Católica de Goiás, para obtenção do título de Bacharel em Ciência da Computação, em 03/12/2020.

Prof.<sup>a</sup> Ma. Ludmilla Reis Pinheiro dos Santos Coordenador(a) de Trabalho de Conclusão de Curso

Banca examinadora:

Orientador: Prof. Dr. Arlindo Rodrigues Galvão Filho

Coorientador: Msc. Isaac Yves Lopes de Macêdo

Prof. Dr. Clarimar José Coelho

Prof. Dr. Rafael Viana de Carvalho

GOIÂNIA 2020

Dedico este trabalho à minha família, meus amigos, meu orientador e coorientador por terem me dado todo o apoio necessário para que eu chegasse até aqui.

## AGRADECIMENTOS

Ao Professor Arlindo Galvão, meu orientador, pelo apoio, paciência e as instruções e conhecimentos passados para que eu conseguisse concluir este trabalho.

Ao Isaac Yves, coorientador, pelo auxílio e por, junto com o Laboratório de Análise Farmacêutica e Ambiental da Universidade Federal de Goiás, disponibilizar as amostras para realização deste estudo.

Aos meus amigos Rafael Justino e Airton Chagas, por me acompanharem nessa trajetória acadêmica, por me auxiliarem nos estudos e por trabalharem comigo.

À minha família e demais amigos, que sempre me apoiaram e acreditaram no meu potencial, sem eles eu não conseguiria chegar até aqui.

Aos professores Clarimar José Coelho e Rafael Viana de Carvalho por participarem da banca.

A todos os integrantes do LCC por toda ajuda e pela ótima convivência durante essa jornada.

"Eu imagino um mundo onde a inteligência artificial nos permitirá ser mais produtivos, viver mais e ter energia mais limpa."

Fei-Fei Li

## RESUMO

O controle de qualidade em segmentos alimentícios e farmacêuticos é de suma importância para evitar riscos à saúde dos consumidores, além disso, reduzem as chances de causar prejuízos de produção. A metodologia clássica utilizada no controle de qualidade para alimentos e fármacos envolve processos químicos que podem ser onerosos. Portanto, o uso de métodos computacionais empregando imagens digitais é uma alternativa para contornar as limitações das metodologias clássicas. O pó de açafrão-da-terra (Curcuma longa) é um condimento utilizado mundialmente, diversos estudos o relacionam a vários benefícios à saúde. Porém, esse produto está sujeito a adulterações feitas com a diluição do pó com materiais mais baratos e de baixa qualidade, o amido está entre os materiais mais utilizados para diluir ao pó de açafrão. Portanto, este trabalho se propõe explorar o uso de métodos computacionais, mais especificamente a aplicação de redes neurais convolucionais para a determinação da concentração de amido em imagens digitais de amostras de açafrão. As redes testadas foram: AlexNet, GoogleNet, ResNet18, ResNet50 e ResNet101, dentre elas, a ResNet50 obteve melhores resultados com erro de raiz quadrático médio de 15,68, erro médio absoluto de 12,41 e coeficiente de determinação de 0,65.

**Palavras-Chave:** Açafrão-da-terra, Curcuma longa, Redes Neurais Convolucionais, ResNet, GoogleNet, AlexNet.

## ABSTRACT

Quality control in food and pharmaceutical segments is of paramount importance to avoid risks to the health of consumers, besides reducing the chances of causing production losses. The classic methodology used in quality control for food and drugs involves chemical processes that can be costly. Therefore, the use of computational methods employing digital images is an alternative to circumvent the limitations of classical methodologies. Turmeric powder (Curcuma longa) is a condiment used worldwide; several studies relate it to several health benefits. However, this product is subject to adulteration made with the dilution of the powder with cheaper and low-quality materials, the starch is among the materials most used to dilute the saffron powder. Therefore, this work proposes to explore the use of computational methods, more specifically the application of convolutional neural networks for the determination of starch concentration in digital images of turmeric samples. The networks tested were: AlexNet, GoogleNet, ResNet18, ResNet50 and ResNet101, among them, ResNet50 obtained better results with a root mean square error of 15.68, mean absolute error of 12.41 and coefficient of determination of 0.65.

*Keywords*: Turmeric, Curcuma longa, Convolutional Neural Networks, ResNet, GoogleNet, AlexNet.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Representação da arquitetura da LeNet-5	.16
Figura 2 - Filtro $3x3$ (em rosa) aplicado à uma imagem $5x5$ (em roxo) resultando er	n 17
	. 17
Figura 3 - Ilustração de duas camadas convolucionais	.18
Figura 4 - Exemplo de p <i>adding</i> de tamanho 1 e valor 0	.18
Figura 5 - Representação da camada de max-pooling com stride 2	.19
Figura 6 - Representação da fully connected	.20
Figura 7 - Arquitetura AlexNet	.21
Figura 8 - Módulo Inception	.22
Figura 9 - Exemplo de convolução 1x1: filtro 1x1x192 aplicado a um mapa de ativação 64x64x192 resultando em outro de tamanho 64x64x1 (com profundidade reduzida)	.23
Figura 10 - Módulo <i>Inception</i> : visão em 3 dimensões	.23
Figura 11 - Aumento do erro de treinamento e teste de redes ao aumentar camada de 20 para 56	as .24
Figura 12 - Bloco residual	.25
Figura 13 - ResNet152: arquitetura	.25
Figura 14 - Amostras de todos as marcas de açafrão (de A a J) na sua composição	0
pura	.28
Figura 15 - Amostras de açafrão da marca A em sua composição pura, 20:80, 40:6 60:40, 80:20 e amido puro, respectivamente	50, .28
Figura 16 - Foto da amostra I Puro	.29
Figura 17 - Demonstração de recorte 150x150 pixels na amostra de açafrão I puro	.30
Figura 18 - Demonstração de recorte 150x150 <i>pixel</i> s para 9 imagens de 50x50 <i>pixe</i> na amostra de açafrão I puro	els .30
Figura 19 - Demonstração de espelhamento na amostra de açafrão I puro	.31
Figura 20 - Demonstração de rotação e corte da imagem 150x150 da amostra de	
açafrão A puro	.31
Figura 21 - AlexNet: gráficos de predito pelo esperado	.37
Figura 22 - GoogleNet: gráficos de predito pelo esperado	.40
Figura 23 - ResNet18: gráficos de predito pelo esperado	.43
Figura 24 - ResNet50: gráficos de predito pelo esperado	.46
Figura 25 - ResNet101: gráficos de predito pelo esperado	.49

## LISTA DE TABELAS

Tabela 1 - AlexNet: predito pelo esperado e acurácia	35
Tabela 2 - GoogleNet: predito pelo esperado e acurácia	39
Tabela 3 - ResNet18: predito pelo esperado e acurácia	42
Tabela 4 - ResNet50: predito pelo esperado e acurácia	45
Tabela 5 - ResNet101: predito pelo esperado e acurácia	48
Tabela 6 - Acurácia das redes	50

## LISTA DE SIGLAS

- CNN Convolutional Neural Network
- GPU Graphics Processing Unit
- HSV Hue, Saturation and Value
- ILSVRC Imagenet Large Scale Visual Recognition Challenge
- LAFAM Laboratório de Análise Farmacêutica e Ambiental da Universidade Federal de Goiás
- MAE Mean absolute error
- RGB Red, Green and Blue
- RMSE Root-mean-square error

<b>SUMÁRIO</b>
----------------

1 INTRODUÇÃO	14
2 FUNDAMENTAÇÃO TEÓRICA	16
2.1 REDES NEURAIS CONVOLUCIONAIS	16
2.1.1 Camada de Convolução	17
2.1.2 Camada de Agrupamento	19
2.1.3 Camada Totalmente Conectada	19
2.1.4 Camada de Saída	20
2.1.5 AlexNet	20
2.1.6 GoogleNet	21
2.1.7 ResNet	24
2.2 HSV	26
3 MATERIAIS E MÉTODOS	27
3.1 ESTUDO DE CASO	27
3.2 MATERIAIS	27
3.3 AMOSTRAGEM	28
3.4 TRATAMENTO DAS IMAGENS	29
3.5 AUMENTO DE DADOS	30
3.6 TREINAMENTO	32
3.7 MÉTRICA DE AVALIAÇÃO DE MODELOS	32
4 RESULTADOS	34
4.1 ALEXNET	34
4.2 GOOGLENET	37
4.3 RESNET18	41
4.4 RESNET50	44
4.5 RESNET101	47
4.6 COMPARATIVO ENTRE AS REDES	50
5 CONCLUSÃO	51

## 1 INTRODUÇÃO

O controle de qualidade em forma de sistema administrativo originou-se no século XX nos Estados Unidos para agregar atributos de qualidade às empresas e aumentar a expectativa da permanência delas no mercado. O aumento de concorrência e a exigência dos consumidores forçaram as empresas a adotar medidas para desenvolverem melhor os seus produtos e serviços (SELEME et al., 2012). Diversos países adotam medidas regulamentares para garantir a qualidade de serviços e produtos, no Brasil, existem órgãos reguladores como a Agência Nacional de Vigilância Sanitária (ANVISA) que são responsáveis pela fiscalização e controle sanitário da comercialização desses serviços e produtos (VALE, 2018). O controle de qualidade em produtos de consumo, como alimentos e fármacos, é de suma importância para evitar riscos à saúde dos consumidores, bem como os prejuízos de produção (ROCHA et al., 2014; PAULA et al., 2017).

De maneira geral, a metodologia clássica utilizada no controle de qualidade para alimentos e fármacos envolve processos químicos que dependem de laboratório. Processos como os cromatográficos podem ser onerosos, destrutivos, consumir insumos e consequentemente podem demandar uma certa quantidade de tempo. Além disso, podem impossibilitar um processo de verificação em larga escala (CHEN et al., 2013).

Uma alternativa aos métodos clássicos é a utilização de modelos computacionais empregando imagens digitais. Existem trabalhos que propõem o uso de imagens digitais para a determinação da concentração de substâncias, como foi proposto por Sorouraddin et al. (2015) gerando determinação de quatro tipos de corantes em produtos comerciais via análise da coloração das imagens; por Shishkin et al. (2003) quantificando substâncias absorvidas por espumas de poliuretano usando processamento digital de imagens; também por Lopez-Molinero et al. (2010) identificando a quantidade de titânio em plástico com redes neurais artificiais. Outro exemplo é a utilização de análise estatística em imagens digitais para determinar o teor de proteína em amostras de arroz (SUN et al., 2008).

Com o avanço do estado da arte, técnicas tradicionais de visão computacional perderam espaço, principalmente em problemas complexos, onde modelos de redes neurais convolucionais ou *convolution neural networks* (CNNs) comumente levam

vantagem (O'MAHONY et al., 2019). As CNNs pertencem a um tipo de modelo de aprendizado de máquina profundo (multicamadas), que tem como principal característica a implementação de camadas convolucionais. Até o momento deste trabalho, o uso de CNNs para determinação da concentração de substâncias foi empregado em trabalhos como: na predição de propriedades do solo usando a combinação de espectros de imagens no visível, infravermelho próximo e infravermelho médio (NG et al., 2019); foi proposto também por Ni, Wang e Tao (2019) para quantização de nitrogênio em imagens no infravermelho próximo de folhas de mudas de pinheiro Masson.

O açafrão-da-terra (*Curcuma longa*) é uma planta herbácea proveniente da Ásia, o pó feito da sua raiz é um condimento consumido mundialmente. Existem diversos estudos sobre benefícios do pó de açafrão à saúde, como: atividades antiinflamatórias, antioxidantes, auxílio no combate de doenças como câncer, artrite, Alzheimer, doenças cardiovasculares e gastrointestinais (SINGLETARY, 2010). Com a ascensão do consumo desses tipos de produtos provenientes de meios industriais, eles estão sujeitos às adulterações, como a diluição de matéria prima mais barata e de menor qualidade para dar volume. Procedimentos como esse podem resultar em um produto de baixa qualidade, com menor ação benéfica à saúde ou até ser nocivo ao consumidor (BANSAL et al., 2015). No caso da raiz do açafrão-da-terra, o amido faz parte de sua composição natural, porém, é comum a adição de amido no pó de açafrão (PARVATHY et al., 2015).

Neste contexto, este trabalho propõe o estudo exploratório do uso de CNNs para determinação da concentração de amido em imagens digitais de amostras de açafrão-da-terra. Como estudo de caso foram utilizadas amostras de açafrão-da-terra de diferentes marcas, em diferentes níveis de adição de amido para treinar e testar algumas das arquiteturas de CNNs existentes: AlexNet, GoogleNet e ResNet.

## 2 FUNDAMENTAÇÃO TEÓRICA

#### 2.1 Redes Neurais Convolucionais

Rede neural convolucional é um modelo de *deep learning* que pertence à classe de redes neurais do tipo *feedfoward* (redes neurais onde os dados percorrem da camada de entrada até a camada de saída e sem ciclos). Um dos primeiros a demonstrar esse conceito foi o pesquisador Yann LeCun, em 1998, ao criar a rede LeNet para classificar dígitos escritos à mão (LECUN, et al., 1998). A rede, cuja arquitetura está representada na Figura 1, tem como estrutura: camadas de convoluções, subamostragem, camada totalmente conectada e etc. As CNNs se destacaram em problemas de classificação de imagens, com um aumento considerável de assertividade em relação aos outros métodos de classificação.

De acordo com LeCun, et al. (1998), é possível construir modelos de reconhecimentos de padrões melhores desenvolvendo mais o aprendizado automático e deixando de lado o desenvolvimento de heurísticas projetadas manualmente. Desde então, técnicas tradicionais de visão computacional perderam espaço, principalmente em problemas complexos, onde as *CNNs* comumente levam vantagem (O'MAHONY et al., 2019).





Para uma rede neural ser convolucional, é necessário que ela contenha convolução em ao menos uma camada. Além da camada de convolução, esse tipo de

Fonte: Adaptado de LeCun, et al., 1998

rede também pode conter outros tipos de camadas, dentre elas: camada de agrupamento, totalmente conectada e saída.

#### 2.1.1 Camada de Convolução

A convolução é uma operação linear entre duas funções que gera uma terceira função (WEISSTEIN, [S.D.]). Considerando que as imagens digitais são matrizes de números (números esses que representam *pixels*), a convolução é o processo de aplicar filtro (em inglês: *kernel*) à uma imagem, que tem como resultado uma imagem filtrada. O filtro é aplicado "deslizando" a imagem de acordo com o passo (em inglês: *stride*), pelas colunas e depois pelas linhas, como no caso da Figura 2 onde o *stride* é 1. Cada passo tem como resultado o somatório da multiplicação entre cada elemento do *kernel* pelo *pixel* correspondente da imagem, considerando a região onde o passo em si está sendo executado. O resultante dos passos tem o tamanho reduzido, conforme a equação abaixo, onde z é a imagem convolucionada decorrente da imagem f em função do *kernel* w (GONZALEZ; WOODS, 2009).

$$z(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) \cdot f(x-s,y-t)$$
(1)

# Figura 2 - Filtro 3x3 (em rosa) aplicado à uma imagem 5x5 (em roxo) resultando em uma imagem 3x3 (em azul)



Fonte: Adaptado de ScienceDirect. Disponível em: https://www.sciencedirect.com/topics/computer-science/convolutional-filter

As camadas de convolução contém *kernels* de tamanhos definidos previamente, que por sua vez, têm os valores atualizados conforme o treinamento da rede. É a partir desse processo que as CNNs conseguem extrair características (ou em inglês: *features*), como mostra a Figura 3, onde a imagem de entrada passa por camadas de convolução resultando em mapas de ativação (PONTI; COSTA, 2017). Por causa da redução do tamanho dos mapas de ativação ao longo das convoluções, é comum utilizar o preenchimento (em inglês: *padding*), que é uma borda adicionada à imagem antes de ser processada pela convolução. O valor e tamanho são definidos de acordo com a aplicação, a Figura 4 mostra um *padding* de tamanho 1 com valor 0 sendo aplicados à imagem, que ao aplicar a convolução, o tamanho do mapa de ativação fica idêntico ao da imagem (DATAHACKER, 2018).



Figura 3 - Ilustração de duas camadas convolucionais

Fonte: PONTI, et al., 2017





Fonte: DataHacker. Disponível em: http://datahacker.rs/what-is-padding-cnn/

#### 2.1.2 Camada de Agrupamento

Camadas de agrupamento (em inglês: *pooling*), comumente são usadas após as camadas de convolução, elas têm o funcionamento parecido ao de convolução em relação à operação deslizante utilizando uma janela e *stride*. Sua ação é de redução de dimensionalidade, reduzindo o tamanho do mapa de características gerado pelas camadas de convolução (ou a camada anterior), diminuindo assim, a quantidade de informações a serem processadas na camada seguinte. Uma técnica bastante utilizada é o agrupamento máximo (em inglês: *max pooling*), onde reduz-se os dados selecionando o maior valor de uma determinada região dada pelo filtro (Figura 5). Outras operações também podem ser feitas, como a média dos valores da janela (em inglês: *average pooling*) (LACERDA, 2019).

Figura 5 - Representação da camada de max-pooling com stride 2



Fonte: Adaptado de Programmer Sought. Disponível em: https://www.programmersought.com/article/62833523524/

## 2.1.3 Camada Totalmente Conectada

As camadas totalmente conectadas (em inglês: fully connected) antecedem a camada de saída, o nome se dá pois todos os neurônios de saída da camada anterior estão conectados em cada um dos neurônios de entrada dessa camada (Figura 6). Ela calcula as probabilidades (em inglês: scores) da entrada estar relacionada às classes. A saída dessa camada é de tamanho igual à quantidade de classes, para gerado cada classe é um valor entre 0 1 para representar е а probabilidade (LACERDA, 2019).

Figura 6 - Representação da fully connected



Fonte: Adaptado de Numahub. Disponível em: https://numahub.com/articles/why-do-you-need-fullyconnected-layer

## 2.1.4 Camada de Saída

Para problemas de classificação, a última camada de uma CNN é composta por n neurônios, sendo n o número de classes, portanto, a quantidade de possibilidades de respostas é discreta. Cada neurônio dessa camada se conecta com todos os neurônios da camada anterior. Após o processamento das camadas anteriores, os mapas de ativação se transformam em um vetor de características, vetor esse que é processado, resultando na predição da rede (O'SHEA; NASH, 2015).

Como as CNNs são modelos de aprendizado supervisionado, essa camada também é responsável pela função de cálculo da perda, que durante o treinamento, atualiza os pesos e *bias* se a predição for incorreta. O algoritmo usado para correção de erro é o *back-propagation*.

Existem problemas que não são resolvidos por modelos de classificação, pois a sua resposta deve ser contínua, onde o resultado só pode ser calculado através de métodos de regressão. Em todas as arquiteturas usadas neste trabalho, a última camada das redes foi substituída por uma camada de regressão.

## 2.1.5 AlexNet

A rede foi proposta em 2012 pelos pesquisadores Alex Krizhevsky, Ilya Sutskever e Geoffrey E. Hinton (KRIZHEVSKY, et al., 2012), obtendo visibilidade por

conseguir bons resultados na edição de 2012 do *Imagenet Large Scale Visual Recognition Challenge* (ILSVRC) (OLGA; JIA et al., 2015).

A Arquitetura da AlexNet é composta por 8 camadas assim como demonstra a Figura 7, foi dívida em duas partes para treinamento paralelo entre duas *GPUs* (KRIZHEVSKY, et al., 2012). A primeira camada é uma convolução de entrada de imagens tamanho 224x224x3, com 96 filtros de tamanho 11x11x3 com *stride* 4 e saída de tamanho 55x55x96. A segunda camada é um *max-pooling* seguido de uma convolução, com entrada de tamanho 55x55x96, o *max-pooling* é de tamanho 27x27x96, com 256 filtros de tamanho 5x5x48 e saída de 27x27x256. As camadas 3, 4 e 5 seguem parecidas com a camada 2. A sexta camada é composta por uma *fully connected* com tamanho de entrada 13x13x128 (transformada em vetor de 1x13x13x128) e multiplicado por uma matriz de tamanho 13x13x128x2048 com a saída de tamanho 1x2048. As camadas 7 e 8 são similares à sexta camada. A saída da rede é dividida em 1000 classes (ANAND, [S.D]).





Fonte: Adaptado de KRIZHEVSKY, et al., 2012

## 2.1.6 GoogleNet

A GoogleNet foi proposta pela empresa Google em 2014, com foco em eficiência computacional. Mesmo sendo mais profunda que a AlexNet, com um total de 22 camadas, ela é 12 vezes menos custosa computacionalmente falando (SZEGEDY et al., 2014), a rede ganhou a edição de 2014 da ILSVRC (OLGA; JIA et al., 2015). Uma das melhorias dessa rede em comparação à AlexNet foi a

implementação do módulo *Inception*. Esse módulo é composto por uma paralelização de filtros de diferentes tamanhos, como 1x1, 3x3 e 5x5 (Figura 8). O princípio da utilização de filtros de tamanhos diferentes parte do pressuposto que pode existir características fortemente correlacionadas em diferentes tamanhos e posições espaciais nas imagens, portanto, utilizar essa técnica extrai melhor as *features* nessas situações (SHAIKH, F, 2018).



Figura 8 - Módulo Inception

Fonte: Adaptado de SZEGEDY et al., 2014

A aplicação de convoluções de tamanho 1x1 servem para a redução da dimensionalidade dos dados reduzindo assim o custo computacional. Como mostra a Figura 9, onde um filtro 1x1x192 aplicado a um mapa de ativação de tamanho 64x64x192, resulta em um mapa de ativação de tamanho 64x64x1. Após o processamento de todos os filtros do módulo, os mapas de ativação são concatenados, passando então, o resultado para a entrada da próxima camada (SAKTHI, 2018). Como exemplo disso, a Figura 10 mostra o módulo *inception* de forma tridimensional, onde os resultados de todas as aplicações de filtros são concatenados no final.

Figura 9 - Exemplo de convolução 1x1: filtro 1x1x192 aplicado a um mapa de ativação 64x64x192 resultando em outro de tamanho 64x64x1 (com profundidade reduzida)









Fonte: Adaptado de MX's Blog. Disponível em: https://x-wei.github.io/dlMOOC\_L3.html

A rede também utiliza o *dropout*, essa é uma técnica introduzida por Hinton, et al. (2012) com o intuito de minimizar o *overfitting* da rede descartando temporariamente e aleatoriamente uma quantidade de neurônios. Esse método permite que a rede aumente a acurácia e a capacidade de generalização. Por exemplo: supondo que uma rede classifique carros e que alguns dos neurônios sejam ativados por *features* relacionadas às rodas, ao descartá-los, a rede aprende com essa situação e deve ser capaz de classificar imagens de carros sem as rodas.

#### 2.1.7 ResNet

Proposta pelos pesquisadores Kaiming He, Xiangyu Zhang, Shaoqing Ren e Jian Sun, a Residual Net (ou ResNet) (HE, et al., 2015), ganhou a edição de 2015 do ILSVRC (OLGA; JIA et al., 2015). Os criadores dessa rede assumem a importância do aumento das camadas das CNNs, porém, mostram que ao aumentar a quantidade de camadas de uma rede, a acurácia dela pode diminuir. Um exemplo disso é a Figura 11, onde para uma determinada rede escolhida pelos autores, o erro de treinamento e teste aumentam ao inserir mais camadas (de 20 para 56). Isso se deve ao problema de degradação de gradiente (ou em inglês: *vanishing gradient*). O gradiente descendente é um método de otimização comumente utilizado em CNNs que calcula derivadas parciais para cada parâmetro da rede, a fim de minimizar os erros (PONTI; COSTA, 2017).

Figura 11 - Aumento do erro de treinamento e teste de redes ao aumentar camadas de 20 para 56



Fonte: Adaptado de He, et al., 2015

A solução proposta pelos criadores da ResNet para o problema de degradação do gradiente foi a implementação de blocos residuais (Figura 12). A partir de uma entrada *x* os dados passam em camadas de convolução e relu, e no final do bloco, o resultado da entrada processada é somada com a entrada original por meio da "conexão de atalho de identidade". Após a utilização dos blocos residuais, a otimização foi mais eficaz e permitiu o empilhamento de mais camadas (HE, et al., 2015).





Fonte: Adaptado de He, et al., 2015

A rede é composta por diversos blocos residuais empilhados de acordo com a quantidade de camadas (Figura 13). Existem diferentes implementações da ResNet, como a ResNet18, a ResNet50, ResNet101 e ResNet152, com 18, 50, 101 e 152 camadas, respectivamente. Além dos blocos residuais, as principais camadas que compõem este tipo de arquitetura são: entrada que recebe imagens de tamanho 224x224x3, *average pooling e fully connected*.





Fonte: Adaptado de Missinglink.ai. Disponível em: https://missinglink.ai/guides/keras/kerasresnet-building-training-scaling-residual-nets-keras/

#### 2.2 HSV

O esquema de cores HSV, também conhecido como HSB (*hue*, *saturation and brightness*, em português: matriz, saturação e brilho) é uma alternativa ao RGB (*red*, *green and blue*). O RGB representa as cores compondo entre vermelho, verde e azul, já o HSV utiliza: matriz para o tipo de cor, dentre todas as cores do espectro; a saturação para o quão menos acinzentada é a cor; e por fim, utiliza o brilho para definir o brilho da cor, entre totalmente preto e brilhante (ALFAMIDIA, 2019). Para problemas relacionados à cor em imagens digitais, o HSV proporcionou melhores resultados em trabalhos como o proposto por Inoue et al. (2019) e também por Du et al. (2005), onde o HSV foi utilizado em conjunto com redes neurais artificiais e outros métodos de aprendizado de máquina.

## **3 MATERIAIS E MÉTODOS**

#### 3.1 Estudo de caso

O açafrão-da-terra (nome binomial: *Curcuma longa*) é uma planta herbácea com raiz tuberosa originária da Ásia, sua raiz seca e moída (em forma de pó) é utilizada na culinária como tempero/corante. O pó de açafrão contém uma concentração natural de amido, porém, o uso de amido para adulteração do pó de açafrão é comum (LEONEL et al., 2003). Como estudo de caso foram utilizadas imagens digitais de amostras de pó de açafrão-da-terra disponibilizadas pelo Laboratório de Análise Farmacêutica e Ambiental da Universidade Federal de Goiás (LAFAM) para explorar o uso de CNNs na determinação da concentração de amido nessas amostras.

#### 3.2 Materiais

As amostras são compostas por 10 marcas diferentes de açafrão (nomeados como A, B, C, D, E, F, G, H, I e J) (Figura 14) e amido puro. Para cada marca de açafrão há a amostra pura, ou seja, sem adição de amido e 4 níveis de mistura com amido com as seguintes proporções: 20:80, 40:60, 60:40 e 80:20 (amido:pó de açafrão), ou seja, a marca A 20:80 é uma mistura de 20% de amido com 80% de açafrão. A Figura 15 mostra a marca A com suas composições e o amido puro.

Foram utilizados dois computadores para o treinamento das redes, sendo o primeiro com um processador Intel Core I9 9900KF, 16 GB de memória RAM DDR4 e uma placa de vídeo GTX 650 1 GB; o segundo com um processador Intel Core I5 3330, 8 GB de memória RAM DDR3 e uma placa de vídeo GTX 760 2GB. Foi utilizado o *software* Matlab na versão 2020a com o pacote *Deep Learning Toolbox*. Para manuseio e realização da aquisição das amostras, foram utilizados os seguintes itens: folhas de papel A4, batoques tamanho M, placas de Petri, lâminas de vidro, pinças, suporte para celular, luminária, borrifador com álcool 70%, borrifador com água, espátulas e o celular Samsung Galaxy S9 Plus para preparação das amostras e captura das imagens.

## Figura 14 - Amostras de todos as marcas de açafrão (de A a J) na sua composição pura



Fonte: Autoria Própria (2020)

Figura 15 - Amostras de açafrão da marca A em sua composição pura, 20:80, 40:60, 60:40, 80:20 e amido puro, respectivamente



Fonte: Autoria Própria (2020)

## 3.3 Amostragem

Para realização da captura das imagens, as amostras foram transferidas do recipiente original para um batoque tamanho M, preenchendo-o por completo e deixando a superfície retilínea. Após isso, foi realizada a captura de 3 fotos com o celular.

O processo para cada amostra foi feito em triplicata, sendo a amostra retornada para o recipiente original, misturada e executado o processo de aquisição novamente. No final, resultou-se em 9 imagens para cada amostra de açafrão e para a amostra de amido, totalizando em 459 imagens semelhantes à Figura 16.

Figura 16 - Foto da amostra I Puro



Fonte: Autoria Própria (2020)

## 3.4 Tratamento das Imagens

Para utilização das imagens, elas foram cortadas com um tamanho de 150x150 *pixels* na região central do batoque usando um *software* de edição de imagens, como demonstra a figura a seguir:

Figura 17 - Demonstração de recorte 150x150 pixels na amostra de açafrão I puro



Fonte: Autoria Própria (2020)

## 3.5 Aumento de Dados

Recomenda-se utilizar uma grande quantidade de dados para treinamento de redes neurais convolucionais, uma solução para trabalhar com dados limitados é a utilização de métodos para aumentar a quantidade desses dados chamada de "aumento de dados" (SHORTEN; KHOSHGOFTAAR, 2019). Como método de aumento de dados, todas as imagens obtidas com o processo de corte de 150x150 foram divididas em 9 imagens de tamanho 50x50 resultando em 27 imagens para cada amostra, como demonstra a figura a seguir:

Figura 18 - Demonstração de recorte 150x150 *pixels* para 9 imagens de 50x50 *pixels* na amostra de açafrão I puro



Fonte: Autoria Própria (2020)

Para os dados de treinamento (todas as amostras da marca A e o amido puro) foi utilizada a técnica de espelhamento horizontal, vertical e horizontal com vertical simultaneamente (Figura 19), resultando em 81 imagens para cada amostra.





Fonte: Autoria Própria (2020)

Ainda para os dados de treinamento, as imagens obtidas no processo de corte 150x150 foram utilizadas para outro processo de aumento de dados: rotação da imagem de 10 em 10 graus por 35 vezes. Para cada rotação, foi feito o corte da região central da imagem no tamanho de 50x50 (Figura 20). Somados com as 81 imagens obtidas no processo anterior, resultou em 639 imagens por amostra.

Figura 20 - Demonstração de rotação e corte da imagem 150x150 da amostra de açafrão A puro



Fonte: Autoria Própria (2020)

#### 3.6 Treinamento

Foram utilizadas as redes: AlexNet, GoogleNet, ResNet18, ResNet50 e ResNet101 pelo Matlab por meio de transferência de aprendizado, que é a técnica de começar o treinamento com os valores dos pesos das redes treinadas, ao invés de atribuir novos valores iniciais. Os valores reais referentes às concentrações de cada marca de açafrão foram obtidos no trabalho feito por Macêdo et al. (2020). Para todas as redes, a última camada foi substituída por uma de regressão. O treinamento foi realizado utilizando as imagens das amostras de açafrão da marca A e o amido puro, convertidas para o espectro de cores HSV e redimensionando-as de acordo com o tamanho da entrada de cada rede. O treinamento foi dividido em duas partes, sendo a primeira parte: treino e teste da marca A com 60% das imagens para treino, 20% para validação e 20% para teste; e a segunda: treino total da marca A com 80% para treino e 20% para validação, onde o teste foi realizado com todas as imagens das outras marcas, ou seja, de B até J. A divisão dos dados de treinamento, validação e teste foi feita de forma aleatória.

A parametrização utilizada para as redes foi: tamanho do *mini-batch* igual a 5, unidade de processamento gráfico ou *Graphics Processing Unit* (GPU) como ambiente de execução, 300 épocas, taxa de aprendizado inicial igual a 0.0001 e frequência de validação sendo o número inteiro mais próximo da divisão entre a quantidade de imagens de treinamento e o tamanho do *mini-batch*.

### 3.7 Métrica de Avaliação de Modelos

Foram utilizados três métricas para avaliação do desempenho dos modelos testados, que são: o erro de raiz quadrático médio (em inglês: root mean square error)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2},$$
(2)

erro médio absoluto (em inglês: mean absolute error)

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |\hat{y}_i - y_i|$$
(3)

e o coeficiente de determinação (ou R2)

$$R^{2} = 1 - \frac{\sum_{i=1}^{N} (y_{i} - \hat{y}_{i})^{2}}{\sum_{i=1}^{N} (y_{i} - \overline{y}_{i})^{2}}$$
(4)

onde *N* é a quantidade de dados,  $\hat{y}_i$  é o valor predito,  $y_i$  é o valor esperado e  $\bar{y}_i$  é a média dos valores esperados. Para o RMSE e MAE, quanto mais próximo de 0 melhor e para o R<sup>2</sup> quanto mais próximo de 1 melhor.

#### 4 RESULTADOS

Este capítulo apresenta os resultados de cada rede e, em seguida, a comparação entre os resultados delas. As tabelas para cada rede contêm os valores esperados das concentrações das diferentes marcas de açafrão foram feitas com base no trabalho de Macêdo et al. (2020), a predição para todas as concentrações de cada marca de açafrão com a média e desvios padrão e, por fim, o resultado das métricas de avaliação referente a cada marca. Ao final do capítulo, há uma tabela que compara o resultado de todas as redes.

## 4.1 AlexNet

A Tabela 1 apresenta o resultado obtido pela rede com as médias e desvios padrão para cada concentração de cada marca de açafrão. O resultado da predição para a marca A (que foi utilizado para treino) foi de 6,42 de RMSE, 5,95 de MAE e 0,94 de R<sup>2</sup>. Para as demais marcas, o melhor resultado foi da marca C, com 10.85 de RMSE, 10,02 de MAE e 0,77 de R<sup>2</sup> e o pior foi da marca F com 28,24 de RMSE, 24,46 de MAE e -0,22 de R<sup>2</sup>. Dentre as redes testadas, foi uma das que apresentaram menores desvios padrão, com o valor máximo de 4,78.

			Esperado						Acurácia				
Marca	X-P ± DP <sup>a</sup>	$X-2 \pm DP^{b}$	$X-4 \pm DP^{b}$	X-6 $\pm$ DP <sup>b</sup>	$X-8 \pm DP^{b}$	X-P ± DP <sup>c</sup>	$X-2 \pm DP^{d}$	$X-4 \pm DP^{d}$	X-6 ± DP <sup>d</sup>	X-8 ± DP <sup>d</sup>	RMSE	MAE	R²
А	29,82 ± 0,56	43,86 ± 0,14	$66,32 \pm 0,12$	86,53 ± 0,08	97,31 ± 0,07	$26,94 \pm 0,74$	34,44 ± 2,71	58,33 ± 3,09	81,51 ± 2,68	92,88 ± 1,47	6,42	5,95	0,94
В	$66,05 \pm 0,36$	72,84 ± 0,12	83,7 ± 0,14	93,48 ± 0,09	98,7 ± 0,05	32,24 ± 1,05	34,81 ± 2,17	50,76 ± 2,68	$72,43 \pm 2,74$	93,05 ± 1,27	28,81	26,3	0,32
С	33,99 ± 2,41	$47,19 \pm 0,27$	68,31 ± 0,13	87,33 ± 0,07	97,47 ± 0,11	28,4 ± 1	42,85 ± 2,88	55,68 ± 3,42	73,45 ± 3,65	83,81 ± 3,13	10,85	10,02	0,77
D	$36,9 \pm 0,27$	49,52 ± 0,11	69,71 ± 0,1	87,88 ± 0,07	97,58 ± 0,02	31,7 ± 0,5	$35,89 \pm 2,3$	49,82 ± 2,39	57,63 ± 4,44	80 ± 3,68	19,14	17,31	0,39
Е	$22,14 \pm 0,44$	37,71 ± 0,14	$62,63 \pm 0,05$	85,05 ± 0,1	97,01 ± 0,03	29,87 ± 0,72	43,28 ± 3,13	51,51 ± 2,41	68,91 ± 2,69	83,83 ± 4,48	11,39	10,75	0,67
F	$27,56 \pm 0,08$	$42,05 \pm 0,1$	$65,23 \pm 0,07$	86,09 ± 0,02	97,22 ± 0,04	$33,03 \pm 0,58$	32,25 ± 1,41	31,96 ± 1,77	44,98 ± 3,42	64,59 ± 3,73	28,24	24,46	-0,22
G	$39,3 \pm 0,61$	51,44 ± 0,12	70,87 ± 0,1	88,35 ± 0,03	97,7 ± 0,05	31,05 ± 1,71	35,54 ± 2,23	48,7 ± 3,26	64,57 ± 5,26	90,28 ± 2,83	16,93	15,5	0,59
н	24,21 ± 1,25	39,37 ± 0,21	63,62 ± 0,11	85,45 ± 0,04	97,09 ± 0,03	$28,34 \pm 0,72$	35,6 ± 3,84	47,87 ± 4,18	65,92 ± 3,15	85,82 ± 4,78	12,55	10,89	0,7
I	53,98 ± 1,67	63,19 ± 0,23	77,91 ± 0,15	91,16 ± 0,07	98,23 ± 0,09	31,34 ± 1,26	33,36 ± 2,74	46,74 ± 3,26	$62,03 \pm 4,43$	88,51 ± 3,52	25,76	24,5	0,37
J	70,32 ± 1,11	76,26 ± 0,22	85,75 ± 0,18	94,3 ± 0,08	98,86 ± 0,08	38,47 ± 3,65	62,89 ± 4,28	80,02 ± 3,62	89,51 ± 2,64	94,99 ± 1,87	15,9	11,92	0,55

Tabela 1 - AlexNet: predito pelo esperado e acurácia

Fonte: Autoria Própria (2020)

<sup>a</sup> Amido por 100 gramas de pó de açafrão sem mistura ± desvio padrão.

<sup>b</sup> Mistura de X-2 até X-8 correspondente à 2:8 até 8:2 (amido:pó de açafrão), respectivamente.

<sup>c</sup> Média do resultado da predição da concentração de amido para amostra de açafrão sem mistura ± desvio padrão.

<sup>d</sup> Média do resultado da predição da concentração de amido para amostra de açafrão com mistura de X-2 até X-8 correspondente à 2:8 até 8:2 amido:pó de açafrão, respectivamente ± desvio padrão.

A Figura 21 mostra os resultados de forma gráfica, com os valores preditos pelos esperados para cada marca. Cada ponto azul representa uma imagem testada e a linha pontilhada vermelha representa o valor ideal, ou seja, quanto mais próximos os pontos estiverem da linha, melhor o resultado. O resultado das imagens testadas das marcas A e C se encontram mais próximos à linha do valor ideal. A marca F tem uma discrepância expressiva à linha ideal, o que justifica o R<sup>2</sup> negativo mostrado na Tabela 1. O resultado do restante das marcas apresenta uma diferença significativa do valor esperado, mas a distribuição dos dados segue um comportamento correlato à linha vermelha.



Figura 21 - AlexNet: gráficos de predito pelo esperado

Fonte: Autoria Própria (2020)

## 4.2 GoogleNet

Como mostra a Tabela 2, o resultado da predição para a marca A foi de 1,78 de RMSE, 1,35 de MAE e 0,99 de R<sup>2</sup>. Depois da marca A, o melhor resultado foi da marca C, com 4,6 de RMSE, 4,09 de MAE e 0,96 de R<sup>2</sup>, ao contrário disso, o pior resultado foi da marca F com 23,75 de RMSE, 20,82 de MAE e 0 de R<sup>2</sup>. Também

apresentou menores desvios padrão em relação às outras redes, com o valor máximo de 4,83.

			Esperado						Acurácia				
Marca	X-P ± DP <sup>a</sup>	$X-2 \pm DP^{b}$	$X-4 \pm DP^{b}$	X-6 $\pm$ DP <sup>b</sup>	$X-8 \pm DP^{b}$	X-P ± DP <sup>c</sup>	$X-2 \pm DP^{d}$	$X-4 \pm DP^{d}$	X-6 $\pm$ DP <sup>d</sup>	$X-8 \pm DP^{d}$	RMSE	MAE	R²
А	29,82 ± 0,56	43,86 ± 0,14	66,32 ± 0,12	86,53 ± 0,08	97,31 ± 0,07	32,85 ± 1,02	41,46 ± 4,72	66,34 ± 2,66	86,01 ± 1,91	96,51 ± 1,3	1,78	1,35	0,99
В	$66,05 \pm 0,36$	72,84 ± 0,12	83,7 ± 0,14	93,48 ± 0,09	98,7 ± 0,05	34,02 ± 1,43	39,36 ± 1,98	53,85 ± 2,64	73,78 ± 2,34	96,3 ± 1,36	26,2	23,49	0,37
С	33,99 ± 2,41	47,19 ± 0,27	68,31 ± 0,13	87,33 ± 0,07	97,47 ± 0,11	31,74 ± 0,98	48,21 ± 3,21	61,76 ± 2,76	81,85 ± 2,61	92,32 ± 2,8	4,6	4,09	0,96
D	$36,9 \pm 0,27$	49,52 ± 0,11	69,71 ± 0,1	87,88 ± 0,07	97,58 ± 0,02	32,24 ± 0,26	$38,9 \pm 2,48$	56,2 ± 1,95	66,36 ± 3,92	86,34 ± 3,36	13,46	12,31	0,66
Е	$22,14 \pm 0,44$	37,71 ± 0,14	$62,63 \pm 0,05$	85,05 ± 0,1	97,01 ± 0,03	31,85 ± 0,28	47,71 ± 2,83	57,26 ± 2,4	74,84 ± 2,48	89,32 ± 3,37	8,79	8,59	0,81
F	$27,56 \pm 0,08$	42,05 ± 0,1	$65,23 \pm 0,07$	86,09 ± 0,02	97,22 ± 0,04	32,72 ± 0,41	33,08 ± 1,36	35,71 ± 1,92	53,35 ± 3,12	69,48 ± 2,77	23,75	20,82	0
G	$39,3 \pm 0,61$	51,44 ± 0,12	70,87 ± 0,1	88,35 ± 0,03	97,7 ± 0,05	32,11 ± 0,92	40,71 ± 2,4	55,18 ± 2,72	67,98 ± 3,49	93,76 ± 2,1	12,99	11,58	0,72
Н	24,21 ± 1,25	39,37 ± 0,21	63,62 ± 0,11	85,45 ± 0,04	97,09 ± 0,03	31,45 ± 0,45	40,38 ± 3,77	56,44 ± 3,47	74,1 ± 2,24	91,27 ± 3,25	7,32	6,52	0,89
I	53,98 ± 1,67	63,19 ± 0,23	77,91 ± 0,15	91,16 ± 0,07	98,23 ± 0,09	$32,23 \pm 0,74$	37,37 ± 2,68	54,03 ± 2,8	67,57 ± 3,05	91,14 ± 2,64	21,53	20,43	0,47
J	70,32 ± 1,11	76,26 ± 0,22	85,75 ± 0,18	94,3 ± 0,08	98,86 ± 0,08	46,6 ± 4,83	68,72 ± 3,43	83,59 ± 2,86	92,51 ± 2,51	97,91 ± 1,44	11,21	7,23	0,68

Tabela 2 - GoogleNet: predito pelo esperado e acurácia

Fonte: Autoria Própria (2020)

<sup>a</sup> Amido por 100 gramas de pó de açafrão sem mistura ± desvio padrão.

<sup>b</sup> Mistura de X-2 até X-8 correspondente à 2:8 até 8:2 (amido:pó de açafrão), respectivamente.

<sup>c</sup> Média do resultado da predição da concentração de amido para amostra de açafrão sem mistura ± desvio padrão.

<sup>d</sup> Média do resultado da predição da concentração de amido para amostra de açafrão com mistura de X-2 até X-8 correspondente à 2:8 até 8:2 amido:pó de açafrão, respectivamente ± desvio padrão.

Para a GoogleNet, ao observar as representações gráficas mostradas pela Figura 22, o resultado das imagens testadas das marcas A e C também se encontram mais próximos à linha do valor ideal. A marca F, por sua vez, tem o pior resultado. A distribuição dos dados das outras marcas segue um comportamento correlato à linha vermelha, mas ainda com uma diferença considerável do valor ideal.



Figura 22 - GoogleNet: gráficos de predito pelo esperado

Fonte: Autoria Própria (2020)

#### 4.3 ResNet18

A Tabela 3 mostra que o resultado da predição para a marca A foi de 3,1 de RMSE, 2,84 de MAE e 0,98 de R<sup>2</sup>. Para os demais dados de teste, o melhor resultado foi da marca E, com 7,34 de RMSE, 6,13 de MAE e 0,91 de R<sup>2</sup>, além disso, o pior foi da marca F com 29,98 de RMSE, 26,26 de MAE e 0,12 de R<sup>2</sup>. A rede, comparada às anteriores, apresentou maiores valores de desvio padrão, com um mínimo de 7,56 e máximo de 14,99.

			Esperado				Predito						
Marca	X-P ± DP <sup>a</sup>	$X-2 \pm DP^{b}$	$X-4 \pm DP^{b}$	X-6 $\pm$ DP <sup>b</sup>	$X-8 \pm DP^{b}$	X-P ± DP <sup>c</sup>	$X-2 \pm DP^{d}$	$X-4 \pm DP^{d}$	X-6 $\pm$ DP <sup>d</sup>	$X-8 \pm DP^{d}$	RMSE	MAE	R²
А	29,82 ± 0,56	43,86 ± 0,14	66,32 ± 0,12	86,53 ± 0,08	97,31 ± 0,07	29,03 ± 11,68	39,81 ± 12,14	68,78 ± 12,51	83,81 ± 14,26	93,14 ± 12,11	3,1	2,84	0,98
В	66,05 ± 0,36	72,84 ± 0,12	83,7 ± 0,14	93,48 ± 0,09	98,7 ± 0,05	15,26 ± 11,72	32,52 ± 11,45	48,15 ± 11,44	57,83 ± 13,51	89,18 ± 12,87	36,96	34,37	0,24
С	33,99 ± 2,41	47,19 ± 0,27	68,31 ± 0,13	87,33 ± 0,07	97,47 ± 0,11	23,39 ± 11,13	38,31 ± 9,71	56,12 ± 13,24	83,51 ± 14,26	89,31 ± 12,2	9,18	8,73	0,88
D	36,9 ± 0,27	49,52 ± 0,11	69,71 ± 0,1	87,88 ± 0,07	97,58 ± 0,02	17,27 ± 8,56	24,33 ± 8	47,38 ± 10,88	66,77 ± 14,51	83,92 ± 12,9	20,74	20,38	0,59
Е	$22,14 \pm 0,44$	37,71 ± 0,14	$62,63 \pm 0,05$	85,05 ± 0,1	97,01 ± 0,03	23,01 ± 10,78	40,47 ± 8,18	50,55 ± 9,62	78,95 ± 12,79	88,18 ± 14,15	7,34	6,13	0,91
F	27,56 ± 0,08	42,05 ± 0,1	65,23 ± 0,07	86,09 ± 0,02	97,22 ± 0,04	27,36 ± 10,44	18,51 ± 8,34	25,19 ± 7,56	46,92 ± 10,18	68,9 ± 12,03	29,98	26,26	0,12
G	39,3 ± 0,61	51,44 ± 0,12	70,87 ± 0,1	88,35 ± 0,03	97,7 ± 0,05	32,86 ± 12,29	34,43 ± 11,84	55,21 ± 13,55	59,56 ± 10,81	92,37 ± 12,8	16,93	14,65	0,58
н	24,21 ± 1,25	39,37 ± 0,21	63,62 ± 0,11	85,45 ± 0,04	97,09 ± 0,03	25,2 ± 11,3	33,64 ± 7,9	50,12 ± 9,39	78,51 ± 12,91	89,71 ± 11,06	7,98	6,91	0,9
I	53,98 ± 1,67	63,19 ± 0,23	77,91 ± 0,15	91,16 ± 0,07	98,23 ± 0,09	26,01 ± 14,4	26,27 ± 9,71	51,16 ± 11,78	73,58 ± 13,11	84,37 ± 12,52	25,93	24,62	0,43
J	70,32 ± 1,11	76,26 ± 0,22	85,75 ± 0,18	94,3 ± 0,08	98,86 ± 0,08	42,86 ± 9,92	63,11 ± 14,99	80,23 ± 14,8	86,21 ± 12,19	95,88 ± 12,16	14,37	11,44	0,57

Tabela 3 - ResNet18: predito pelo esperado e acurácia

Fonte: Autoria Própria (2020)

<sup>a</sup> Amido por 100 gramas de pó de açafrão sem mistura ± desvio padrão.

<sup>b</sup> Mistura de X-2 até X-8 correspondente à 2:8 até 8:2 (amido:pó de açafrão), respectivamente.

<sup>c</sup> Média do resultado da predição da concentração de amido para amostra de açafrão sem mistura ± desvio padrão.

<sup>d</sup> Média do resultado da predição da concentração de amido para amostra de açafrão com mistura de X-2 até X-8 correspondente à 2:8 até 8:2 amido:pó de açafrão, respectivamente ± desvio padrão.

É possível observar que nos gráficos mostrados pela Figura 23, a ResNet18 apresentou uma maior dispersão entre as predições para as mesmas concentrações, ou seja, um maior desvio padrão. Para as marcas A e E, é possível observar uma maior concentração das predições próximo à linha do ideal. A marca F continua seguindo um comportamento fora do esperado em relação à linha vermelha. As demais marcas seguem o comportamento correlato à linha que representa o valor ideal.



Figura 23 - ResNet18: gráficos de predito pelo esperado

Fonte: Autoria Própria (2020)

43

#### 4.4 ResNet50

O resultado da predição mostrado na Tabela 4 para a marca A foi de 2,32 de RMSE, 1,96 de MAE e 0,99 de R<sup>2</sup>. Subsequente à marca A, o melhor resultado foi da marca H, com 5,2 de RMSE, 4,56 de MAE e 0,96 de R<sup>2</sup>, já o pior resultado foi da marca F com 19.23 de RMSE, 19.23 de MAE e 0.13 de R<sup>2</sup>. Assim como a ResNet18, a rede apresentou desvios padrão maiores, com um valor mínimo de 5,06 e um máximo de 11,57.

			Esperado					Predito				Acurácia	
Marca	X-P ± DP <sup>a</sup>	$X-2 \pm DP^{b}$	$X-4 \pm DP^{b}$	X-6 $\pm$ DP <sup>b</sup>	X-8 $\pm$ DP <sup>b</sup>	X-P ± DP <sup>c</sup>	$X-2 \pm DP^{d}$	$X-4 \pm DP^{d}$	X-6 $\pm$ DP <sup>d</sup>	X-8 $\pm$ DP <sup>d</sup>	RMSE	MAE	R²
А	29,82 ± 0,56	43,86 ± 0,14	66,32 ± 0,12	86,53 ± 0,08	97,31 ± 0,07	29,11 ± 9,1	44,07 ± 9,08	69,58 ± 7,85	83,83 ± 10,58	100,22 ± 10,13	2,32	1,96	0,99
В	66,05 ± 0,36	72,84 ± 0,12	83,7 ± 0,14	93,48 ± 0,09	98,7 ± 0,05	33,64 ± 8,42	43,76 ± 6,35	57,19 ± 6,62	73,35 ± 7,95	81,51 ± 10,57	25,69	25,06	0,3
С	33,99 ± 2,41	47,19 ± 0,27	68,31 ± 0,13	87,33 ± 0,07	97,47 ± 0,11	27,86 ± 6,13	46,31 ± 6	55,09 ± 7,82	83,17 ± 7,98	101,18 ± 8,54	6,99	5,62	0,93
D	36,9 ± 0,27	49,52 ± 0,11	69,71 ± 0,1	87,88 ± 0,07	97,58 ± 0,02	30,06 ± 5,61	44,64 ± 5,83	50,88 ± 5,06	73,01 ± 8,54	92,83 ± 8	11,57	10,03	0,77
Е	$22,14 \pm 0,44$	37,71 ± 0,14	62,63 ± 0,05	85,05 ± 0,1	97,01 ± 0,03	26,24 ± 6,14	41,68 ± 5,31	47,33 ± 5,22	$74,79 \pm 6,4$	94,95 ± 9,88	8,67	7,14	0,88
F	27,56 ± 0,08	42,05 ± 0,1	65,23 ± 0,07	86,09 ± 0,02	97,22 ± 0,04	36,83 ± 7,1	47,86 ± 8,43	36,49 ± 6,29	59,37 ± 6,97	83,46 ± 9,37	19,23	16,86	0,13
G	39,3 ± 0,61	51,44 ± 0,12	70,87 ± 0,1	88,35 ± 0,03	97,7 ± 0,05	33,32 ± 7,39	43,23 ± 7,59	54,91 ± 7,18	71,97 ± 8,34	96,29 ± 9,33	11,21	9,59	0,79
н	24,21 ± 1,25	39,37 ± 0,21	63,62 ± 0,11	85,45 ± 0,04	97,09 ± 0,03	25,05 ± 5,28	45,22 ± 6,13	58,58 ± 7,58	77,3 ± 8,07	100,03 ± 9,57	5,2	4,56	0,96
I	53,98 ± 1,67	63,19 ± 0,23	77,91 ± 0,15	91,16 ± 0,07	98,23 ± 0,09	33,83 ± 6,74	38,24 ± 7,57	56,23 ± 6,61	69,35 ± 8,7	97,64 ± 10,31	19,87	17,84	0,54
J	70,32 ± 1,11	76,26 ± 0,22	85,75 ± 0,18	94,3 ± 0,08	98,86 ± 0,08	51,7 ± 7,93	75,93 ± 7,5	86,7 ± 8,5	97,03 ± 11,25	105,14 ± 11,57	8,88	5,78	0,77

Tabela 4 - ResNet50: predito pelo esperado e acurácia

Fonte: Autoria Própria (2020)

<sup>a</sup> Amido por 100 gramas de pó de açafrão sem mistura ± desvio padrão.

<sup>b</sup> Mistura de X-2 até X-8 correspondente à 2:8 até 8:2 (amido:pó de açafrão), respectivamente.

<sup>c</sup> Média do resultado da predição da concentração de amido para amostra de açafrão sem mistura ± desvio padrão.

<sup>d</sup> Média do resultado da predição da concentração de amido para amostra de açafrão com mistura de X-2 até X-8 correspondente à 2:8 até 8:2 amido:pó de açafrão, respectivamente ± desvio padrão.

Referente à ResNet50, os gráficos mostrados na Figura 24, assim como para a ResNet18, apresentaram uma maior dispersão entre os dados de mesma concentração. As predições para as marcas A e H se mostraram mais próximas ao ideal. A marca F, assim como nas anteriores, mantém um resultado menos adequado. As demais marcas seguem o comportamento correlato à linha que representa o valor ideal, mas ainda sem uma aproximação significativa da linha que representa os valores ideais.



Figura 24 - ResNet50: gráficos de predito pelo esperado

Fonte: Autoria Própria (2020)

#### 4.5 ResNet101

O resultado da predição mostrado na Tabela 5 para a marca A foi de 3,53 de RMSE, 3,05 de MAE e 0,98 de R<sup>2</sup>. Após a marca A, o melhor resultado foi da marca C, com 10.27 de RMSE, 8.26 de MAE e 0.84 de R<sup>2</sup>, já o pior resultado foi da marca F com 23.96 de RMSE, 23.42 de MAE e -4.6 de R<sup>2</sup>. Assim como as demais arquiteturas da ResNet utilizadas nesse trabalho, a rede apresentou desvios padrão maiores, com um valor mínimo de 7,07 e um máximo de 31,75.

			Esperado				Acurácia						
Marca	X-P ± DP <sup>a</sup>	$X-2 \pm DP^{b}$	$X-4 \pm DP^{b}$	X-6 $\pm$ DP <sup>b</sup>	$X-8 \pm DP^{b}$	X-P ± DP <sup>c</sup>	$X-2 \pm DP^{d}$	$X-4 \pm DP^{d}$	X-6 $\pm$ DP <sup>d</sup>	X-8 $\pm$ DP <sup>d</sup>	RMSE	MAE	R²
А	29,82 ± 0,56	43,86 ± 0,14	66,32 ± 0,12	86,53 ± 0,08	97,31 ± 0,07	30,16 ± 7,77	41,64 ± 7,02	69,24 ± 7,41	82,39 ± 8,81	102,92 ± 10,33	3,53	3,05	0,98
В	$66,05 \pm 0,36$	72,84 ± 0,12	83,7 ± 0,14	93,48 ± 0,09	98,7 ± 0,05	57,18 ± 22,8	50,13 ± 11,87	67,11 ± 12,41	75,79 ± 10,46	99,25 ± 16,13	15,38	13,28	0,49
С	33,99 ± 2,41	$47,19 \pm 0,27$	68,31 ± 0,13	87,33 ± 0,07	97,47 ± 0,11	31,43 ± 9,1	47,61 ± 8,04	52,23 ± 9,51	73,43 ± 11,85	105,8 ± 11,26	10,27	8,26	0,84
D	$36,9 \pm 0,27$	49,52 ± 0,11	69,71 ± 0,1	87,88 ± 0,07	97,58 ± 0,02	42,73 ± 12,79	52,58 ± 8,41	52,44 ± 9,18	59,09 ± 9,55	83,02 ± 12,5	16,63	13,9	0,05
Е	$22,14 \pm 0,44$	37,71 ± 0,14	62,63 ± 0,05	85,05 ± 0,1	97,01 ± 0,03	36,65 ± 9,18	43,85 ± 7,73	48,27 ± 8,18	63,17 ± 9,08	93,31 ± 11,26	13,76	12,12	0,55
F	$27,56 \pm 0,08$	$42,05 \pm 0,1$	$65,23 \pm 0,07$	86,09 ± 0,02	97,22 ± 0,04	52,23 ± 31,75	65,07 ± 11,87	$49,89 \pm 9,04$	54,88 ± 10,19	74,38 ± 10,14	23,96	23,42	-4,6
G	39,3 ± 0,61	51,44 ± 0,12	70,87 ± 0,1	88,35 ± 0,03	97,7 ± 0,05	49,96 ± 15,15	47,63 ± 11,59	55,08 ± 10,93	74,84 ± 11,75	90,88 ± 11,51	11,01	10,12	0,61
н	24,21 ± 1,25	39,37 ± 0,21	63,62 ± 0,11	85,45 ± 0,04	97,09 ± 0,03	31,89 ± 7,07	43,18 ± 7,61	48,08 ± 9,08	68,82 ± 9,95	96,02 ± 15,16	10,89	8,94	0,78
I	53,98 ± 1,67	63,19 ± 0,23	77,91 ± 0,15	91,16 ± 0,07	98,23 ± 0,09	50,24 ± 13,38	42,57 ± 9,88	57,05 ± 15,33	68,84 ± 11,19	92,08 ± 11,53	16,79	14,74	0,45
J	70,32 ± 1,11	76,26 ± 0,22	85,75 ± 0,18	94,3 ± 0,08	98,86 ± 0,08	53,23 ± 9,6	73,99 ± 11,66	83,71 ± 9,91	102,9 ± 14,01	115,99 ± 16,24	11,57	9,43	0,72

Tabela 5 - ResNet101: predito pelo esperado e acurácia

Fonte: Autoria Própria (2020)

<sup>a</sup> Amido por 100 gramas de pó de açafrão sem mistura ± desvio padrão.

<sup>b</sup> Mistura de X-2 até X-8 correspondente à 2:8 até 8:2 (amido:pó de açafrão), respectivamente.

<sup>c</sup> Média do resultado da predição da concentração de amido para amostra de açafrão sem mistura ± desvio padrão.

<sup>d</sup> Média do resultado da predição da concentração de amido para amostra de açafrão com mistura de X-2 até X-8 correspondente à 2:8 até 8:2 amido:pó de açafrão, respectivamente ± desvio padrão.

Os gráficos da Figura 25 mostram, assim como nas outras duas arquiteturas de ResNet utilizadas, uma dispersão expressiva dos resultados referentes à mesma concentração. A marca A, apesar do desvio padrão, dispõe do resultado das predições em cima da linha de valores ideais. A marca C segue como o que segundo melhor se encaixa na linha. Já a marca F, assim como nas redes anteriores, mantém um resultado impreciso. O comportamento das demais marcas segue paralelo à linha do ideal, mas também sem uma proximidade relevante aos valores esperados.



#### Figura 25 - ResNet101: gráficos de predito pelo esperado

Fonte: Autoria Própria (2020)

## 4.6 Comparativo entre as redes

A tabela a seguir mostra os resultados das métricas de avaliação para todas as redes e para todas as marcas de açafrão. O resultado da predição de todas as imagens foi utilizado para o cálculo, diferente dos resultados apresentados nas tabelas anteriores (feito com as médias).

	AlexNet				ogleNet	t	ResNet18			R	esNet50		ResNet101		
Marca	RMSE	MAE	R²	RMSE	MAE	R²	RMSE	MAE	R²	RMSE	MAE	R²	RMSE	MAE	R²
Α	6,82	5,96	0,93	3,2	2,41	0,98	12,9	10,05	0,78	9,64	7,76	0,88	9,04	7,06	0,89
В	28,89	26,3	0,32	26,27	23,49	0,36	38,91	35,05	0,22	26,93	25,17	0,28	21,69	17,48	0,32
С	11,25	10,09	0,75	5,27	4,45	0,94	15,22	12,42	0,73	10,12	7,95	0,86	14,33	11,38	0,74
D	19,37	17,31	0,38	13,73	12,31	0,65	23,56	21,11	0,53	13,37	11,06	0,72	19,71	15,94	0,04
Е	11,76	10,77	0,65	9,14	8,59	0,8	13,43	10,7	0,75	11	9,18	0,82	16,5	13,78	0,46
F	28,35	24,46	-0,22	23,84	20,82	0	31,54	27,97	0,11	20,7	17,52	0,11	29,29	24,72	-1,22
G	17,24	15,5	0,58	13,22	11,59	0,71	20,88	17,42	0,47	13,74	11,42	0,71	16,44	13	0,41
Н	13,06	11,06	0,68	7,86	6,97	0,88	13,26	10,66	0,77	9,07	7,12	0,89	14,87	12,05	0,65
I	25,95	24,5	0,36	21,67	20,43	0,47	28,71	25,38	0,38	21,44	19,43	0,5	20,83	17,65	0,35
J	16,24	11,94	0,54	11,66	7,58	0,66	19,29	15,64	0,43	12,97	10,26	0,62	17	12,98	0,55
*	18,84	15,24	0,49	15,11	11,34	0,62	22,89	18,16	0,46	15,68	12,41	0,65	18,28	14,19	0,47

Tabela 6 - Acurácia das redes

Fonte: Autoria Própria (2020)

\* Resultados totais calculados para todas as marcas utilizando todos os dados de predição de cada rede.

## **5 CONCLUSÃO**

Este trabalho propôs uma comparação entre algumas arquiteturas de CNN para controle de qualidade na determinação da concentração de amido em imagens digitais de amostras de açafrão de diferentes marcas. Considerando o R<sup>2</sup>, a ResNet50 obteve o melhor resultado geral, seguido da GoogleNet, AlexNet, ResNet101 e ResNet18, respectivamente. Ademais, a sequência do melhor para o pior, considerando o RMSE e o MAE foi: GoogleNet, ResNet50, ResNet101, AlexNet e ResNet18. O açafrão da marca F teve o pior resultado em todas as redes, entre as duas melhores, a ResNet50 foi melhor na predição para essa marca. Comparando os três resultados da ResNet (ResNet18, ResNet50 e ResNet101), pode-se concluir que o aumento das camadas para essa arquitetura não incidiu em uma melhor acurácia.

Para trabalhos futuros, recomenda-se o teste de outras arquiteturas de CNN, combinado a uma quantidade maior de dados por meio de obtenção de mais amostras e/ou melhorias nas técnicas de aumento de dados. Sugere-se também o treinamento com diferentes marcas de açafrão com o objetivo de aumentar a capacidade de generalização das redes.

## REFERÊNCIAS

ANAND, Prashant. Architecture of AlexNet and its current use. **OpenGenus**, [S. I.], [S.D.]. Disponível em: https://iq.opengenus.org/architecture-and-use-of-alexnet/. Acesso em: 14 de nov. de 2020.

BANSAL, Sangita; SINGH, Apoorva; MANGALB, Manisha; MANGAL, Anupam K; KUMAR, Sanjiv. Food Adulteration: Sources, Health Risks and Detection Methods. **Critical Reviews in Food Science and Nutrition**, [S. I.], p. 1174-1189, 9 jun. 2015.

CHEN, Quansheng; ZHANG, Chaojie; ZHAO, Jiewen; OUYANG, Qin. Recent advances in emerging imaging techniques for non-destructive detection of food quality and safety. **Trends in Analytical Chemistry**, [S. I.], v. 52, p. 261-274, dez. 2013.

CNN Padding. **DataHacker**, [S. I.], 01 de nov. 2018. Disponível em: http://datahacker.rs/what-is-padding-cnn/. Acesso em: 13 de nov. de 2020.

DU, Cheng-Jin et al. Comparison of three methods for classification of pizza topping using different colour space transformations. **Journal of Food Engineering**, [S. I.], p. 277-287, 1 maio 2005.

GONZALEZ, Rafael C.; WOODS, Richard E. **Processamento Digital de Imagens**. 3. ed. [S. I.]: Pearson Education do Brasil, 2009.

HE, Kaiming; ZHANG, Xiangyu; REN, Shaoqing; SUN, Jian. DEEP RESIDUAL LEARNING FOR IMAGE RECOGNITION. **IEEE Conference on 291 Computer Vision and Pattern Recognition**, [S. I.], p. 770-778, 10 dez. 2015. DOI https://doi.org/10.1109/CVPR.2016.90.

HINTON, Geoffrey E. et al. Improving neural networks by preventing co-adaptation of feature detectors. **ArXiv**, [S. I.], p. 1-18, 3 jul. 2012.

INOUE, Takashi; NISHIO, Yoshifumi. Applications of Color Image Processing Using Three-Layer Cellular Neural Network Considering HSB Model. **Proceedings of International Joint Conference on Neural Networks**, [S. I.], p. 1335-1342, 19 jun. 2019.

KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. ImageNet Classification with Deep Convolutional Neural Networks. **Advances in neural information processing systems**, [S. I.], p. 1-6, 2012.

LACERDA, Lucas. Deep Learning & Visão Computacional — REDES NEURAIS CONVOLUCIONAIS. **Medium**, [S. I.], 10 de ago. 2019. Disponível em: https://medium.com/@lucaaslb/deep-learning-visão-computacional-redes-neurais-convolucionais-c21f19f5ec34. Acesso em: 14 de nov. de 2020.

LECUN, Yann; BOTTOU, Léon; BENGIO, Yoshua; HAFFNER, Patrick. GRADIENT-BASED LEARNING APPLIED TO DOCUMENT RECOGNITION. **Proc. Of The IEEE**, [S. I.], nov. 1998. DOI https://doi.org/10.1109/5.726791.

LEONEL, Magali; SARMENTO, Silene B.S; CEREDA, Marney P. New starches for the food industry: Curcuma longa and Curcuma zedoaria. **Carbohydrate Polymers**, [S. I.], v. 54, n. 3, p. 385-388, 27 ago. 2003. DOI https://doi.org/10.1016/S0144-8617(03)00179-6.

LOPEZ-MOLINERO, Angel; LIÑAN, David; SIPIERA, Daniel; FALCON, Raquel. CHEMOMETRIC INTERPRETATION OF DIGITAL IMAGE COLORIMETRY APPLICATION FOR TITANIUM DETERMINATION IN PLASTICS. **Microchemical Journal**, [S. I.], p. 380-385, 1 jul. 2010. DOI https://doi.org/10.1016/j.microc.2010.06.013.

MACÊDO, Isaac Yves Lopes de; MACHADO, Fabio Bahls; RAMOS, Gabrielle Santos; COSTA, André Gabriel do Carmo; BATISTA, Rayssa Dias; FILHO, Arlindo Rodrigues Galvão; ASQUIERI, Eduardo Ramirez; DE SOUZA, Aparecido Ribeiro; DE OLIVEIRA, Anselmo Elcana; GIL, Eric de Souza. Starch adulteration in turmeric samples through multivariate analysis with infrared spectroscopy. **Food Chem**, [S. I.], v. 340, p. 1-8, 25 ago. 2020. DOI https://doi.org/10.1016/j.foodchem.2020.127899.

Modelos de cor: HSB ou HSV. **AlfaMidia**, [S. I.], 15 de jan. 2019. Disponível em: http://www.alfamidia.com.br/15-01-2019-modelos-de-cor-hsb-ou-hsv/. Acesso em: 04 de nov. de 2020.

NG, Wartini; MINASNY, Budiman; MONTAZEROLGHAEM, Maryam; PADARIAN, Jose; FERGUSON, Richard; BAILEY, Scarlett; MCBRATNEY, Alex B. CONVOLUTIONAL NEURAL NETWORK FOR SIMULTANEOUS PREDICTION OF SEVERAL SOIL PROPERTIES USING VISIBLE/NEAR-INFRARED, MID-INFRARED, AND THEIR COMBINED SPECTRA. **Geoderma**, [S. I.], p. 251-267, 2019. DOI https://doi.org/10.1016/j.geoderma.2019.06.016.

O'MAHONY, Niall et al. DEEP LEARNING VS. TRADITIONAL COMPUTER VISION. In: O'MAHONY, Niall. Advances in Intelligent Systems and Computing: Selected Papers from the International Conference on Computer Science and Information Technologies. [S. I.]: **Natalia Shakhovska**, 2019. v. 943, p. 487-498.

O'SHEA, Keiron; NASH, Ryan. An Introduction to Convolutional Neural Networks. **ArXiv**, [S. I.], p. 1-10, 26 nov. 2015.

OLGA, Russakovsky; JIA, Deng et al. IMAGENET LARGE SCALE VISUAL RECOGNITION CHALLENGE. International Journal of Computer Vision. v. 115. [S. I.], p. 1-19, 11 abr. 2015. DOI https://doi.org/10.1007/s11263-015-0816-y.

PARVATHY, V. A.; SWETHA, V. P. et al. Detection of plant-based adulterants in turmeric powder using DNA barcoding. **Pharmaceutical Biology**, [S. I.], v. 53, p. 1774-1779, 5 mai. 2015. DOI https://doi.org/10.3109/13880209.2015.1005756.

PAULA, LUANA NASCIMENTO DE; ALVES, ADRIANO ROSA; NANTES, ELIZA ADRIANA SHEUER. A IMPORTÂNCIA DO CONTROLE DE QUALIDADE EM INDÚSTRIA DO SEGMENTO ALIMENTÍCIO. **Revista Conhecimento Online**, [S. I.], 2017. DOI 10.25112/rco.v2i0.1077.

PONTI, Moacir Antonelli; COSTA, Gabriel B. Paranhos da. Como funciona o deep learning. **Tópicos em gerenciamento de dados e informações,** 2017, [S.I.], p. 63-88, 2017.

ROCHA, TIAGO GALDINO; GALENDE, SHARIZE BETONI. A IMPORTÂNCIA DO CONTROLE DE QUALIDADE NA INDÚSTRIA FARMACÊUTICA. **Revista UNINGÁ Review**, [S. I.], 13 out. 2014. DOI 10.25112/rco.v2i0.1077.

SAKTHI, R. Talented Mr. 1X1: Comprehensive look at 1X1 Convolution in Deep Learning. **Analytics Vidhya**, [S. I.], 18 de out. 2018. Disponível em: https://medium.com/analytics-vidhya/talented-mr-1x1-comprehensive-look-at-1x1-convolution-in-deep-learning-f6b355825578. Acesso em: 04 de nov. de 2020.

SELEME, Robson et al. Controle da Qualidade. [S. l.]: InterSaberes, 2012.

SHAIKH, F. Deep Learning in the Trenches: Understanding Inception Network from Scratch. **Analytics Vidhya**, [S. I.], 18 de out. 2018. Disponível em:

https://www.analyticsvidhya.com/blog/2018/10/understanding-inception-network-from-scratch/. Acesso em: 04 de nov. de 2020.

SHISHKIN, Yu. L.; DMITRIENKO, S. G.; MEDVEDEVA, O. M.; BADAKOVA, S. A.; PYATKOVA, L. N. Use of a Scanner and Digital Image-Processing Software for the Quantification of Adsorbed Substances. **Journal of Analytical Chemistry**, [S. I.], p. 102-106, 21 maio 2003.

SHORTEN, Connor; KHOSHGOFTAAR, Taghi M. A survey on Image Data Augmentation for Deep Learning. **Journal of Big Data**, [S. I.], p. 1-48, 6 jul. 2019. DOI https://doi.org/10.1186/s40537-019-0197-0.

SINGLETARY, Keith. Turmeric: An Overview of Potential Health Benefits. Nutrition Today, [S. l.], v. 45, n. 5, p. 216-225, 1 out. 2010.

SOROURADDIN, Mohammad-Hossein; SAADATI, Masoud; MIRABI, Fariba. SIMULTANEOUS DETERMINATION OF SOME COMMON FOOD DYES IN COMMERCIAL PRODUCTS BY DIGITAL IMAGE ANALYSIS. Journal of Food and Drug Analysis (JFDA), [S. 1-8, 2 jan. 2015. DOI I.], р. http://dx.doi.org/10.1016/j.jfda.2014.10.007.

SUN, JIAN-PING; HOU, CAI-YUN; FENG, JING; WANG, XU. DETERMINATION OF THE PROTEIN CONTENT IN RICE BY THEDIGITAL CHROMATIC METHOD. **Journal of Food Quality**, [S. I.], v. 31, n. 2, p. 250-263, 28 abr. 2008.

SZEGEDY, Christian et al. Going Deeper with Convolutions. **2015 IEEE Conference** on Computer Vision and Pattern Recognition, [S. I.], p. 1-8, 17 set. 2014.

VALE, Mayara. O QUE É A ANVISA? O QUE ELA FISCALIZA?. **Mayara Vale Consultora de Alimentos**, [S. I.], 2015. Disponível em: https://consultoradealimentos.com.br/consultoria/o-que-e-a-anvisa/. Acesso em: 06 de dez. de 2020.

WEISSTEIN, Eric W. CONVOLUTION. **MathWorld--A Wolfram**, [S. I.], [S.D.]. Disponível em: https://mathworld.wolfram.com/Convolution.html. Acesso em: 20 de out. de 2020.