



UNIVERSIDADE DO ESTADO DA BAHIA – UNEB
DEPARTAMENTO DE CIÊNCIAS EXATAS E DA TERRA- CAMPUS I
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA APLICADA

FABRINA OLIVEIRA PARANHOS

**REDES NEURAIS CONVOLUCIONAIS PARA AVALIAÇÃO DA
ADULTERAÇÃO DE *SPIRULINA* ATRAVÉS DE IMAGENS DIGITAIS**

SALVADOR

2023

FABRINA OLIVEIRA PARANHOS

**REDES NEURAIS CONVOLUCIONAIS PARA AVALIAÇÃO DA
ADULTERAÇÃO DE *SPIRULINA* ATRAVÉS DE IMAGENS DIGITAIS**

Dissertação apresentada ao programa de Pós-Graduação em Química Aplicada da Universidade do Estado da Bahia, orientada pelo Prof. Dr. Fabio de Souza Dias como parte dos requisitos para obtenção do título de Mestre em Química.

Área: Química Analítica.

SALVADOR

2023

FICHA CATALOGRÁFICA
Biblioteca Professor **Edivaldo Machado Boaventura** - UNEB – Campus I
Bibliotecária: Célia Maria da Costa – CRB5/918

P223r Paranhos, Fabrina Oliveira

Redes neurais convolucionais para avaliação da adulteração de spirulina através de imagens digitais / Fabrina Oliveira Paranhos . – Salvador, 2023.

78 f. : il.

Orientador: Fabio de Souza Dias

Dissertação (Mestrado Acadêmico). Universidade do Estado da Bahia.
Departamento de Ciências Exatas e da Terra. Programa de Pós-Graduação em
Química Aplicada – PGQA. Campus I. 2023.

Contém referências.

1. Spirulina. 2. Contaminação de alimentos. 3. Processamento de imagens – Técnicas digitais. 4. Inteligência artificial. I. Dias, Fábio de Souza. II. Universidade do Estado da Bahia. Departamento de Ciências da Vida. Campus I. III. Título.

CDD: 543

FABRINA OLIVEIRA PARANHOS

REDES NEURAIS CONVOLUCIONAIS PARA AVALIAÇÃO DA ADULTERAÇÃO DE *SPIRULINA* ATRAVÉS DE IMAGENS DIGITAIS

Dissertação apresentada ao programa de Pós-Graduação em Química Aplicada da Universidade do Estado da Bahia, orientada pelo Prof. Dr. Fabio de Souza Dias como parte dos requisitos para obtenção do título de Mestre em Química.

Área: Química Analítica.

APROVADA EM 27/09/2023.

BANCA EXAMINADORA

Prof. Dr. Fabio de Souza Dias (Orientador)

Universidade do Estado da Bahia

Prof. Dr. Anibal de Freitas Santos Junior (Avaliador Interno)

Universidade do Estado da Bahia

Prof^a. Dr^a. Elane Santos da Boa Morte (Avaliadora Externa)

Instituto Federal de Educação, Ciência e Tecnologia da Bahia

AGRADECIMENTOS

Agradeço, primeiramente, a Deus. Ele que sempre me apresenta as melhores oportunidades, me sustenta em todos os projetos que abraço e me dá forças para conciliar tantas demandas simultaneamente.

Agradeço a meu marido Alan e minha princesa Isa. É por nosso futuro que tantos sacrifícios momentâneos são feitos! Em breve continuaremos colhendo os frutos de nossas conquistas. Eu amo vocês.

Agradeço a minha mãe e as minhas tias por serem fontes de inspiração para mim, com a força, garra e determinação que possuem. Vocês me ensinam todos os dias através do exemplo! Amo vocês!

Obrigada, também, a todos os meus amigos que seguraram minha mão, me incentivaram e sempre me colocaram para cima quando achei que não tinha mais forças para seguir. Carol, Jessica, Mylena, Elivaldo, Talita, Taíssa e Lane, o meu muito obrigada!

Quero agradecer imensamente a meu orientador, professor Fábio, por ter acreditado em mim, por traçar estratégias sempre avaliando minhas condições no momento! Obrigada professor! Tenho certeza que se não fossem a sua sensibilidade e perspicácia eu teria desistido! A Academia precisa de mais professores como você! És uma inspiração da profissional em Educação que almejo ser!

Por fim, agradeço ao Programa de Pós-Graduação em Química Aplicada pela oportunidade e à todos os professores que contribuíram para a culminância desse projeto de vida! Porque muito mais do que um título de Mestra, findo um ciclo de uma realização pessoal.

Meu muito obrigada!

RESUMO

A *Arthrospira (Spirulina) platensis* é um alimento que vem ganhando espaço na indústria de suplementos por seu alto valor nutritivo. É a cianobactéria mais utilizada no consumo humano e animal e contém alto teor de proteínas e aminoácidos e quantidades significativas de betacaroteno, vitaminas do complexo B, ferro, minerais essenciais e ácido gama-linolênico. Dada sua importância, é um alimento com alto valor de mercado e que pode ser alvo de adulterações, comprometendo a segurança alimentar de seus consumidores. Assim detectá-las se faz de extrema importância e necessidade. Este estudo objetivou avaliar de forma controlada, adulteração de *Spirulina* com bicarbonato de sódio, em percentuais de 10%, 15% e 25% m/m. A escolha do adulterante se deu pelo fato do mesmo participar do processo de cultivo desta microalga, por ser de baixo custo e não higroscópico. Um sistema de captação de imagens digitais foi montado, testes de otimização do mesmo foram realizados. Realizou-se a captação de imagens digitais das amostras a 0%, 10%, 15% e 25% m/m. Os dados fornecidos pelas imagens digitais nos canais RGB foram analisados através do *software ChemoStat* e os gráficos de análise de componentes principais (PCA) e análise de agrupamento hierárquico (HCA) foram plotados. Estes métodos quimiométricos de análise exploratória mostraram-se eficientes quando os estudos foram realizados com poucas amostras, contudo quando do aumento destas (120 amostras com captações realizadas em triplicata, totalizando 360 captações), os gráficos não as classificaram bem, houve aglomeração de dados, dificultando a análise dos mesmos. Neste momento, a Inteligência Artificial, mais especificamente as Redes Neurais Convolucionais, foi utilizada. Os algoritmos de classificação utilizados, ResNet-18 e ResNet-50, apresentaram resultados satisfatórios na classificação das amostras. Ambos tiveram uma média de 98,9% de precisão na validação e 88,5% e 98%, respectivamente, na média da precisão dos testes. Desse modo, as Redes Neurais Convolucionais associadas às imagens digitais foram muito eficientes na classificação das amostras adulteradas frente ao uso de PCA e HCA, podendo ser uma aliada com grande potencial de crescimento e diversificação em análises, amplamente utilizadas nas diversas áreas dentro da Química.

Palavras-chave: *Spirulina*, Adulteração de Alimentos, Imagens Digitais, Inteligência Artificial, Redes Neurais Convolucionais.

ABSTRACT

Arthrospira (Spirulina) platensis is a food that has been gaining ground in the supplement industry due to its high nutritional value. It is the cyanobacteria most used in human and animal consumption and contains a high content of proteins and amino acids and significant amounts of beta-carotene, B vitamins, iron, essential minerals and gamma-linolenic acid. Given its importance, it is a food with high market value and can be subject to adulteration, compromising the food safety of its consumers. Therefore, detecting them is extremely important and necessary. This study aimed to evaluate, in a controlled manner, adulteration of *Spirulina* with sodium bicarbonate, in percentages of 10%, 15% and 25% m/m. The adulterant was chosen because it participates in the cultivation process of this microalgae, as it is low cost and non-hygroscopic. A digital image capture system was assembled and optimization tests were carried out. Digital images of samples were captured at 0%, 10%, 15% and 25% m/m. The data provided by digital images in RGB channels were analyzed using *ChemoStat software* and principal component analysis (PCA) and hierarchical cluster analysis (HCA) graphs were plotted. These chemometric methods of exploratory analysis proved to be efficient when studies were carried out with few samples, however when these were increased (120 samples with captures carried out in triplicate, totaling 360 captures), the graphs did not classify them well, there was data agglomeration, making their analysis difficult. At this time, Artificial Intelligence, more specifically Convolutional Neural Networks, was used. The classification algorithms used, ResNet-18 and ResNet-50, presented satisfactory results in classifying the samples. Both had an average of 98.9% validation accuracy and 88.5% and 98%, respectively, in average test accuracy. In this way, Convolutional Neural Networks associated with digital images were very efficient in classifying adulterated samples compared to the use of PCA and HCA, and can be an ally with great potential for growth and diversification in analyses, widely used in different areas within Chemistry.

Keywords: *Spirulina*, Food Adulteration, Digital Images, Artificial Intelligence, Convolutional Neural Networks.

LISTA DE FIGURAS

Figura 1:	Estrutura da <i>Spirulina</i> em forma de espiral, originando seu nome.	20
Figura 2:	Mulheres Kanembous secando a <i>Spirulina</i> . Registro feito em 1963.	21
Figura 3:	Mulheres Kanembous colhendo e secando a <i>Spirulina</i> de forma tradicional. Registro feito em 2004.	21
Figura 4:	Espectro Eletromagnético.	25
Figura 5:	Círculos cromáticos dos sistemas de cores.	25
Figura 6:	Representação de imagem digital bidimensional. 6.a a imagem está decomposta no verde, 6.b no vermelho, 6.c no azul e 6.d representa uma imagem digital bidimensional oriunda da composição das três cores do sistema RGB.	27
Figura 7:	Imagem digital com ampliação para visualização dos <i>pixel</i> .	28
Figura 8:	Qualidade da imagem x quantidade <i>pixels</i> .	28
Figura 9:	Estrutura básica neurônio artificial. O vetor de entrada é multiplicado por pesos. Os elementos ponderados e um limiar de ativação (<i>bias</i>) são somados e o potencial de ativação resultante alimenta uma função de ativação. A saída é o resultado da transformação do potencial de ativação.	34
Figura 10:	Tipos de rede neural profunda. 10.a rede neural simples. 10.b rede neural multicamadas.	34
Figura 11:	Arquitetura básica de uma rede neural convolucional.	36
Figura 12:	Divisão em porções das amostras de <i>Spirulina</i> com os percentuais de adulteração.	40
Figura 13:	Placa de polietileno com tampa, de cor transparente, medindo 5 x 5 x 1 cm contendo amostra de <i>Spirulina</i> .	40
Figura 14:	Caixa para captação das amostras. 14.a caixa parte interna, medindo 36x22x13 cm. 14.b luminária localizada a 20 cm centro da captação das imagens. 14.c caixa parte externa.	42

Figura 15:	Amostras de <i>Spirulina</i> com e sem tratamento no <i>ChemoStat</i> . 15.a amostra 0% sem tratamento, 15.b amostra adulterada em 10% m/m sem tratamento, 15.c amostra adulterada em 15% m/m sem tratamento, 15.d amostra adulterada em 25% m/m sem tratamento, 15.e foto da amostra 0% tratada no <i>ChemoStat</i> (500x500 <i>pixels</i> e ROI 150X50), 15.f foto da amostra adulterada em 10% m/m tratada no <i>ChemoStat</i> (500x500 <i>pixels</i> e ROI 150X50), 15.g foto da amostra adulterada em 15% m/m tratada no <i>ChemoStat</i> (500x500 <i>pixels</i> e ROI 150X50) e 15.h foto da amostra adulterada em 25% m/m tratada no <i>ChemoStat</i> (500x500 <i>pixels</i> e ROI 150X50) (Autoria Própria).	44
Figura 16:	Histograma de cores das amostras. 16.a: amostra 0%. 16.b: amostra adulterada em 10% m/m. 16.c: amostra adulterada em 15% m/m e 16.d: amostra adulterada em 25% m/m.	45
Figura 17:	Arquitetura dos algoritmos das redes neurais residuais ResNet18 e ResNet50.	66
Figura 18:	Matriz de Confusão ResNet18.	67
Figura 19:	Matriz de Confusão ResNet50.	67
Figura 20:	Dados da Matriz de Confusão ResNet18.	68
Figura 21:	Dados da Matriz de Confusão ResNet50.	68

LISTA DE TABELAS

Tabela 1:	Cores primárias com respectivos comprimentos de onda. Combinação das cores primárias e cor secundária resultante.	26
Tabela 2:	Descrição das amostras de <i>Spirulina</i> usadas. Origem, quantidade adquirida em cada localidade e data de aquisição das amostras.	38
Tabela 3:	Meio Zarrouk modificado.	39
Tabela 4:	Quantificação, peso e abreviação atribuída as amostras após a divisão.	41
Tabela 5:	Descrição dos testes de otimização do sistema de captação.	43
Tabela 6:	Principais parâmetros utilizados na CNN.	67
Tabela 7:	Comparação da precisão da validação e precisão do teste para as arquiteturas ResNet18 e ResNet50 entre categorias.	69

LISTA DE GRÁFICOS

Gráfico 1:	Quantidade de publicações x áreas de publicações. Estudos utilizando as Redes Neurais Convolucionais. Levantamento feito no <i>Science Direct</i> em publicações de 2013 a 2024	37
Gráfico 2:	PCA Teste 1 – Área de captação 32 x 32 (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m)	47
Gráfico 3:	HCA Teste 1 – Área de captação 32 x 32 (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana	47
Gráfico 4:	PCA Teste 2 – Área de captação 64 x 64 (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m)	48
Gráfico 5:	HCA Teste 2 – Área de captação 64 x 64 (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana	48
Gráfico 6:	PCA Teste 3 – Altura do porta-amostra de 1,3 cm da placa + 1,6 cm de isopor (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m)	49
Gráfico 7:	HCA Teste 3 – Altura do porta-amostra de 1,3 cm da placa + 1,6 cm de isopor (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m). Método Complete Linkage. Distância Euclidiana	50
Gráfico 8:	PCA Teste 4 - Altura do porta-amostra de 1,3 cm da placa + 3,2 cm de isopor (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m)	50
Gráfico 9:	HCA Teste 4 – Altura do porta-amostra de 1,3 cm da placa + 3,2 cm de isopor isopor (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m). Método Complete Linkage. Distância Euclidiana	51
Gráfico 10:	PCA Teste 5 – Amostra Prensada (SP em verde amostra	52

	0% e SV em vermelho amostra adulterada em 25% m/m)	
Gráfico 11:	HCA Teste 5 – Amostra Prensada (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana	52
Gráfico 12:	PCA Teste 6 – Amostra Sem Prensa (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m)	53
Gráfico 13:	HCA Teste 6 – Amostra Sem Prensa (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana	53
Gráfico 14:	PCA Teste 7 - PCA: Auto Histograma R, G, (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m)	54
Gráfico 15:	HCA Teste 7 - Complete Histograma RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana	55
Gráfico 16:	HCA Teste 7 - Single Histograma RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Single Linkage. Distância Euclidiana	55
Gráfico 17:	HCA Teste 7 - Average Histograma RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Average Linkage. Distância Euclidiana	58
Gráfico 18:	PCA Teste 8 - PCA: Mean Histograma R, G, B (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m)	58
Gráfico 19:	HCA Teste 8 - Complete Histograma RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana	57
Gráfico 20:	HCA Teste 8 - Single Histograma RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Single Linkage. Distância Euclidiana	57
Gráfico 21:	HCA Teste 8 - Average Histograma (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m)	58

	Método Average Linkage. Distância Euclidiana	
Gráfico 22:	PCA Teste 9 - PCA: Auto Single Channels R, G, (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m)	58
Gráfico 23:	HCA Teste 9 - Complete Single Channels RGB ((SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana	59
Gráfico 24:	HCA Teste 9 - Single Single Channels RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Single Linkage. Distância Euclidiana	59
Gráfico 25:	HCA Teste 9 - Average Single Channels RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Average Linkage. Distância Euclidiana	60
Gráfico 26:	PCA Teste 10 - PCA: Mean Single Channels R, G, (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m)	60
Gráfico 27	HCA Teste 10 - Complete Single Channels RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana	61
Gráfico 28	HCA Teste 10 - Single Single Channels RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Single Linkage. Distância Euclidiana	61
Gráfico 29:	HCA Teste 10 - Average Single Channels RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Average Linkage. Distância Euclidiana	62
Gráfico 30:	PCA - 80 amostras (SP em verde amostra 0%, SD em roxo amostra adulterada em 10% m/m e SV em vermelho amostra adulterada em 25% m/m)	63
Gráfico 31:	HCA - 80 amostras (SP em verde amostra 0%, SD em	63

roxo amostra adulterada em 10% m/m e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana

Gráfico 32: PCA – 360 captações (SP em verde amostra 0%, SD em roxo amostra adulterada em 10% m/m, SQ em rosa amostra adulterada em 15% m/m e SV em vermelho amostra adulterada em 25% m/m) 64

Gráfico 33: HCA – 360 captações (SP em verde amostra 0%, SD em roxo amostra adulterada em 10% m/m, SQ em rosa amostra adulterada em 15% m/m e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana 65

LISTA DE SIGLAS

UV	Raios Ultravioleta
HPLC	Cromatografia Líquida de Alta Eficiência
RMN	Ressonância Magnética
PCR	Reação em Cadeia de Polimerase
RGB	Red, Green and Blue (Vermelho, Verde e Azul)
CMY	Cyan, Magenta and Yellow (Ciano, Magenta e Amarelo)
UV-Vis	Ultravioleta e Visível
ID	Imagens Digitais
CMYK	Cyan, Magenta, Yellow and Black (Ciano, Magenta, Amarelo e Preto)
pH	Potencial Hidrogeniônico
NFC	Comunicação por campo de proximidade
PCA	Análise por Componentes Principais
HCA	Agrupamento por Análise Hierárquica
DD-SIMCA	Modelagem independente suave orientada por dados de analogia de classe
OC-PLS	Mínimos quadrados parciais de classe
IA	Inteligência Artificial
CNN	Redes Neurais Convolucionais
UFRB	Universidade Federal do Recôncavo Baiano
SERS	Espectroscopia Raman de superfície aprimorada
ROI	Região de Interesse
SP	Amostras Puras
SD	Amostras Adulteradas em 10%
SQ	Amostras Adulteradas em 15%
SV	Amostras Adulteradas em 25%
LED	Diodo emissor de luz
ResNet	Redes neurais residuais

SUMÁRIO

1	INTRODUÇÃO	17
2	OBJETIVOS	19
2.1	Objetivo Geral.....	19
2.2	Objetivo Específico.....	19
3	FUNDAMENTAÇÃO TEÓRICA	19
3.1	<i>Spirulina</i>	19
3.2	Adulteração de alimentos.....	22
3.3	Cor e sistemas de cores.....	24
3.4	Imagens Digitais e a Química Analítica.....	28
3.5	Análise Multivariada.....	30
3.6	Inteligência Artificial.....	31
4	MATERIAIS E MÉTODOS	38
4.1	Amostras.....	38
4.2	Sistema de captação.....	41
4.3	Avaliação das variáveis do sistema de captação e tratamento das imagens digitais.....	42
5	RESULTADOS E DISCUSSÕES	44
5.1	Amostras.....	44
5.2	Avaliação das variáveis do sistema de captação.....	46
5.2.1	Área de captação.....	46
5.2.2	Altura do porta-amostra.....	49
5.2.3	Disposição da amostra.....	51
5.2.4	Método de Aglomeração de Dados.....	54
5.3	PCA, HCA e CNN.....	62
6	CONCLUSÃO	69
	REFERÊNCIAS	71

1 INTRODUÇÃO

A *Spirulina* é um alimento que vem ganhando espaço na indústria de suplementos por seu alto valor nutritivo. *Arthrospira platensis* e *Arthrospira maxima* são os dois tipos mais comercializados da *Spirulina* (LAFARGA *et al.* 2020). Trata-se de uma alga azul-esverdeada comestível, não fixadora de nitrogênio, que possui um filo de microrganismos que estão relacionados com as bactérias, mas são capazes de realizar a fotossíntese (HENRIKSON, 2009; LAFARGA *et al.* 2020).

É a cianobactéria mais utilizada no consumo humano e animal. Contém alto teor de proteínas e aminoácidos (cerca de 65%) contra 15% a 25% de carne de animal ou peixe, 35% de soja e 12% de ovos (COSTA *et al.*, 2019; HENRIKSON, 2009). Além das proteínas, o consumo da *Spirulina* proporciona quantidades de betacaroteno, vitaminas do complexo B, ferro, minerais essenciais e ácido gama-linolênico (HENRIKSON, 2009).

Uma das preocupações de governos, órgãos fiscalizadores e instituições não governamentais é a segurança alimentar da população (ESPINOZA *et al.*, 2015). Fraudes alimentares são cada vez mais comuns e perigosas (SPINK; MOYER, 2011) e podem ser classificadas em adulteração, manipulação, furto e desvio (ESPINOZA *et al.*, 2015; SPINK; MOYER, 2011).

Métodos analíticos tradicionais, como os físico-químicos, espectroscópicos, cromatográficos e de ressonância magnética são importantes aliados na detecção de fraudes alimentares.

Métodos não tradicionais de análise, como o uso de imagens digitais estão em ascensão e ganham visibilidade dentro da Química Analítica. Imagens digitais já foram utilizadas no controle de qualidade do meio ambiente: Choodum *et al.* (2013) e Zarghampour *et al.* (2020); em análise forense: Thanakiatkrai *et al.* (2013) e da Silva *et al.* (2018); em avaliação da qualidade de produtos farmacêuticos: Lamarca *et al.* (2019) e Coutinho *et al.* (2019) e no controle de qualidade de combustível: Ribeiro *et al.* (2019) e João *et al.* (2019). Alguns estudos já realizaram análise de alimentos utilizando imagens digitais, como os trabalhos de Botelho *et al.* (2017), Lima *et al.* (2020), Pierini *et al.* (2016) e Reile *et al.* (2020).

Alcántara (2018) realizou uma revisão sobre os métodos de análise de elementos potencialmente tóxicos na produção da *Spirulina*, mas até o momento não há trabalhos sobre a análise de adulteração desse alimento, fazendo com que este estudo seja inovador.

O objetivo deste estudo foi utilizar Redes Neurais Convolucionais para avaliar adulteração em *Spirulina* através de imagens digitais. A proposta deste trabalho trata-se de uma metodologia inovadora, visto que não há nenhum estudo que tenha associado imagens digitais à inteligência artificial na análise de adulteração de alimentos. A escolha do analíto justifica-se pela importância nutricional que a *Spirulina* tem, além de seu alto valor de mercado, podendo ser alvo de adulteração. Por esse motivo identificar adulterações nesse alimento é de extrema importância para garantir a segurança alimentar de seus consumidores.

O trabalho foi estruturado da seguinte maneira: na Fundamentação Teórica, falou-se sobre a *Spirulina*, seus benefícios e propriedades; posteriormente abordou-se sobre a adulteração de alimentos seus malefícios à saúde pública e da importância de sua detecção, bem como um levantamento de estudos que já realizaram análises de adulteração alimentar. Há também uma exposição de conceitos de cor, seus sistemas e elementos que compõem uma imagem digital, embasando, desse modo, uma discussão sobre imagens digitais e sua aplicação na Química Analítica. A Análise Multivariada também foi estudada. Nesse momento houve exposição de diversos trabalhos que utilizam deste recurso da Quimiometria para analisar adulteração em alimentos. Para finalizar a Fundamentação Teórica, abordou-se sobre a Inteligência Artificial, seu conceito, tipos de redes e aplicações nas mais diversas áreas de pesquisa, inclusive na Química. A Metodologia foi apresentada no capítulo 4 e os Resultados e Discussões no capítulo 5.

2. OBJETIVOS

2.1 Objetivo Geral

Avaliar adulteração em *Spirulina* através de imagens digitais associadas às redes neurais convolucionais.

2.2 Objetivos Específicos

- Propor uma metodologia de baixo custo, via imagens digitais, para detecção de adulteração em alimentos;
- Analisar a eficiência da Inteligência Artificial na identificação de adulteração alimentar;
- Demonstrar como as imagens digitais e a Inteligência Artificial são aliadas em análises químicas *in situ*.

3. FUNDAMENTAÇÃO TEÓRICA

3.1 *Spirulina*

A *Spirulina* é um alimento que vem ganhando espaço na indústria de suplementos por seu alto valor nutritivo (LAFARGA *et al.*, 2020). *Spirulina platensis* é uma microalga azul-esverdeada; uma cianobactéria que pode chegar a ter comprimento entre 50 e 500 μm e largura entre 3 e 4 μm . *Spirulina* vem do latim e significa espiral ou hélice, daí a origem de seu nome, pois possui filamentos em forma de espiral (Figura 1) (JUNG *et al.*, 2019).

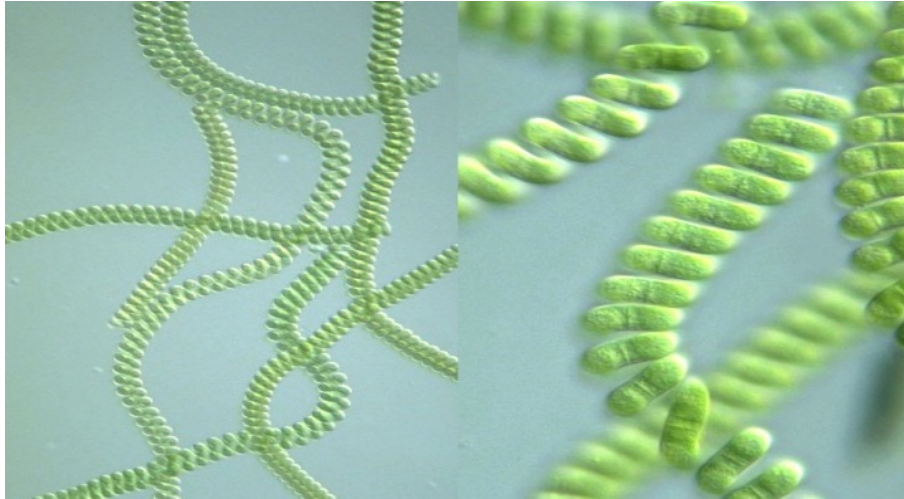


Figura 1: Estrutura da *Spirulina* em forma de espiral, originando seu nome (http://protist.i.hosei.ac.jp/PDB/Images/Prokaryotes/Oscillatoriaceae/Spirulina_2.jpg).

A *Spirulina* tem origem em lagoas da África e da América Latina e foi usada como alimento no México pelos astecas e pelos índios Kanembous até o século XVI (COSTA, *et al.*, 2019; JUNG *et al.*, 2019). O ficologista francês Dangeard foi o primeiro a realizar estudos sobre a *Spirulina*, em 1940. Nesse estudo, ele relatou a existência de bolos endurecidos de algas verdes secas ao sol, Dihé na língua local, consumido pelos povos Kanembous que viviam às margens do Lago Chade. Na Figura 2 tem-se registros feitos em 1963 de mulheres do povo Kanembous secando os bolos de *Spirulina*. Já a figura 3 foi um registro feito em 2004, onde mostra as mulheres deste mesmo povo colhendo e secando a *Spirulina* de forma tradicional.



Figura 2: Mulheres Kanembous secando a *Spirulina*. Registro feito em 1963 (<https://www.spirulinasource.com/wp-content/uploads/2023/03/Africa-Spirulina-Industry-3-scaled.jpg>).



Figura 3: Mulheres Kanembous colhendo e secando a *Spirulina* de forma tradicional. Registro feito em 2004 (<https://www.spirulinasource.com/wp-content/uploads/2023/03/Africa-Spirulina-Industry-4-scaled.jpg>).

Quase 25 anos depois dos estudos realizados por Dangeard, o botânico J. Leonard, participante de uma expedição belga às margens do lago Chade, notou os bolos de algas verdes e assim, a microalga foi se popularizando, embora já fosse bastante consumida e vendida pelos povos supracitados (HENRIKSON, 2009). Mas, foi em 1967 que foi reconhecida como alimento pela Associação Internacional de Microbiologia Aplicada (COSTA *et al.*, 2019).

No que tange a composição bioquímica, a *Spirulina* tem um valor de proteínas bastante elevado, chegando a 65% em sua composição de proteínas

e aminoácidos, 20% desta microalga é composta de carboidratos, 7% de minerais e 5% de gorduras. Além disso, é rica em vitaminas A, C, E, K e do complexo B. Possui altos teores de ferro, cálcio, fósforo e magnésio (HENRIKSON, 2009).

Além de seu alto valor nutricional, essa microalga pode ser usada para fins medicinais. Já existem estudos que mostram a eficiência na sua utilização para prevenção e controle de doenças cardiovasculares, doenças inflamatórias, alergias e infecções virais (DARTSCH, 2008; KARKOSE *et al.*, 2008; KULSHRESHTHA *et al.*, 2008). Estudos realizados por Loseva e Dardynskaya (1993) mostram efetividade no uso de *Spirulina* no tratamento de enjoos provocados pela exposição à radiação pós desastre de Chernobyl.

Por essas razões, a *Spirulina* é um alimento com alto valor de mercado. 150 gramas de *Spirulina* orgânica em pó pode chegar a R\$ 86,00, enquanto ela em cápsulas chega a custar R\$ 140,00 com 200 doses. Assim, analisar e identificar adulterações neste alimento são de extrema importância para garantir a qualidade do alimento a seus consumidores.

3.2 Adulteração de alimentos

A adulteração de alimentos é uma das formas de fraude que comprometem a segurança alimentar da população por todo o mundo (EVERSTINE, 2013). Fraude alimentar é o termo usado para designar as mais diversas formas de manipulação indevida nos alimentos e são comumente classificadas em adulteração, manipulação, furto, desvio, simulação e falsificação (ESPINOZA *et al.*, 2015; SPINK; MOYER, 2011). Spink *et al.* (2016) compilaram os tipos de fraude e suas características. Segundo os autores, adulteração diz respeito à adição de uma substância ou componente alimentar que resulta em um produto fraudulento; manipulação ocorre quando uma embalagem fraudulenta é produzida; o furto acontece quando um produto original é roubado e vendido como adquirido por vias legais; o desvio se caracteriza pela comercialização de produtos originais em mercados clandestinos; a simulação diz respeito a um produto parecer legítimo, mas não é; e quanto à falsificação esta ocorre quando tanto o produto quanto a

embalagem são fraudados de forma a se assemelhar ao produto original (SPINK *et al.*, 2016).

Fraudes alimentares podem ser detectadas conforme sua quantidade, qualidade (sensorial, nutricional e/ou tecnológica), pureza (contaminação, adulteração), conservação e identidade (substituição de uma espécie por outra) (ESPINOZA *et al.*, 2015). No Brasil, de 2007 a 2017, dentre os tipos de fraudes relatadas em produtos alimentícios e compilados por Tíbola *et al.* (2018), têm-se que 50% diz respeito a diluição (adulteração), 22% a substituição, 14% aprimoramento artificial, falsificação e rotulagem incorreta (ambos com 7%).

As fraudes alimentares são enquadradas como crimes lucrativos (TIBOLA *et al.*, 2018). Assim, produtos alimentícios que possuem valor de mercado são alvos de fraude alimentar frequentemente. O objetivo dos fraudadores é enganar os consumidores a pensar que estão comprando produtos autênticos e seguros (GALVIN-KING *et al.* 2020; GALVIN-KING *et al.* 2018). Tibola *et al.* (2018) fizeram um levantamento bibliográfico dos estudos acerca dos alimentos mais fraudados no Brasil entre 2007 e 2017. Nesses 10 anos, 38% dos relatos são de adulterações em leite e derivados, tendo como adulterante mais comum à adição de água; 21% em óleos vegetais, onde há a diluição e substituição do azeite de oliva por azeites refinados e outros óleos, como óleo de soja; bebidas e café - ambas com 10% - neste caso, nas bebidas, incluindo cerveja e suco de frutas, os estudos apontam que a maior incidência de fraudes são na rotulagem e no caso do café é na adição de cevada, milho e arroz; peixes e frutos do mar com 9% dos casos de adulteração estudados, onde há a substituição de uma espécie de peixe de alto custo por outra de menor preço; e 7% em carnes, aqui foram relatados casos de substituição de carne bovina por soja e de carne suína por frango em linguiças (TÍBOLA *et al.*, 2008).

Adulterações alimentares podem ser detectadas por análises físico-químicas, como os trabalhos realizados por Ribeiro *et al.* (2012) quando analisaram adulteração em queijo parmesão ralado avaliando acidez, proteína, lipídios, umidade, cinzas, índice de peróxidos e teste com lugol e realizando análises microbiológicas para coliformes em queijo ralado tipo parmesão vendidos na cidade de Ponta Grossa/PR; Lima *et al.* (2021) por meio de métodos físico-químicos estudaram a detecção da fosfatase alcalina em leite

bovino, bubalino e caprino; e por Caldeira *et al.* (2010) em análise de leite onde foram avaliados acidez titulável, densidade, gordura, proteína, cinzas e índice crioscópico.

Técnicas cromatográficas (AUED-PIMENTEL *et al.*, 2015), espectroscópicas (AUED-PIMENTEL *et al.*, 2008) e de ressonância magnética (RIBEIRO *et al.*, 2017) são comumente utilizadas para detectar adulteração em alimentos. Aued-Pimentel *et al.* (2008) utilizaram espectroscopia UV e cromatografia líquida de alta eficiência (HPLC) para detectar adulteração em azeite de oliva; Aued-Pimentel *et al.* (2015) através da cromatografia gasosa determinaram teores de ácidos graxos, vitamina E, vitamina A em amostras de óleo de coco; Ribeiro *et al.* (2017) utilizaram a ressonância magnética (RMN) para realizar análises de adulteração em café; Galvin-King *et al.*, (2020) analisaram páprica picante através da espectroscopia molecular; Yamada *et al.* (2021) realizaram um estudo microscópio de páprica com o objetivo de identificar fraudes e materiais estranhos; e Lima *et al.* (2021) utilizaram PCR multiplex como método para autenticação de leites bubalino e caprino.

3.3 Cor e sistemas de cores

Na Química, para compreender o conceito de cor, é preciso compreender o conceito de luz. A luz visível ao olho humano é uma onda eletromagnética pertencente a uma faixa do espectro eletromagnético (Figura 4) (MARTINS *et al.*, 2015). Essa luz visível possui comprimentos de onda que são percebidos pelos olhos humanos através das cores (MARTINS *et al.*, 2015).

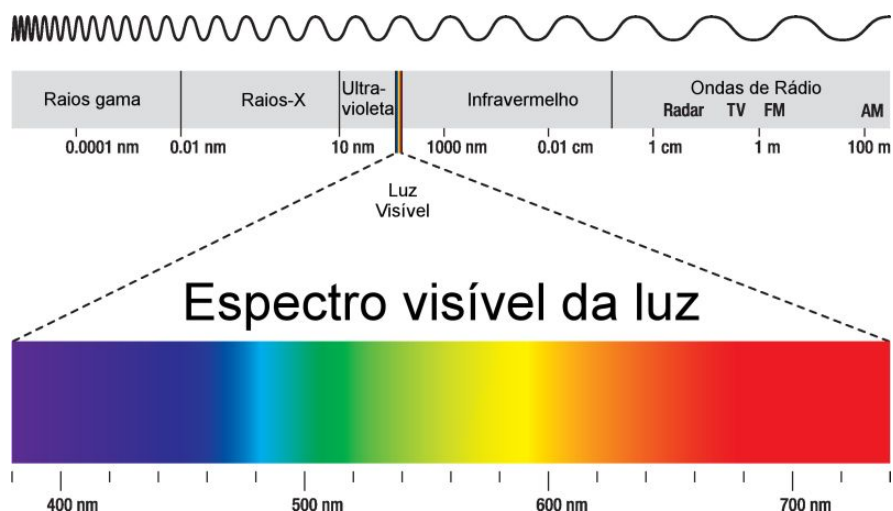


Figura 4: Espectro Eletromagnético (<https://www.infoescola.com/fisica/espectro-eletromagnético/>)

O espectro eletromagnético que corresponde às cores que os olhos humanos conseguem enxergar vai da cor violeta à vermelha (400 nm a 700 nm).

O sistema RGB (Red, Green e Blue) é o que regulamenta as cores dos objetos que emitem luz e o sistema CMY (Cyan, Magenta e Yellow) rege os corpos opacos (ROCHA, 2010). No sistema RGB, se as três cores forem somadas em proporções corretas, obtém-se a cor branca. Já no sistema CMY, ao somar as três cores surge à cor preta (ROCHA, 2010). A figura 5 ilustra a mistura das cores nestes dois sistemas.

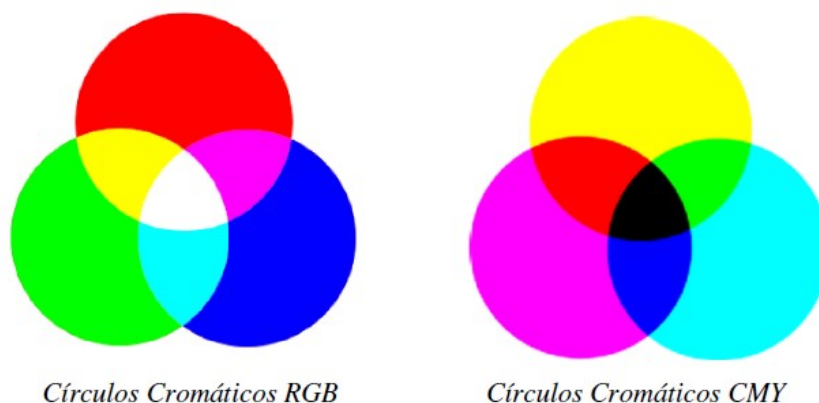


Figura 5: Círculos cromáticos dos sistemas de cores (Rocha, 2010).

No sistema de cores RGB a combinação das cores primárias (vermelho, verde e azul) resultam nas cores secundárias amarelo, magenta e ciano (Tabela 1).

Primária	λ (nm)	Combinação de Primárias	Secundária Resultante
Vermelho	700,0	Vermelho + Verde	Amarelo
Verde	546,1	Vermelho + Azul	Magenta
Azul	435,8	Verde + Azul	Ciano

Tabela 1: Cores primárias com respectivos comprimentos de onda. Combinação das cores primárias e cor secundária resultante (Autoria Própria).

O sistema RGB é utilizado em cinema, fotografia, televisão, fotografia digital, tela de computadores e a formação dessas imagens em terminais de vídeo ocorrem por emissão de radiação no espectro eletromagnético na faixa do UV-Vis. Já o CMY é o sistema utilizado em impressões coloridas em baixa escala (impressoras domésticas) e em artes plásticas (QUEIROZ *et al.*, 2006; ROCHA, 2010).

Na figura 6, há a decomposição da imagem digital cromática no sistema RGB. Na figura 6.a a imagem está decomposta no verde, 6.b no vermelho e 6.c no azul. A figura 6.d representa uma imagem digital bidimensional oriunda da composição das três cores do sistema RGB (QUEIROZ *et al.*, 2006).

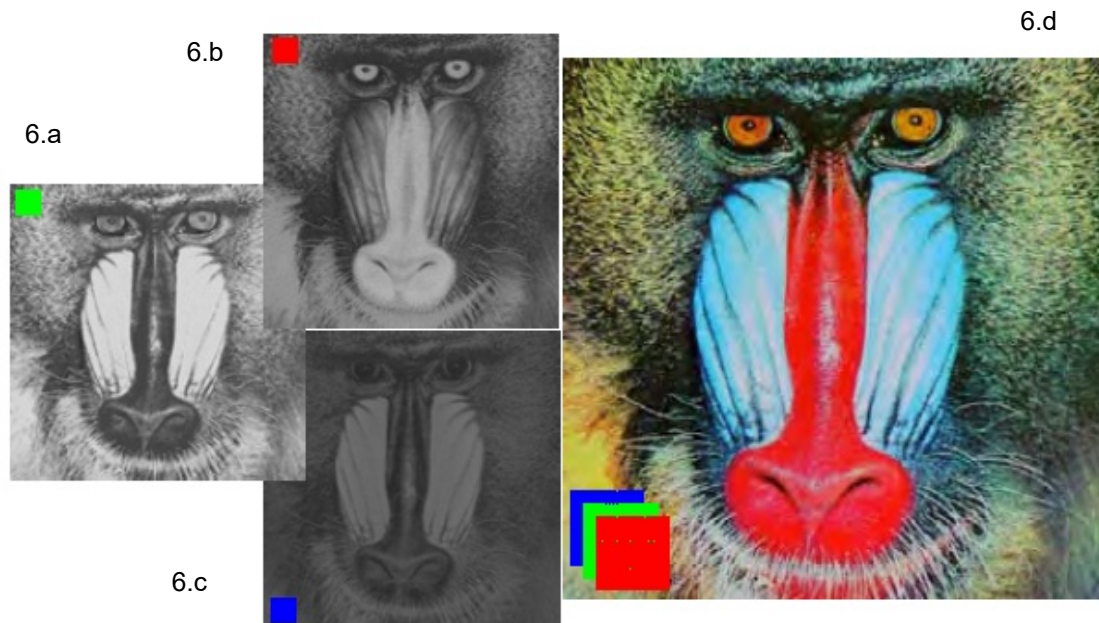


Figura 6: Representação de imagem digital bidimensional. 6.a a imagem está decomposta no verde, 6.b no vermelho, 6.c no azul e 6.d representa uma imagem digital bidimensional oriunda da composição das três cores do sistema RGB (QUEIROZ *et al.*, 2006).

Filho *et al.* (1999) abordaram que uma imagem monocromática é uma função matemática bidimensional, da intensidade da luz, do tipo $f(x,y)$, onde (x, y) são as coordenadas espaciais proporcionais ao brilho ou escala de cinza em determinado ponto (FILHO *et al.*, 1999). Cada ponto em uma grade bidimensional é chamado de *pixel*. O elemento de imagem ou *pixel* é a menor unidade de uma imagem digital (Figura 7). *Pixel* é uma abreviação do termo *picture element* (QUEIROZ *et al.*, 2006; LEÃO, 2005). Já em uma imagem digital cromática (colorida) no RGB, o *pixel* é um vetor de coordenadas que representam as intensidades das cores vermelha, verde e azul (QUEIROZ *et al.*, 2006). É a quantidade de *pixel* que determina a qualidade da imagem, assim, quanto maior for essa quantidade, melhor será a qualidade da imagem (LEÃO, 2005). A figura 8 representa essa relação qualidade da imagem x quantidade de *pixels*.



Figura 7: Imagem digital com ampliação para visualização dos *pixels* (LEÃO, 2005).

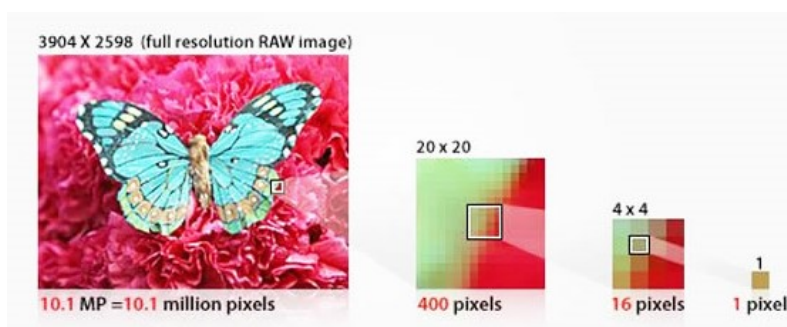


Figura 8: Qualidade da imagem x quantidade *pixels* (<https://fotografiamais.com.br/wp-content/uploads/2019/03/o-que-e-pixel-megapixel.jpg>)

3.4 Imagens Digitais e a Química Analítica

As imagens digitais (ID) estão ganhando cada vez mais espaço dentro da Química Analítica. Botelho *et. al.* (2017) desenvolveram um instrumento de baixo custo para determinação de corante vermelho em balas utilizando imagens digitais; Lima *et. al.* (2020) realizaram análise de vinhos tintos do Vale do São Francisco, por meio de histogramas de cores (obtidas através de ID); Pierini *et. al.* (2016) utilizaram as imagens digitais associadas à quimiometria para rastrear a origem de própolis argentina; e Reile *et. al.* (2020) analisaram, qualitativa e quantitativamente, amostras de catchup para detectar a presença de corante Sudan. São exemplos de estudos realizados utilizando imagens digitais na análise de alimentos. Contudo, a utilização desse recurso é bem mais ampla.

Bock *et al.* (2020) fizeram um levantamento da empregabilidade das imagens digitais em diversas áreas de análise. Controle de qualidade do meio ambiente (CHOODUM *et al.* 2013; ZARGHAMPOUR *et al.* 2020), análise forense (THANAKIATKRAI *et al.* 2013; DA SILVA *et al.* 2018), avaliação da qualidade de produtos farmacêuticos (LAMARCA *et al.* 2019; COUTINHO *et al.* 2019), controle de qualidade de combustível (RIBEIRO *et al.* 2019; JOÃO *et al.* 2019) são algumas possibilidades de uso desta.

O pioneirismo na utilização de imagens digitais para análises químicas é atribuído a Martinez e coautores (2008) quando captaram imagens digitais por meio de *scanner* e de uma câmera de *smartphone* para analisar glicose e proteína em amostras artificiais de urina. Esse trabalho utilizou o *scanner* e o *smartphone* apenas para digitalização ou captação das imagens, o tratamento de dados foi feito por meio de *software* de edição de imagens, o Adobe Photoshop®. Para o ensaio de glicose, os autores converteram as imagens digitais digitalizadas ou fotografadas em escala de cinza de oito *bits* e selecionaram a zona de teste apropriada do dispositivo. Já para o ensaio de proteína, as imagens foram convertidas em cor CMYK e medida a intensidade do ciano (MARTINEZ *et al.*, 2008).

Em 2011, têm início os primeiros estudos onde o *smartphone* foi utilizado para captação e para o tratamento de dados (BOCK *et al.*, 2020). Desde então as imagens digitais mostram-se como um recurso prático e eficiente em análises químicas, pois não demandam grandes equipamentos e podem ser realizadas fora do laboratório.

Há estudos que também utilizam o *smartphone*, mas apenas para captação das imagens, ficando o tratamento de dados realizados em *softwares* auxiliares. Thongprajukaew *et al.* (2014) determinaram amilase em intestino de peixes utilizando imagens digitais e tratamento de dados via Adobe Photoshop®. Franco *et al.* (2016) analisaram metanol em aguardente, as imagens foram captadas em *smartphone* e tratadas no ImageJ®. Ribeiro *et al.* (2019) estudaram a degradação oxidativa de óleo lubrificante via imagens digitais; contudo nesse estudo os autores associaram os dados obtidos pelas imagens digitais com análises quimiométricas, utilizando o Matlab®.

Delaney *et al.* (2011) realizaram a combinação de microfluídica de papel com detecção eletroquimioluminescente. As imagens foram analisadas usando

um programa escrito em Python[®], que separou os *pixels* em vermelho, verde e azul e produziu os valores totais para cada uma das intensidades. Ludwig *et al.* (2014) analisaram biomarcadores em extratos de leite usando um imunoenensaio de fluorescência de microesferas. Nesse estudo as captações e análises foram feitas no próprio *smartphone*, por meio do aplicativo chamado GotMilk[®]. Koh *et al.* (2016) avaliaram quantitativamente o volume total de perda de suor, o pH, bem como a concentração de cloreto, lactato e glicose no suor via análise colorimétrica e analisados via sistema integrado de comunicação de campo próximo (NFC). Marinho *et al.* (2019) também realizaram suas captações e análises no próprio *smartphone* quando quantificou etanol em bebidas destiladas utilizando o aplicativo *App Color Grab*[®].

No que tange trabalhos de análise de adulteração em microalgas, não foram encontrados até o momento no *Science Direct*, enfatizando ainda mais o caráter inovador deste estudo.

3.5 Análise Multivariada

Em 1970, a Quimiometria surge como uma ferramenta para solucionar alguns problemas dentro da Química (FERREIRA, 2022). É uma área dentro da química que utiliza de métodos matemáticos e estatísticos para planejar e otimizar experimentos químicos, além de auxiliar na obtenção do máximo de dados possíveis em uma análise química (FERREIRA *et al.*, 1999).

Ferramentas quimiométricas podem ser associadas aos métodos de análise para obter uma melhor avaliação da adulteração dos alimentos (GRIGOLO *et al.* 2018). Em alguns desses recursos tecnológicos as análises podem ocorrer de forma uni ou multivariada, o que as diferem são conjunto de variáveis analisadas. Enquanto na análise univariada apenas uma característica é analisada por vez, a multivariada analisa conjuntos similares ou dissimilares de um conjunto de características ao mesmo tempo (GRIGOLO *et al.*, 2018).

Análise por Componentes Principais (PCA) e Análise Hierárquica por Agrupamento (HCA) são as técnicas de análise multivariadas mais conhecidas na Quimiometria (PANERO *et al.*, 2009). Cebi *et al.* (2017) detectaram L-Cisteína em farinha de trigo usando microespectroscopia Raman combinada de

PCA e HCA. Garcia *et al.* (2019) avaliaram biomassa vegetal mais favorável para a produção de pellets de biocombustível via duas técnicas não supervisionadas, o PCA e o HCA. Panero *et al.* (2009) utilizaram essas técnicas de análise multivariada para realizar a discriminação geográfica do quiabo oriundo do Rio Grande do Norte e de Pernambuco.

A PCA e HCA também foram utilizadas nos estudos de Prado *et al.* (2022) com o objetivo de avaliar o estresse por cádmio, chumbo e alumínio em manjerição (*Ocimum basilicum L.*) utilizando ácido caféico, ácido rosmarínico, fenólicos totais, flavonóides totais e massa seca total como resposta. Neste trabalho, o manjerição foi submetido a diferentes concentrações dos metais. Os dados foram analisados via análise de variância, PCA e HCA e a partir deles conclui-se que o manjerição é sensível ao alumínio e pode tolerar de forma moderada concentrações de chumbo e cádmio quando cultivado de forma hidropônica. Além disso, a presença de chumbo e cádmio provocou aumento na síntese dos ácidos fenólicos, principalmente o ácido rosmarínico.

Outra metodologia para análise de adulterações em alimentos é via imagens digitais associada à Quimiometria, como o estudo de Araújo *et al.* (2017) onde foram analisadas amostras de açaí (*Euterpe oleracea*) comercializadas na cidade de Marabá (PA). Nesse estudo, as amostras de açaí obtidas em feiras livres foram adulteradas em laboratório com farinha de trigo e farinha de mandioca e posteriormente analisadas. Por meio de capturas de imagens digitais foi possível realizar análises por histograma, saturação. Além disso, foram realizadas análises por meio de PCA, por modelagem independente suave orientada por dados de analogia de classe (DD-SIMCA) e mínimos quadrados parciais de classe (OC-PLS). Os dois modelos de classe utilizados foram eficientes na classificação e distinção entre as amostras adulteradas e não adulteradas do açaí.

3.6 Inteligência Artificial

Ao longo da história da humanidade, a sociedade evoluiu. Concomitantemente a essa evolução ocorreu o avanço. O homem cada vez mais tenta copiar os mecanismos de raciocínio lógico advindos do cérebro humano para o funcionamento de máquinas (SILVA *et al.*, 2019). Já existem

linhas de estudos que afirmam que o momento atual trata-se da quarta revolução industrial (SILVA *et al.*, 2019; LUDERMIR, 2021). A característica principal dessa revolução é o avanço no desenvolvimento de máquinas que utilizam a tecnologia da Inteligência Artificial (IA) (LUDERMIR, 2021).

Schleder *et. al.* (2021) aborda que a IA pode ser descrita como técnicas computacionais que se assemelham à inteligência humana. Para Damaceno *et al.* (2018) a IA diz respeito à produção de máquinas que utilizam algoritmos complexos para atuar nas mais diversas áreas sociais e auxiliam na tomada de decisões a partir dos dados que foram disponibilizados, além de resolver problemas e criar soluções (SILVA *et al.*, 2019). A IA pode ser dividida em duas técnicas principais: o aprendizado de máquina (*Machine Learning*) e aprendizado profundo (*Deep Learning*) (SCHLEDER *et. al.*, 2021).

O *Machine Learning* possui algoritmos que trabalham com equações definidas previamente, que são capazes de organizar dados fornecidos na entrada e criar soluções para resolver uma situação-problema proposta na saída (DAMACENO *et al.*, 2018). Ainda segundo Damaceno *et al.*, (2018) um exemplo de aplicação do aprendizado de máquinas é a classificação de e-mails em Spam ou não. Para ocorrer tal separação, o software recebe e-mails classificados como Spam, com o objetivo de criar um padrão, e a partir daí consegue classificar os e-mails de usuários em Spam ou não.

O aprendizado de máquinas pode ainda ser subdividido em três tipos: supervisionado, não supervisionado e por reforço (BOCHIE *et al.*, 2020). No aprendizado supervisionado o algoritmo precisa receber um pacote de informações sobre o que ele precisa distinguir na entrada e saída dos dados informados. Nesse tipo de aprendizado o algoritmo deve gerar rótulos e formar uma classificação para que quando surgirem novos exemplos ele seja capaz de separar os dados corretamente (BOCHIE *et al.*, 2020; LUDERMIR, 2021). No aprendizado não supervisionado não há o pacote de informações fornecido previamente, os chamados rótulos. Neste caso o algoritmo separa os dados fornecidos por características em comum destes e classifica internamente de acordo a suas semelhanças (BOCHIE *et al.*, 2020; LUDERMIR, 2021). Já a aprendizagem por reforço trabalha com sistema de punições e recompensas. Não existe um levantamento de dados na entrada e saída, mas uma alimentação de dados constante entre o sistema e o ambiente (BOCHIE *et al.*,

2020). Segundo Ludermir (2021) esse tipo de aprendizado é o mais comumente utilizado em jogos e em robôs.

O Aprendizado Profundo (*Deep Learning*) é um tipo de técnica oriunda do Aprendizado de Máquinas (*Machine Learning*), porém é mais complexa. Nela a máquina tem a capacidade de reconhecer fala e identificar imagens (DAMACENO *et al.*, 2018). A vantagem do aprendizado profundo em detrimento ao aprendizado de máquina é a necessidade mínima de pré-processamento ou segmentação dos dados fornecidos (SANTOS *et al.*, 2019). Um exemplo de uso do *Deep Learning* é no reconhecimento de fala, nele a máquina realiza análises de padrões na voz do usuário, como intensidade, frequência e volume, assim quando esse usuário falar, a máquina irá identificar sua voz (DAMASCENO *et al.*, 2018).

A diferença básica entre as técnicas é que o aprendizado de máquina utiliza técnicas para análise automatizada de dados, como agrupamento, classificação e regressão, já o aprendizado profundo diz respeito a redes neurais que aprendem um conjunto de dados (SCHLEDER *et al.*, 2021).

O *Deep Learning* é comumente chamado de Redes Neurais Profundas, pois possuem estruturas que se assemelham aos neurônios humanos (BOCHIE *et al.*, 2020). Na figura 9 tem-se a estrutura básica de um neurônio artificial. Nela x são os vetores de entrada, cada vetor vai ter um peso w_k a ser multiplicado, esses elementos ponderados ($x.w_k$) são somados ao *bias*. O *bias* é um limiar de ativação e tem como função aumentar o grau de liberdade, fornecendo ao sistema maior adaptação ao conhecimento fornecido a ele. $\varphi(.)$ é a função de ativação que tem como objetivo mapear a saída y em um intervalo entre 0 e 1 (BOCHIE *et al.*, 2020; SOARES *et al.*, 2011).

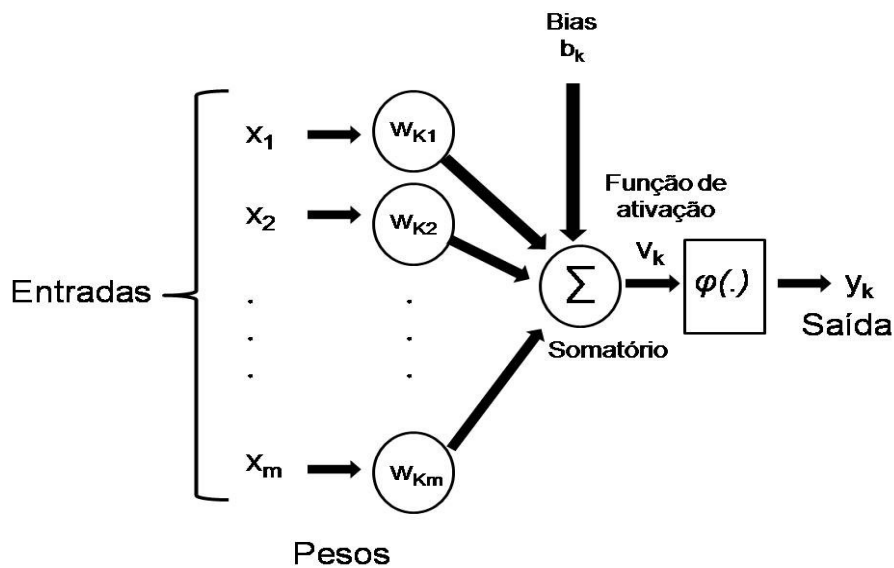


Figura 9: Estrutura básica neurônio artificial. O vetor de entrada é multiplicado por pesos. Os elementos ponderados e um limiar de ativação (*bias*) são somados e o potencial de ativação resultante alimenta uma função de ativação. A saída é o resultado da transformação do potencial de ativação. (Autoria Própria).

A organização de um neurônio consiste em unidades de processamento chamadas de camadas. Um conjunto de camadas é chamado de cadeia. A entrada e a saída são a primeira e a última camada de um neurônio e ainda existem as camadas internas que podem ser chamadas de camadas ocultas. As redes neurais profundas podem ter estrutura simples, com apenas entrada e saída (figura 10.a) e podem apresentar múltiplas camadas (Figura 10.b) (BOCHIE *et al.*, 2020).

10. a. Rede Neural Simples

10. b. Rede Neural Multicamadas

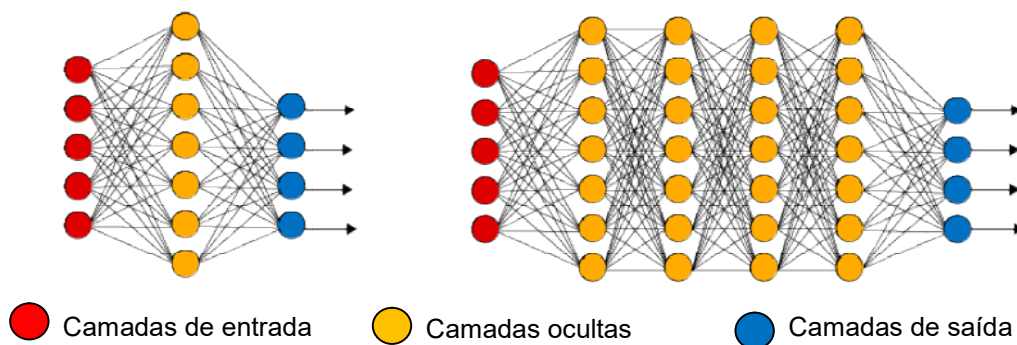


Figura 10: Tipos de rede neural profunda. 10.a rede neural simples. 10.b rede neural multicamadas (<https://www.deeplearningbook.com.br/o-que-sao-redes-neurais-artificiais-profundas/>).

As redes neurais profundas podem ser classificadas em Redes Neurais sem Realimentação, Recorrentes, Autoassociativas e Convolucionais (BOCHIE *et al.*, 2020). Segundo Bochie *et al.* (2020) nas redes neurais sem realimentação, as informações fornecidas na entrada seguem um fluxo e sentido único para saída. No entanto, esse tipo de rede é trabalhosa quando os dados são sequenciais. Desse modo, surge à rede neural recorrente, nela os dados inseridos podem ser reutilizados a qualquer momento, por se tratar de um fluxo cíclico de informações. A rede neural autoassociativa tem como função principal criar na saída uma informação mais próxima possível daquela que deu entrada. Já as redes neurais convolucionais são redes de aprendizado profundo que são utilizadas quando o problema possui muitas variáveis a serem analisadas (BOCHIE *et al.*, 2020). As redes neurais convolucionais (CNN) foram incorporadas a este estudo com o objetivo de dar ainda mais credibilidade e confiabilidade nos resultados obtidos.

As CNN são um tipo de rede neural artificial que processam dados e tem como uma de suas funções, o processamento de imagens (RAWAT *et al.*, 2017). Ela tem a capacidade de detectar e classificar objetos em uma imagem. Uma imagem é composta, conforme supracitado, por uma série de *pixels* dispostos em uma grade com valores que destacam a cor que cada *pixel* deve possuir (RAWAT *et al.*, 2017). Além disso, ela é capaz de capturar dependências espaciais e temporais em uma imagem no RGB e aplica filtros importantes durante a aquisição das características da imagem (RAWAT *et al.*, 2017). Segundo Santos *et al.* (2010), a CNN produz dados relevantes a partir de imagens captadas e pode ser usado para classificar, detectar e reconhecer imagens e vídeos.

Nas arquiteturas das redes neurais convolucionais o fluxo de informações ocorre em apenas uma direção, desde suas entradas até suas saídas. De uma forma geral, a arquitetura de uma CNN é formada por camadas convolucionais e de *pooling*, que são agrupadas em módulos, ou blocos, seguidos por uma ou mais camadas totalmente conectadas. Os blocos são frequentemente empilhados uns sobre os outros para formar um modelo profundo (LECUN *et al.*, 1998).

Cada CNN possui uma arquitetura básica com pelo menos quatro camadas: camada de entrada, camadas convolucionais, camadas de *pooling* e

camadas totalmente conectadas (HE *et al.*, 2018; RAWAT *et al.*, 2017). A figura 11 traz uma representação de uma estrutura básica de uma rede neural convolucional. Nela, a camada de entrada é aquela que pega a imagem, redimensiona para passar para camadas adicionais a fim de extrair as características principais; a camada convolucional é aquela que atua como um filtro para imagens encontrando um mapa de características que resume a presença de características detectadas na entrada; camada de *pooling* tem como objetivo reduzir a resolução espacial dos mapas de características. Esta etapa captura imagens grandes e as dimensiona preservando as informações mais importantes nelas dentro dessa ativação; já as camadas totalmente conectadas convertem os mapas de características 2D em um vetor de características 1D. Este vetor de características é a entrada para uma série de camadas totalmente conectadas, realizando o cálculo do produto interno (HE *et al.*, 2018; RAWAT *et al.*, 2017).

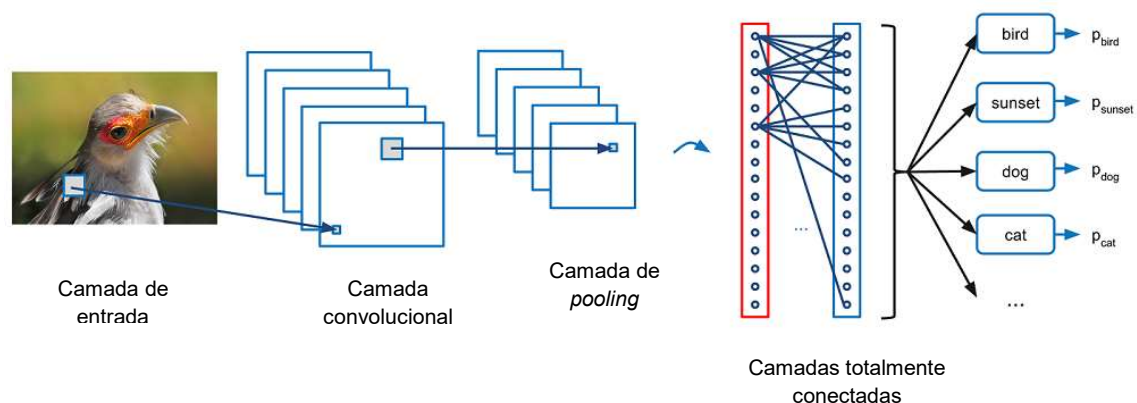


Figura 11: Arquitetura básica de uma rede neural convolucional (<https://www.deeplearningbook.com.br/as-10-principais-arquiteturas-de-redes-neurais/>).

A CNN está em ascensão e já é um recurso utilizado em diversas áreas. Um levantamento realizado na base de dados da *Science Direct* mostrou que, de 2013 a 2024, há uma crescente no número de publicações que utilizaram esse recurso da IA (Gráfico 1).

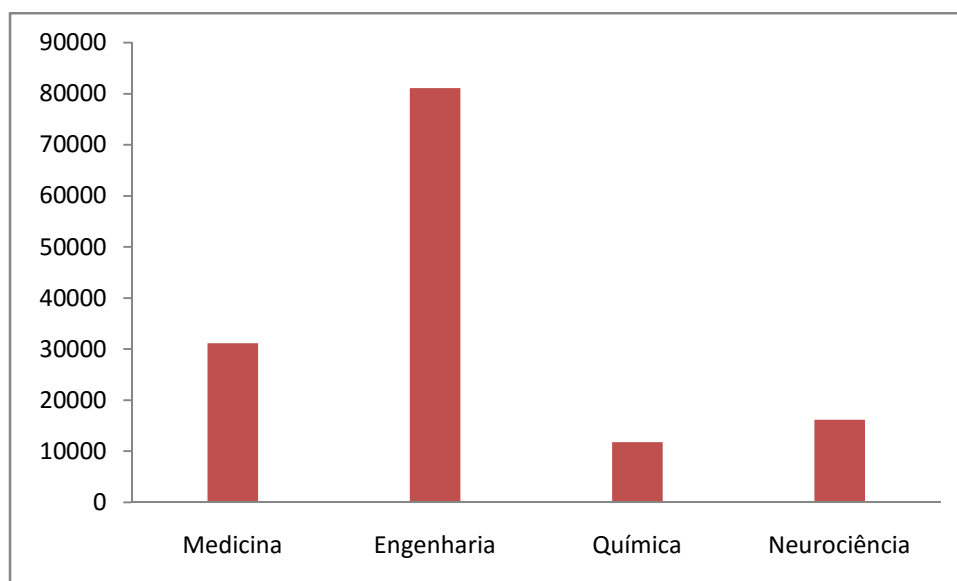


Gráfico1: Quantidade de publicações x áreas de publicações. Estudos utilizando as Redes Neurais Convolucionais. Levantamento feito no *Science Direct* em publicações de 2013 a 2024 (Autoria Própria).

São alguns exemplos de estudos utilizando a IA em diversas áreas. Na área da saúde, Vogado *et al.* (2019) propuseram diagnosticar leucemia através da CNN. No campo da perícia criminal, Souza *et al.* (2020) realizaram uma análise comparativa de CNN no reconhecimento de cenas. Na área jurídica, Mota *et al.* (2020) classificaram páginas iniciais de petições utilizando as redes neurais convolucionais. Na agronomia, Neto *et al.* (2019) detectaram e contaram os frutos verdes a partir de imagens digitais de pés de laranja através da CNN.

Dentro da Química, a utilização da CNN é ampla e vasta. Luo *et al.* (2023) combinaram o modelo de redes neurais convolucionais (CNN) e a espectroscopia vibracional representada pela espectroscopia Raman de superfície aprimorada (SERS) para identificar e analisar simultaneamente cada componente nos espectros de seis misturas de microplásticos comuns. Shanling *et al.* (2023) propuseram uma metodologia para avaliação da qualidade da borracha sintética em tempo real utilizando as CNN. Képeš *et al.* (2024) aplicaram diversas técnicas de interpretação de dados (conjunto de dados de calibração estendido coletados para o instrumento de espectroscopia de ruptura induzida por *laser* ChemCam) com o objetivo de compreender a tomada de decisão de redes neurais convolucionais quanto a classificação

desses dados. Liu *et al.* (2023) propuseram um novo método de modelagem de bateria de fluxo redox de vanádio usando otimização de texugo de mel assistido por CNN-BiLSTM. Esse tipo de rede neural aprende diretamente a relação comportamental entre corrente, vazão, estado de carga e tensão com excelente precisão, evitando o uso de parâmetros que sofram variação.

Esses são apenas alguns exemplos de estudos que utilizam as CNN na Química, desse modo, constata-se o quão atual, amplo e crescente é a junção dessas duas áreas de conhecimento.

4 MATERIAIS E MÉTODOS

4.1 Amostras

Para este estudo três amostras de *Spirulina* foram utilizadas. A primeira foi cedida pelo Laboratório de Cultivo de Microalgas e Plâncton da Universidade Federal do Recôncavo Baiano (UFRB), *campus* Cruz das Almas/BA; a segunda foi adquirida na loja de produtos naturais da cidade de Salvador/BA; e a terceira em loja de produtos naturais da cidade de Alagoinhas/BA. A Tabela 2 traz a origem, quantidade adquirida em cada localidade e data de aquisição das amostras.

Amostra	Origem	Quantidade	Data da Aquisição
A	URFB	60 gramas	18/02/2022
B	Salvador	60 gramas	13/02/2022
C	Alagoinhas	60 gramas	15/02/2022

Tabela 2: Descrição das amostras de *Spirulina* usadas. Origem, quantidade adquirida em cada localidade e data de aquisição das amostras (Autoria Própria).

A amostra de *Spirulina* adquirida na UFRB é cultivada em meio Zarrouk modificado (MARTINS *et al.*, 2014), conforme composição da Tabela 3.

Reagentes	Marca	g.L ⁻¹
Carbonato de sódio PA	Synth	16,8
Nitrato de sódio PA	Exodo	2,5
Fosfato de potássio dibásico anidro PA	Synth	0,5
Sulfato de potássio PA	Synth	1,0
Cloreto de sódio PA	Synth	1,0
Sulfato de magnésio PA	Synth	0,2
Cloreto de cálcio anidro PA	ACS	0,04
Sulfato de ferro II PA	ACS	0,01
EDTA sal dissolvido (2H ₂ O) PA	Exodo	0,08

Tabela 3: Meio Zarrouk modificado (MARTINS *et al.*, 2014).

Além do meio, existem condições ideais de cultivo de microalgas em laboratório. A temperatura entre 30 e 40° C, pH 8, iluminação e fonte de carbono para garantir a fotossíntese (FERREIRA, 2020). Essa amostra foi armazenada em vidro âmbar e conservada em geladeira. As outras duas amostras foram compradas em casas de produtos naturais, conforme supracitado.

As amostras foram divididas em porções e adulteradas com bicarbonato de sódio PA, da marca Synth, em percentuais de 0%, 10%, 15% e 25% m/m. A figura 12 detalha como essa divisão foi realizada.

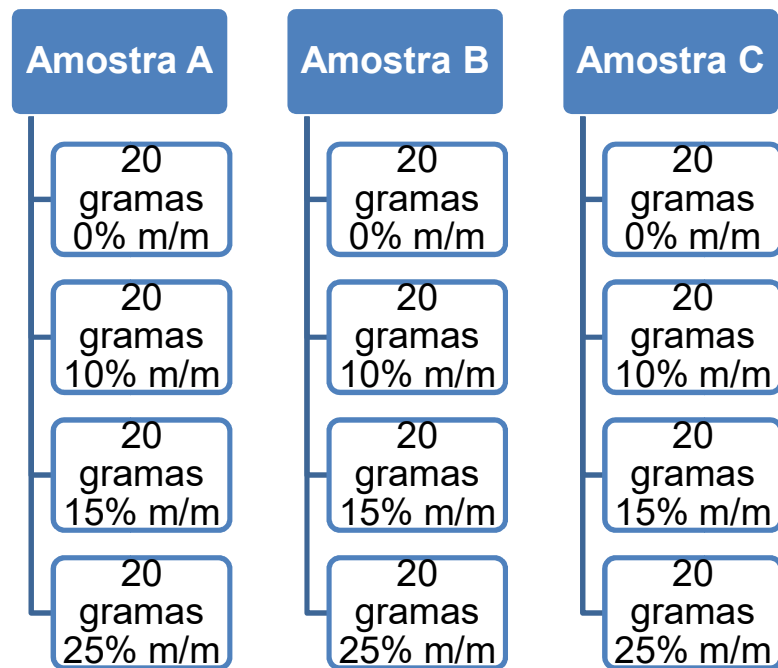


Figura 12: Divisão em porções das amostras de Spirulina com os percentuais de adulteração (Autoria Própria).

De posse das amostras adulteradas em seus respectivos percentuais, 02 (dois) gramas de cada fração foram pesadas e armazenadas em placa de polietileno, de cor transparente, medindo 5 x 5 x 1 cm para que as captações fossem realizadas (Figura 13).



Figura 13: Placa de polietileno com tampa, de cor transparente, medindo 5 x 5 x 1 cm contendo amostra de *Spirulina* (Autoria Própria).

Desse modo, um total de 120 placas foram preenchidas com 02 (dois) gramas de *Spirulina* adulterados em 0%, 10%, 15% e 25% m/m. Para facilitar a visualização dos dados e dos gráficos, abreviações e cores foram atribuídas para cada tipo de amostras. A tabela 4 traz a quantificação, peso, abreviações e cores de cada amostra.

Amostra	Quantidade Final	Peso	Abreviação	Cor
0%	30 placas	02 gramas	SP	Verde
10% m/m	30 placas	02 gramas	SD	Roxa
15% m/m	30 placas	02 gramas	SQ	Rosa
25% m/m	30 placas	02 gramas	SV	Vermelho

Tabela 4: Quantificação, peso, abreviação e cores atribuídas as amostras após a divisão (Autoria Própria).

4.2 Sistema de captação

Para a captação das imagens digitais foi utilizado um sistema composto por uma caixa de papelão pintada com tinta guache fosca branca e revestida com isopor, com dimensões de 36 x 22 x 13 cm (Figura 14. a). A luz interna utilizada foi de uma mini luminária portátil contendo três LEDs brancos, com funcionamento à pilha do tipo AAA, da marca LLUM com potência de 0,5W e dimensões 7 x 7 x 2 cm (Figura 14.b). A mesma foi posicionada na parte interna da caixa, do lado esquerdo a 20 cm do local onde as amostras foram colocadas para captação. Na parte externa superior da caixa foi acoplada uma capa de celular para o modelo de *smartphone* com o objetivo de deixar o local das captações fixo (Figura 14.c).

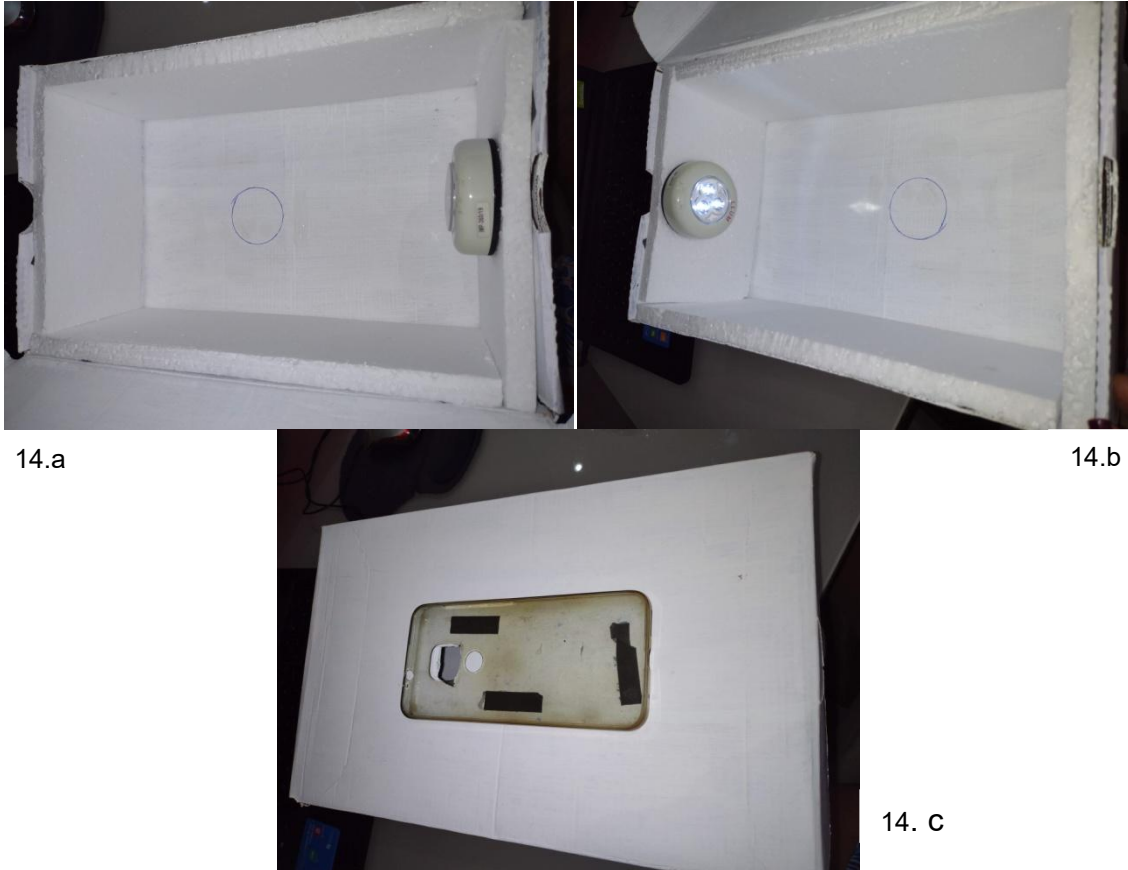


Figura 14: Caixa para captação das amostras. 14.a caixa parte interna, medindo 36x22x13 cm. 14.b luminária localizada a 20 cm centro da captação das imagens. 14.c caixa parte externa (Autoria Própria).

O *smartphone* utilizado na captação das imagens foi do modelo Moto g(9) *play* da marca Motorola®. O mesmo possui memória RAM de 4 GB, memória ROM de 64 GB, processador 4x 2.0 GHz Kryo 260 Gold + 4x 1.8 GHz Kryo 260 Silver, tela com resolução de 1600 x 720 *pixels*, câmera traseira tripla de 48 mega *pixels* + 2 mega *pixels* + 2 mega *pixels*, a câmera de 48 mega *pixels* permite captar fotos com 8000 x 6000 *pixels* de resolução, conforme informações contidas nas configurações do aparelho.

Algumas variáveis, aqui chamados de testes de otimização, foram analisadas para que o sistema de captação de imagens fosse o mais eficaz possível.

4.3 Avaliação das variáveis do sistema de captação e tratamento das imagens digitais

Diversos itens foram analisados até que se chegasse a um sistema próximo da idealidade: área de captação, altura do porta-amostra, disposição da amostra e método de aglomeração de dados sofreram variação e foram analisadas de forma univariada, conforme tabela 5. Para realização desta etapa apenas duas concentrações das amostras, 0% e 25% m/m, foram utilizadas, pois neste momento o objetivo não era a análise das adulterações, mas sim avaliação das variáveis.

Teste	Item Avaliado
1	Área de captação das fotos de 32 x 32
2	Área de captação das fotos de 64 x 64
3	Altura do porta-amostra de 1,3 cm do coletor + 1,6 cm de isopor
4	Altura do porta-amostra de 1,3 cm do coletor + 3,2 cm de isopor
5	Disposição da amostra: com prensa
6	Disposição da amostra: sem prensa
7	PCA: Auto Histograma R, G, B HCA: Average Histograma RGB Complete Histograma RGB Single Histograma RGB
8	PCA: Mean Histograma R, G, B HCA: Average Histograma RGB Complete Histograma RGB Single Histograma RGB
9	PCA: Auto Single Channels R, G, B HCA: Average Single Channels RGB Complete Single Channels RGB Single Single Channels RGB
10	PCA: Mean Single Channels R, G, B HCA: Average Single Channels RGB Complete Single Channels RGB Single Single Channels RGB

Tabela 5: Descrição dos testes de otimização do sistema de captação (Autoria Própria).

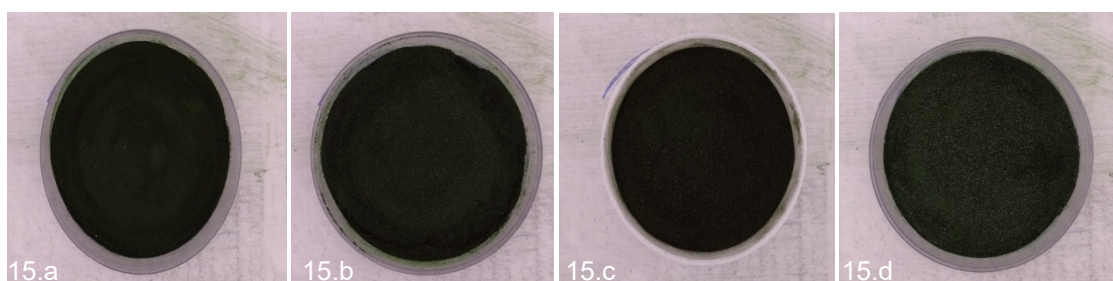
Ao finalizar os testes de otimização, definiu-se um sistema de captação adequado para este estudo, bem como o método de tratamento das imagens captadas. Usando o software *ChemoStat*[®], versão para *Windows* V2 build 260821.ba3285 (Helfer, 2015), as imagens captadas foram tratadas e utilizou-se duas técnicas de análises exploratórias para plotagem dos gráficos de PCA e HCA.

Outro recurso utilizado neste trabalho foi a Inteligência Artificial, através das Redes Neurais Convolucionais (CNN), via MatLab[®] Neural Network Toolbox (Rawat et al., 2017) versão para *Windows* matlab2020b. Nele, a camada de entrada da rede neural convolucional foi composta por 360 imagens das amostras, divididas em categorias: 90 imagens das amostras 0%, 90 imagens das amostras adulteradas em 10% m/m de bicarbonato de sódio, 90 imagens das amostras com 15% m/m de adulteração do mesmo contaminante e 90 imagens das amostras contaminadas com 25% m/m de NaHCO₃.

5 RESULTADOS E DISCUSSÕES

5.1 Amostras

As adulterações da Spirulina foram realizadas de forma controlada com bicarbonato de sódio (NaHCO₃), com percentuais de 0%, 10%, 15% e 25% m/m. O adulterante foi escolhido por não ser higroscópico, facilitando seu manuseio; e de baixo custo. A figura 15 traz fotos das amostras com e sem tratamento no *ChemoStat*.



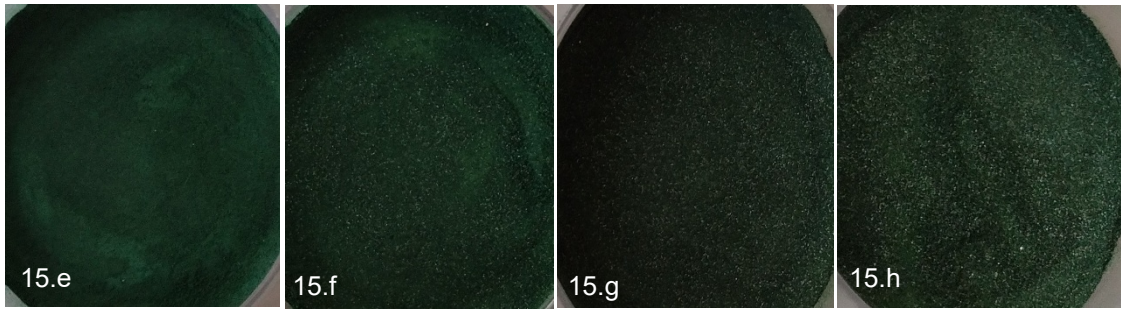
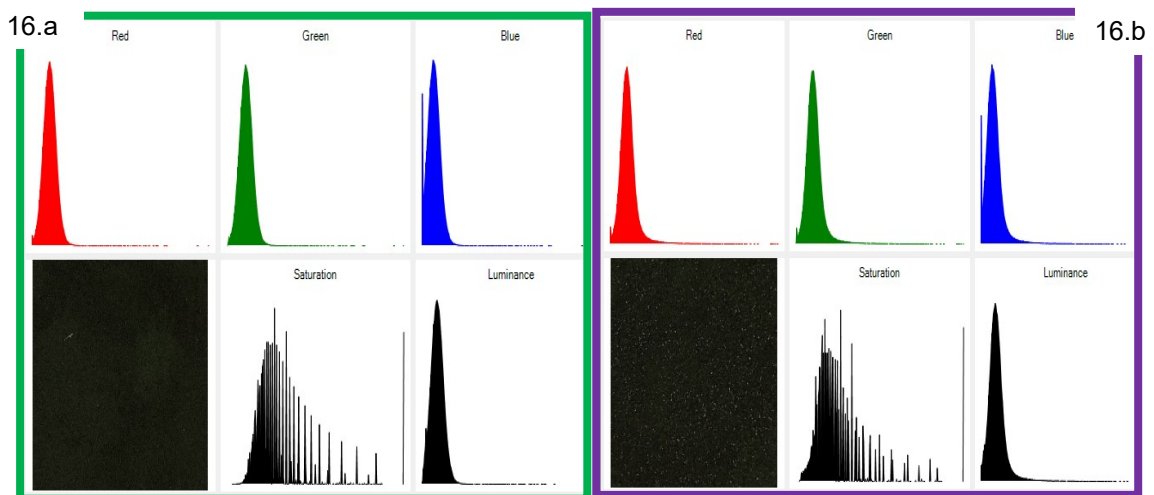


Figura 15: Amostras de *Spirulina* com e sem tratamento no *ChemoStat*. 15.a amostra 0% sem tratamento, 15.b amostra adulterada em 10% m/m sem tratamento, 15.c amostra adulterada em 15% m/m sem tratamento, 15.d amostra adulterada em 25% m/m sem tratamento, 15.e foto da amostra 0% tratada no *ChemoStat* (500x500 pixels e ROI 150X50), 15.f foto da amostra adulterada em 10% m/m tratada no *ChemoStat* (500x500 pixels e ROI 150X50), 15.g foto da amostra adulterada em 15% m/m tratada no *ChemoStat* (500x500 pixels e ROI 150X50) e 15.h foto da amostra adulterada em 25% m/m tratada no *ChemoStat* (500x500 pixels e ROI 150X50), (Autoria Própria).

Observou-se que, visualmente, as amostras não apresentaram uma diferença tão aparente, a ponto de diferenciar as amostras 0% das demais, adulteradas. Assim, o histograma das amostras (Figura 16) foi realizado para reinterar as diferenças existentes nos tipos de amostras. O histograma é um gráfico que mostra a distribuição do nível de brilho de cada cor no RGB.



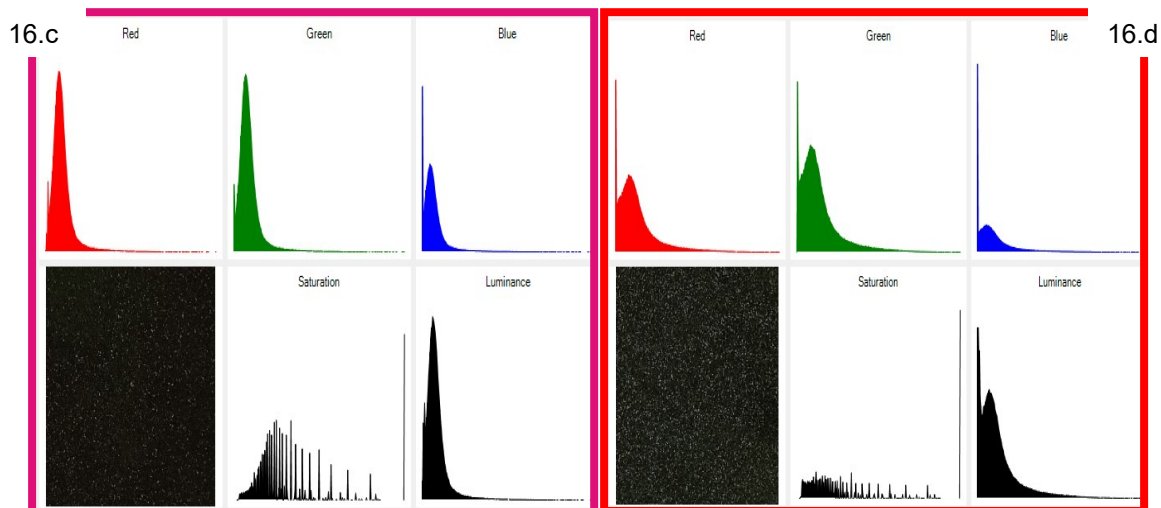


Figura 16: Histograma de cores das amostras. 16.a: amostra 0%. 16.b: amostra adulterada em 10% m/m. 16.c: amostra adulterada em 15% m/m e 16.d: amostra adulterada em 25% m/m (Autoria Própria).

Analisando os histogramas de cores das amostras, verificou-se que existe uma diferença nos histogramas das amostras 0% e adulteradas com 15% e 25%, o que não ocorreu com as amostras adulteradas em 10%. Isso porque as amostras com maiores percentuais de adulteração estão mais diluídas, fazendo com que as cores do histograma diminuam.

5.2 Avaliação das variáveis do sistema de captação

Conforme supracitado, variáveis foram analisadas com o objetivo de obter um sistema de captação mais próximo do ideal. Área de captação das imagens de 32 x 32 e 64 x 64; altura do porta-amostra de 1,3 cm da placa + 1,6 cm de isopor e de 1,3 cm da placa + 3,2 cm de isopor; disposição da amostra e método de aglomeração dos dados foram avaliados.

5.2.1 Área de captação

A área de captação das imagens foram de 32 x 32 no teste 1 e 64 x 64 no teste 2. Os gráficos 2 e 3 trazem o PCA e o HCA do teste 1 e os gráficos 4 e 5 do teste 2.

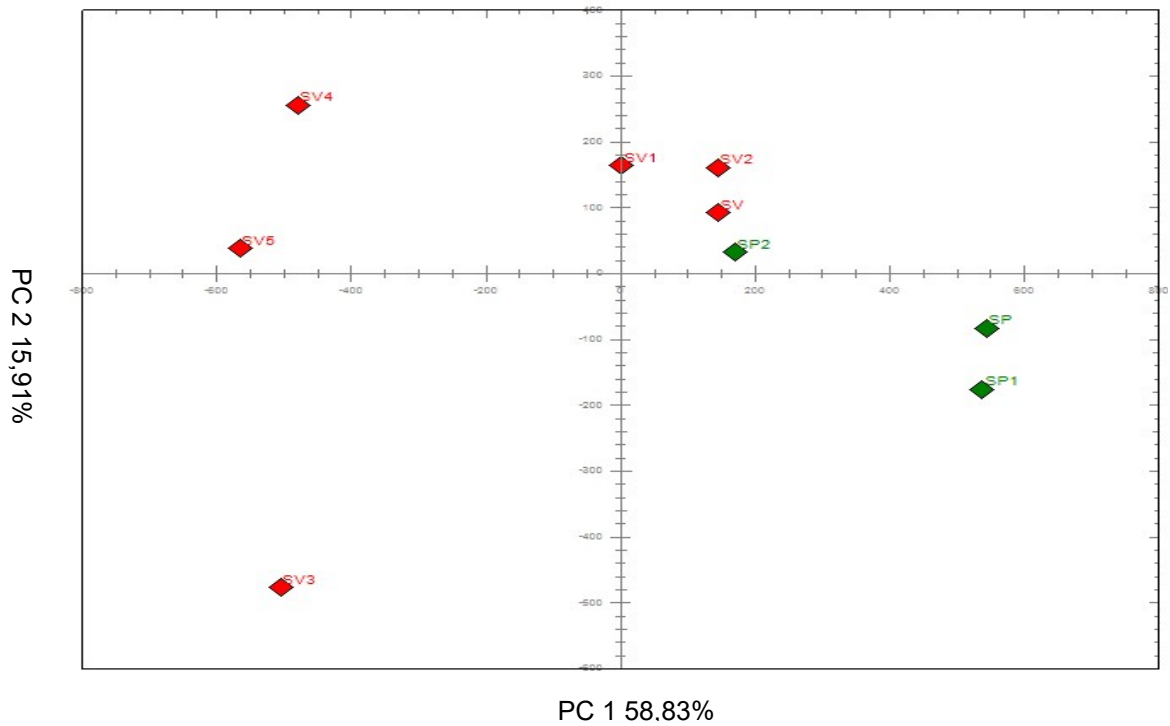


Gráfico 2: PCA Teste 1 – Área de captação 32 x 32 (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

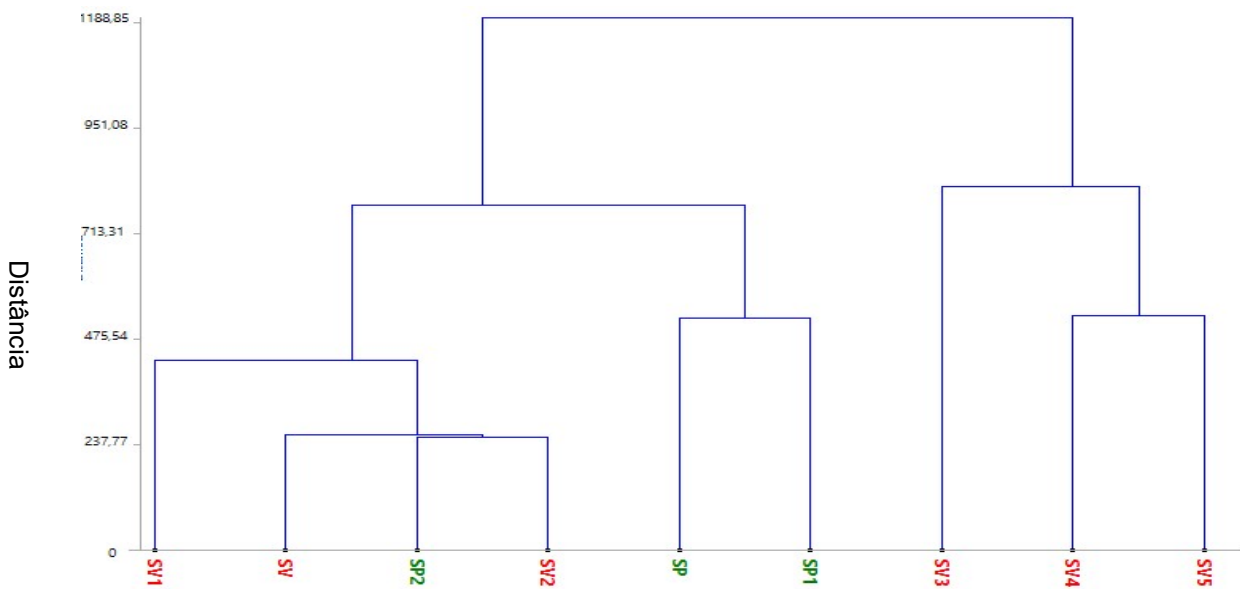


Gráfico 3: HCA Teste 1 – Área de captação 32 x 32 (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana (Autoria Própria).

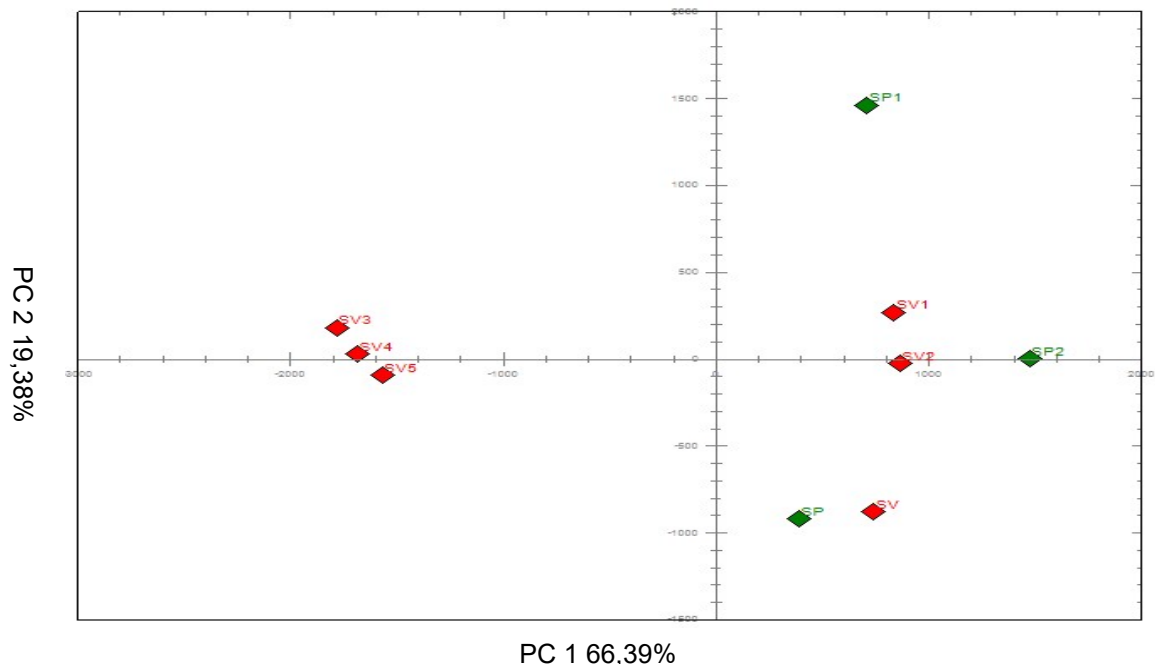


Gráfico 4: PCA Teste 2 – Área de captação 64 x 64 (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

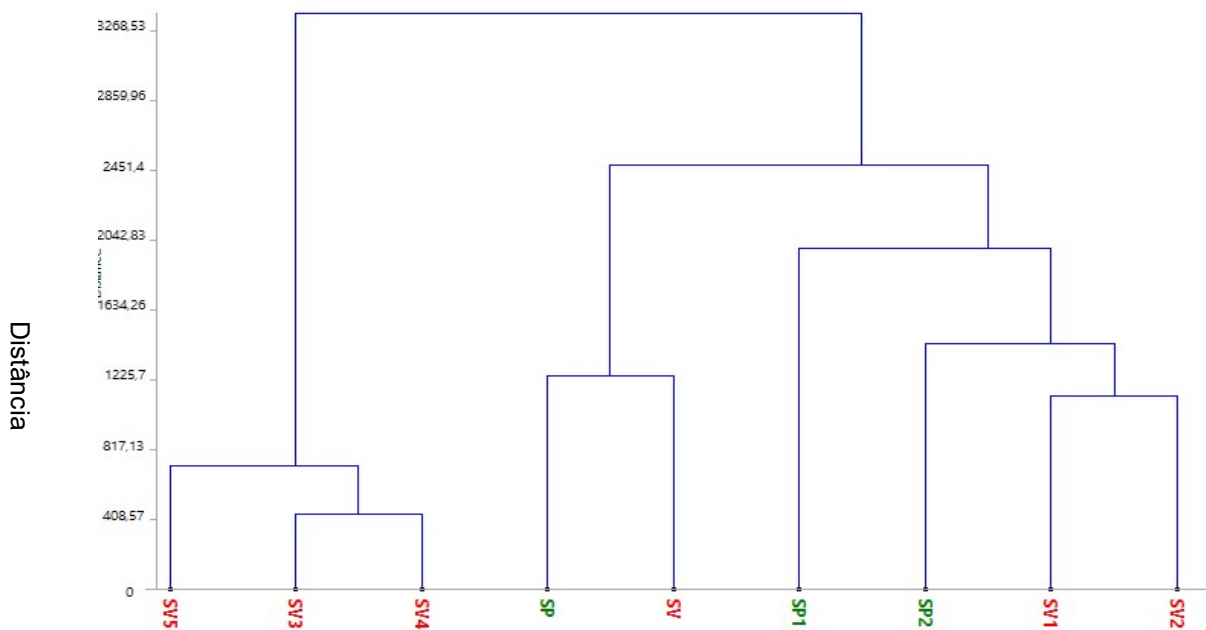


Gráfico 5: HCA Teste 2 – Área de captação 64 x 64 (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana (Autoria Própria).

Comparando-se o PCA do teste 1 e do teste 2, observou-se que no teste 2 as amostras ficaram menos dispersas, desse modo a região de interesse

para captação das fotos escolhida para este estudo foi a 64 x 64. Essas áreas de captações nada mais são do que as matrizes dos *pixels* capazes de serem captados nas amostras. Toda fotografia é composta por *pixel*, este é “o menor ponto visível em uma imagem” (MORENO *et. al.*, 2016). Dessa forma, a área de captação 64 X 64 é melhor do que 32 X 32, pois houve o aumento da matriz do *pixel* da imagem. Quanto maior o *pixel*, maior a qualidade da fotografia (MORENO *et. al.*, 2016).

5.2.2 Altura do porta-amostra

Dois testes foram realizados para analisar se a altura do porta-amostra interferiria na captação das imagens, desse modo os testes 3 e 4 avaliaram essa variável. No teste 3, a altura do porta-amostra foi 1,3 cm do coletor onde a amostra ficou disposta e acrescentou-se uma altura de 1,6 cm feita de isopor. Já no teste 4 dobrou-se a altura acrescida, a mesma foi de 1,6 cm para 3,2 cm também feita com isopor. Os gráficos 6 e 7 foram plotados com os dados do teste 3 e os gráficos 8 e 9, do teste 4.

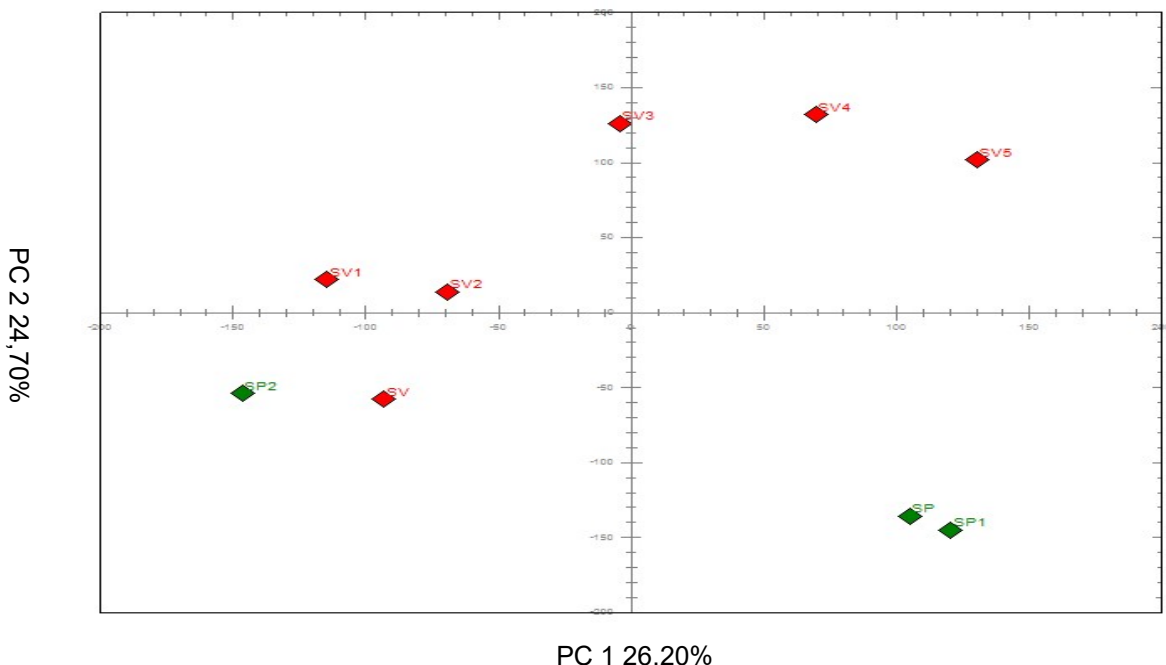


Gráfico 6: PCA Teste 3 – Altura do porta-amostra de 1,3 cm da placa + 1,6 cm de isopor (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

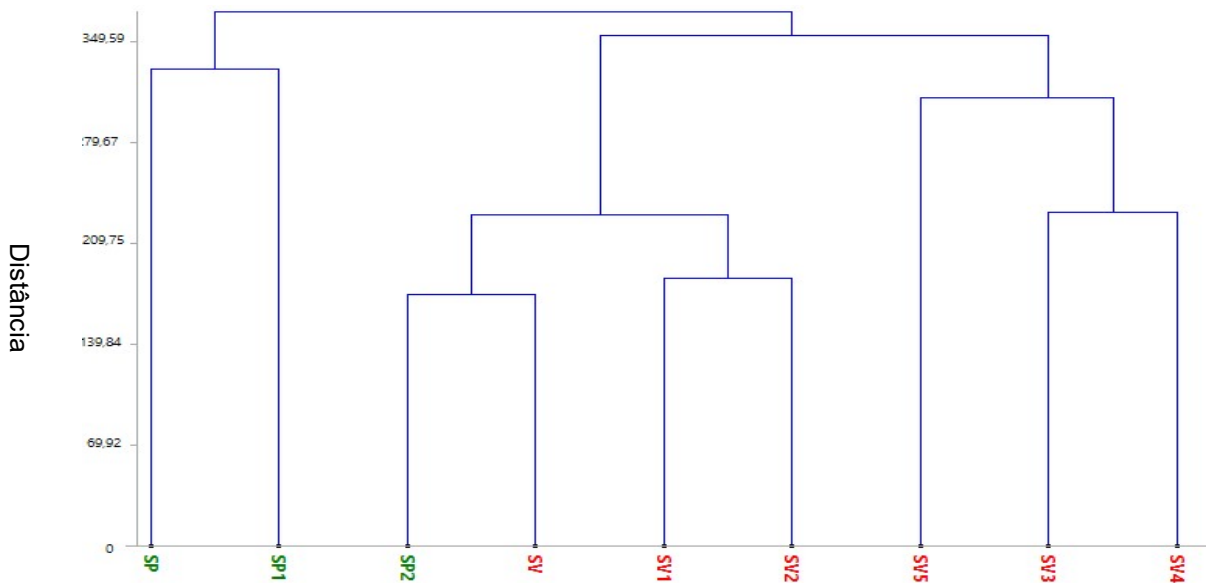


Gráfico 7: HCA Teste 3 – Altura do porta-amostra de 1,3 cm da placa + 1,6 cm de isopor (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m). Método Complete Linkage. Distância Euclidiana (Autoria Própria).

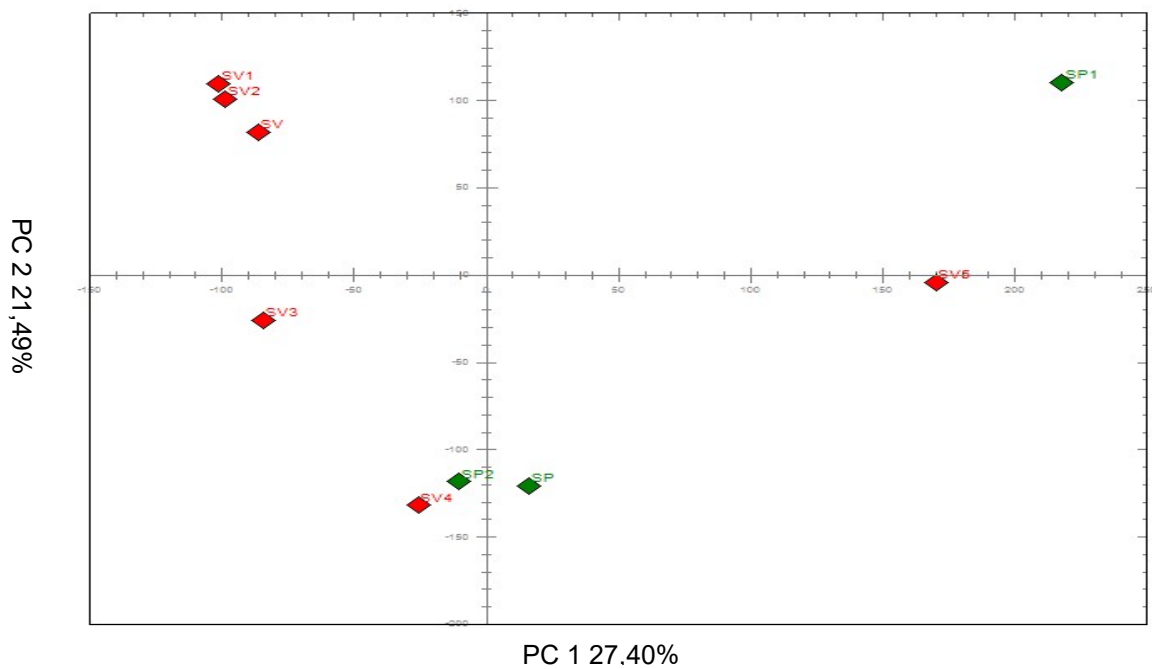


Gráfico 8: PCA Teste 4 - Altura do porta-amostra de 1,3 cm da placa + 3,2 cm de isopor (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

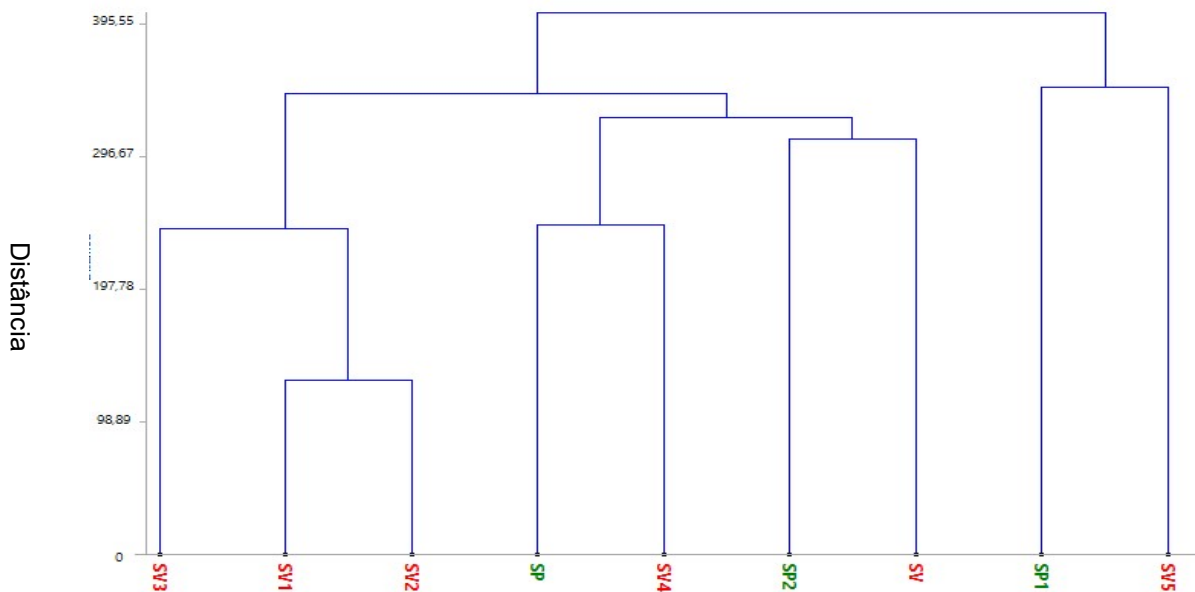


Gráfico 9: HCA Teste 4 – Altura do porta-amostra de 1,3 cm da placa + 3,2 cm de isopor isopor (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m). Método Complete Linkage. Distância Euclidiana (Autoria Própria).

Percebeu-se que em nenhum dos gráficos destes testes apresentaram boa separação, o que levou a escolha do porta-amostra sem quaisquer alturas, a não ser a própria – 1,3 cm. A distância entre a câmera e o objeto a ser fotografado influencia a qualidade da captação. O erro aumenta à medida que a distância entre a câmera e o objeto a ser fotografado diminuem (PAUL *et. al.*, 1993). Assim, apenas a altura do porta-amostra foi a mais adequada e a escolhida para a realização desse trabalho.

5.2.3 Disposição da amostra

Os testes 5 e 6 foram realizados para analisar a disposição da amostra. O teste 5 a amostra foi prensada mecanicamente até completa homogeneização da superfície. Já no teste 6, a amostra não foi prensada, apresentando em sua superfície rugosidade e não linearidade. Os gráficos 10 e 11 foram plotados com os dados da amostra prensada e os gráficos 12 e 13 da amostra sem prensa.

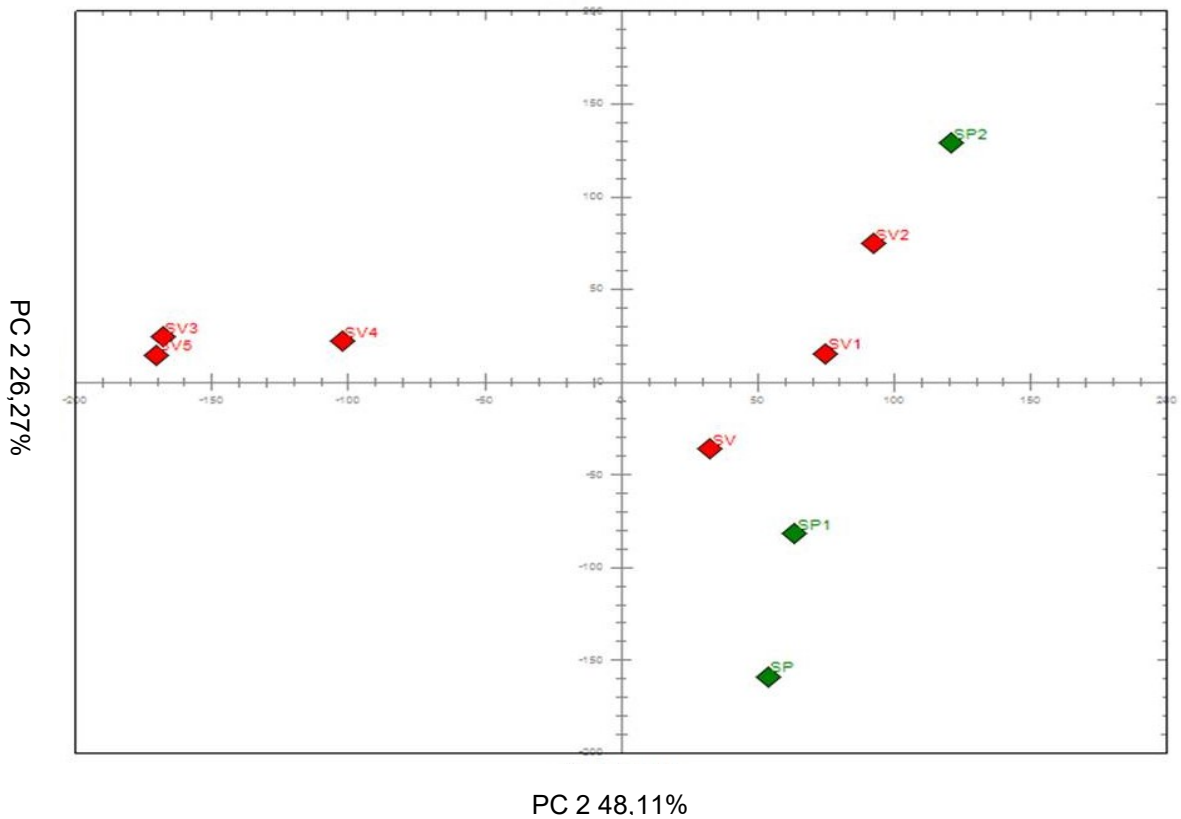


Gráfico 10: PCA Teste 5 – Amostra Prensada (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

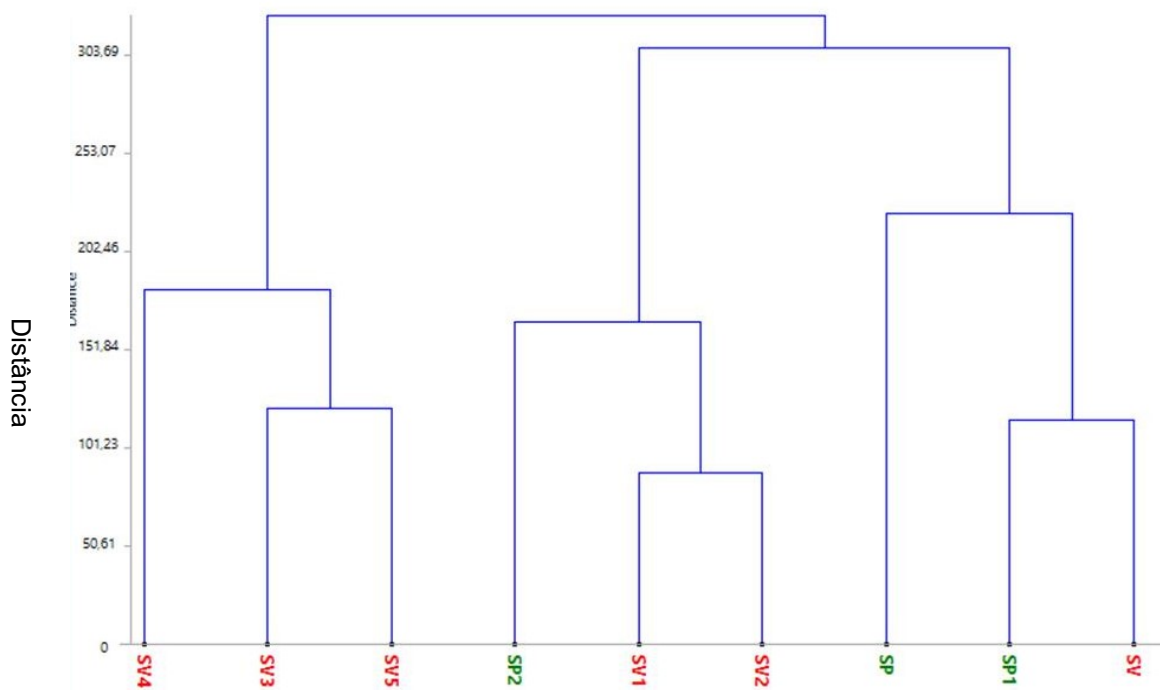


Gráfico 11: HCA Teste 5 – Amostra Prensada (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana (Autoria Própria).

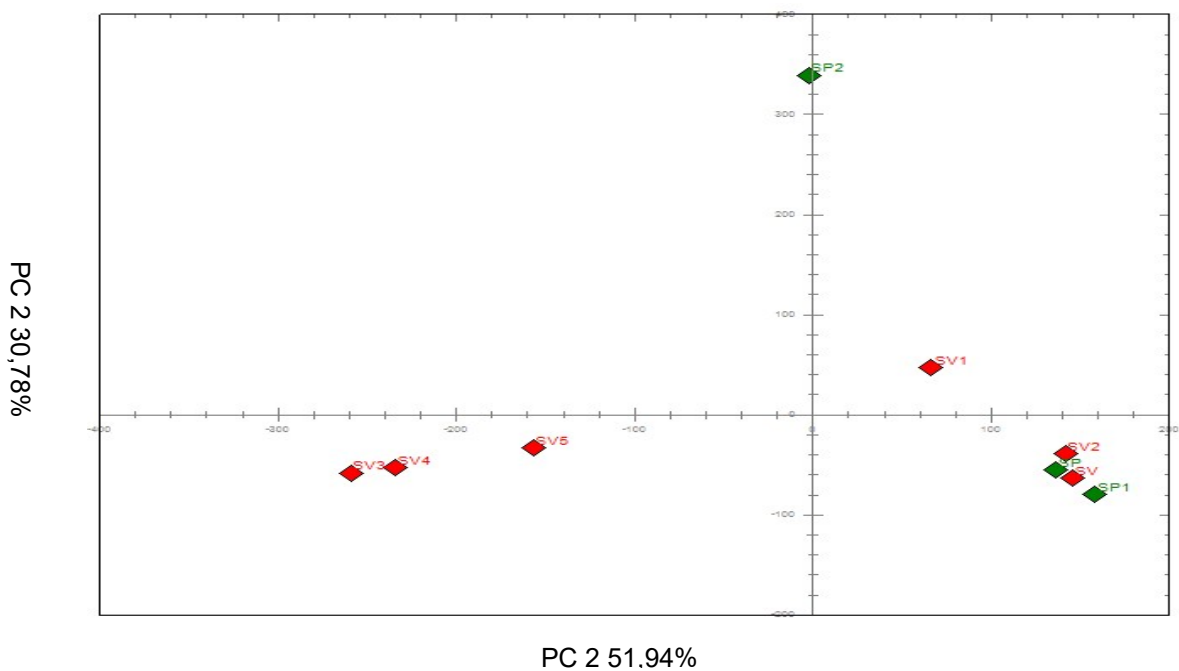


Gráfico 12: PCA Teste 6 – Amostra Sem Prensa (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

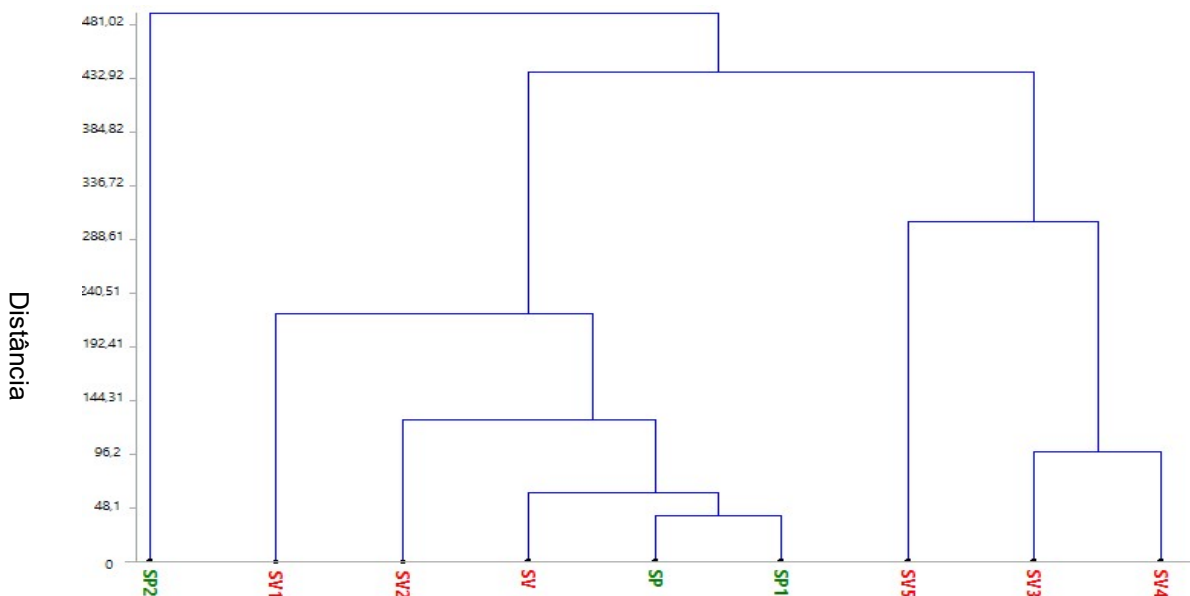


Gráfico 13: HCA Teste 6 – Amostra Sem Prensa (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana (Autoria Própria).

A partir das análises dos gráficos 10, 11, 12 e 13, escolheu-se prensar a amostras, pois as amostras sem prensa não foram classificadas tão bem. Além

disso, a amostra prensada evita rugosidade e deformidades na superfície desta, sendo a mais adequada.

5.2.4 Método de aglomeração de dados

Os métodos de aglomeração de dados no ChemoStat foram avaliados, alternando no PCA: Auto Histograma, Mean Histograma, Auto Single Channels e Mean Single Channels. A diferença entre Auto e Mean consiste em: na auto os dados autoescalados e no mean centralizado os dados na média.

No HCA foram avaliados: average, complete e single linkage. Average Linkage, Complete Linkage e Single Linkage são agrupamentos hierárquicos, em português, significam ligação média, ligação completa e ligação única, respectivamente. Na ligação média há o cálculo da distância média entre as classes antes de mesclá-las; a ligação completa calcula a distância máxima entre classes antes de mesclá-los; já na ligação simples, é a distância mínima que é calculada entre as classes antes de mesclá-las.

O teste 7 utilizou o Método Auto Histograma, nas seguintes configurações: PCA: Auto Histograma R, G, B e HCA: Average Histograma RGB; Complete Histograma RGB; e Single Histograma RGB.

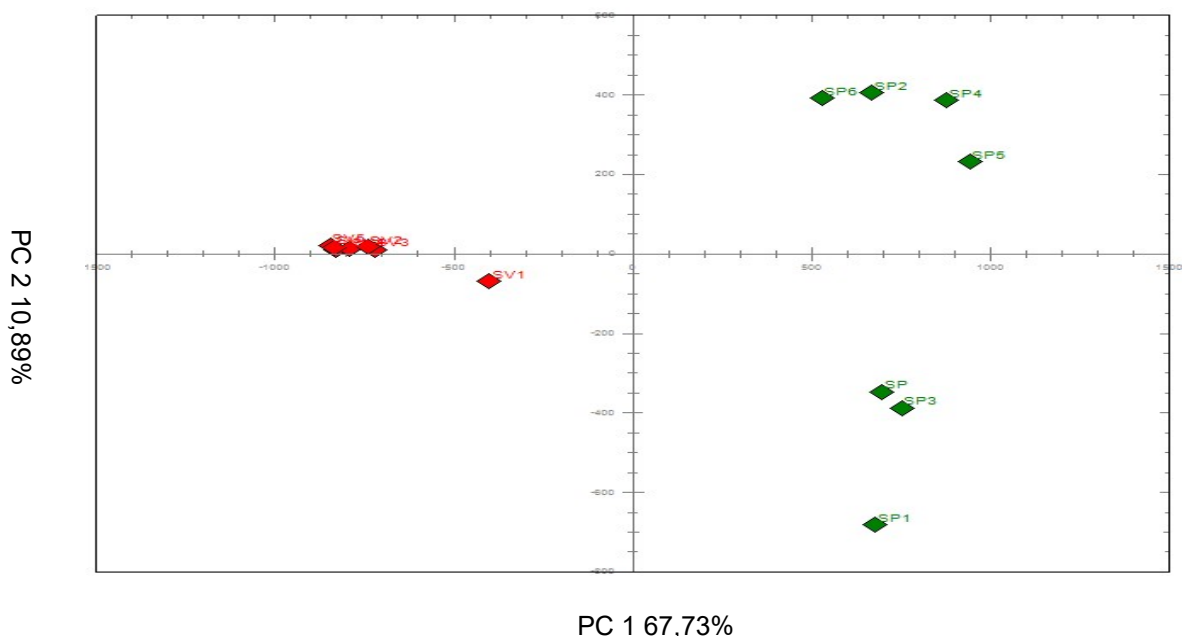


Gráfico 14: PCA Teste 7 - PCA: Auto Histograma R, G, (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

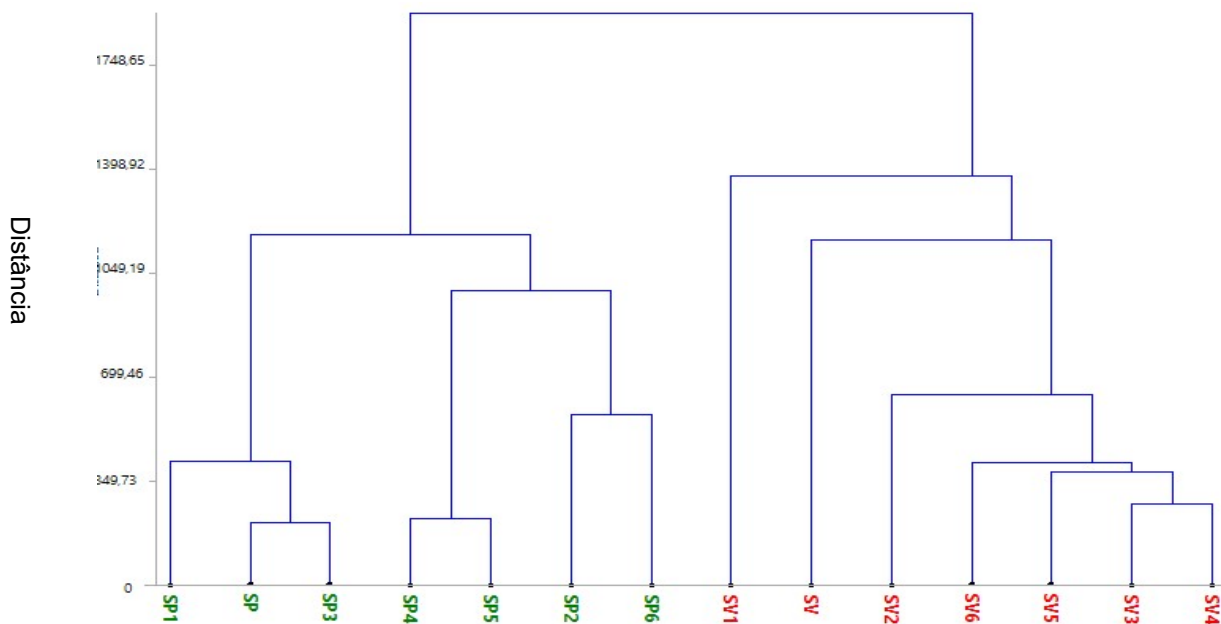


Gráfico 15: HCA Teste 7 - Complete Histograma RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana (Autoria Própria).

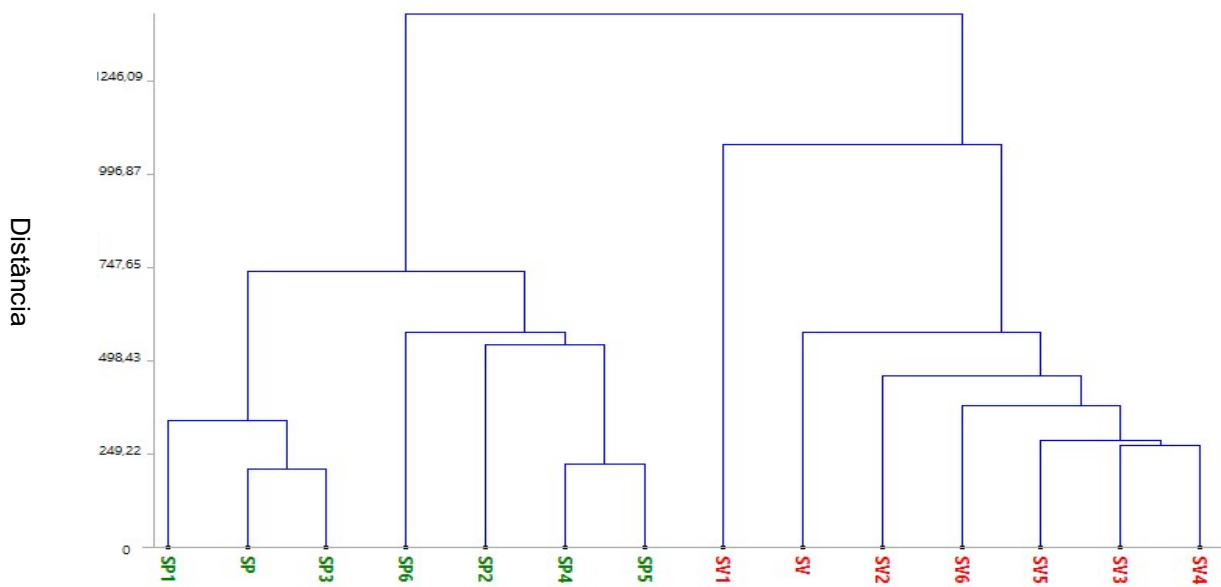


Gráfico 16: HCA Teste 7 - Single Histograma RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Single Linkage. Distância Euclidiana (Autoria Própria).

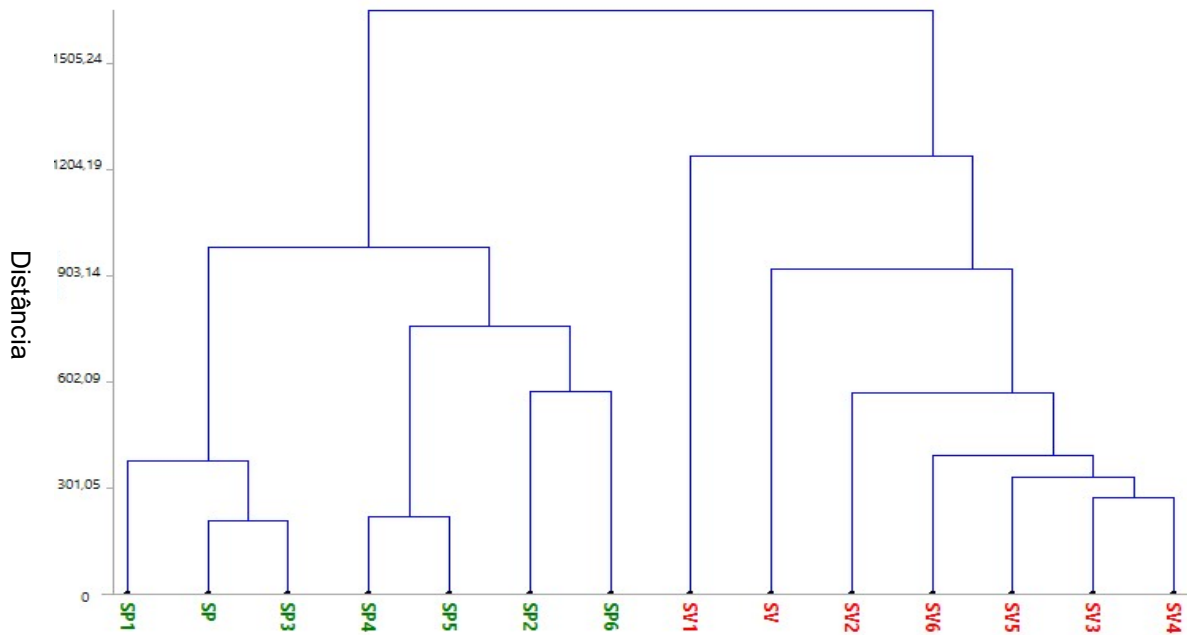


Gráfico 17: HCA Teste 7 - Average Histograma RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Average Linkage. Distância Euclidiana (Autoria Própria).

No oitavo teste, o Método Mean Histograma foi utilizado. PCA: Mean Histograma R, G, B e HCA: Average Histograma RGB; Complete Histograma RGB; e Single Histograma RGB.

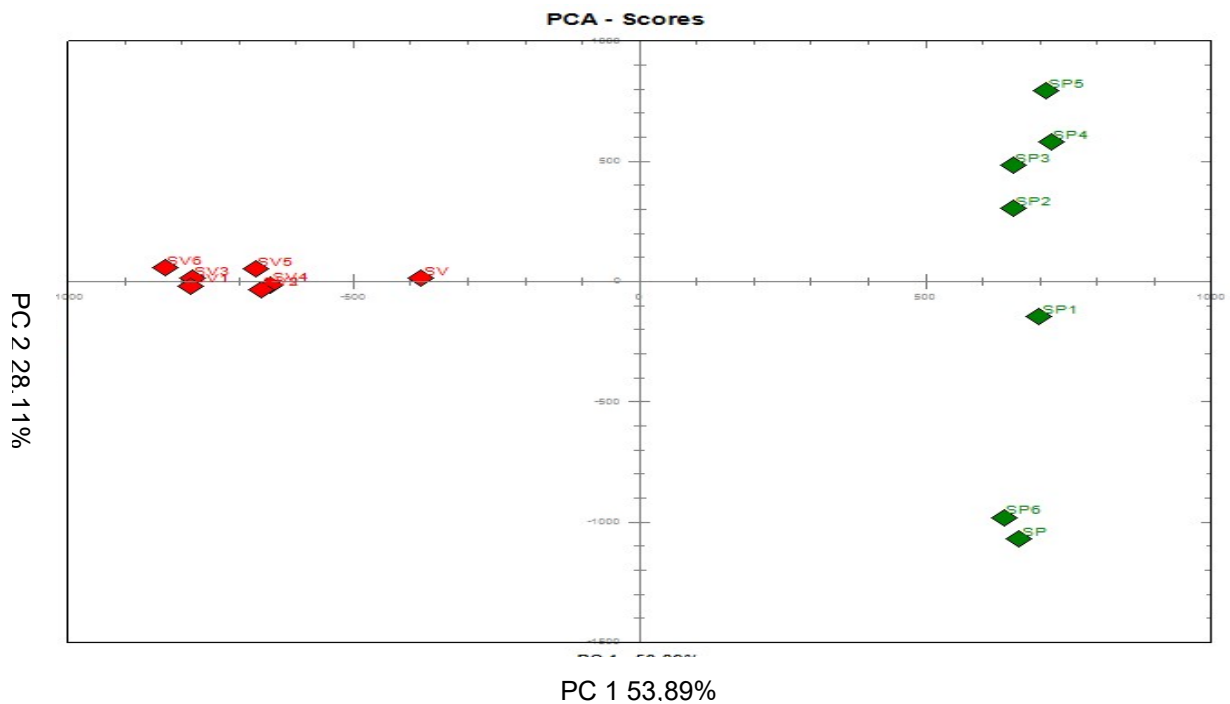


Gráfico 18: PCA Teste 8 - PCA: Mean Histograma R, G, B (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

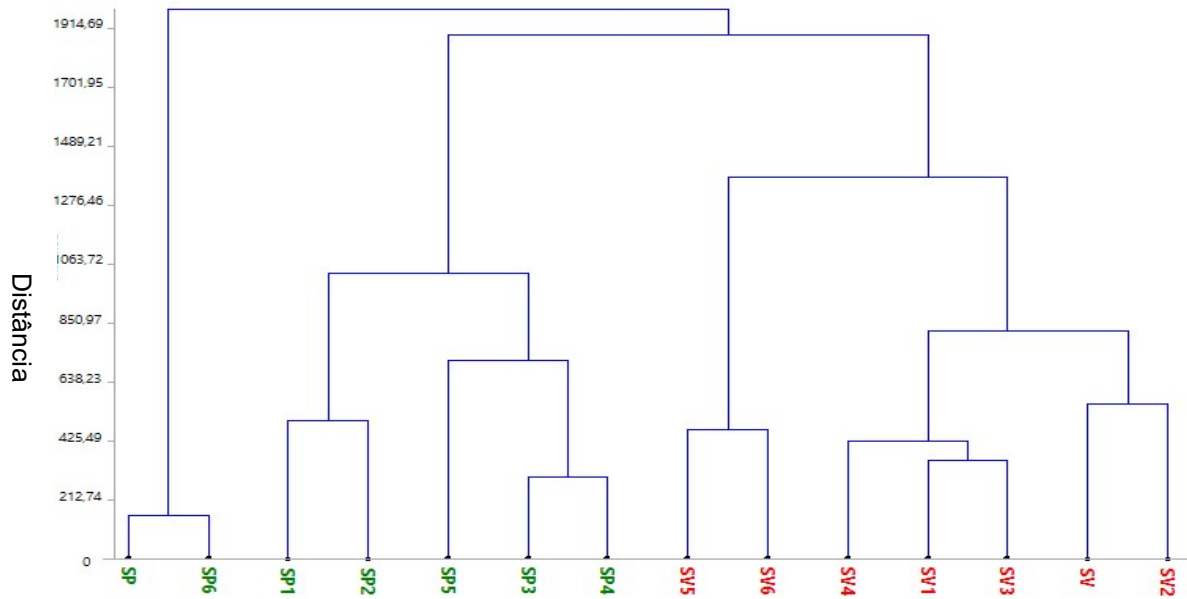


Gráfico 19: HCA Teste 8 - Complete Histograma RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana (Autoria Própria).

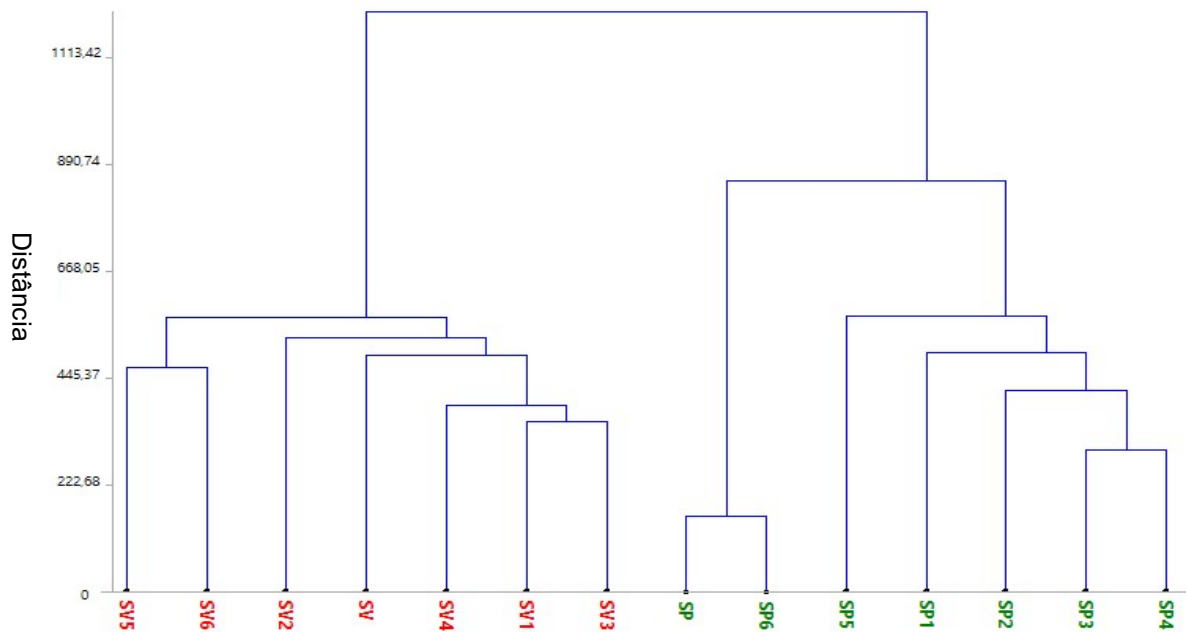


Gráfico 20: HCA Teste 8 - Single Histograma RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Single Linkage. Distância Euclidiana (Autoria Própria).

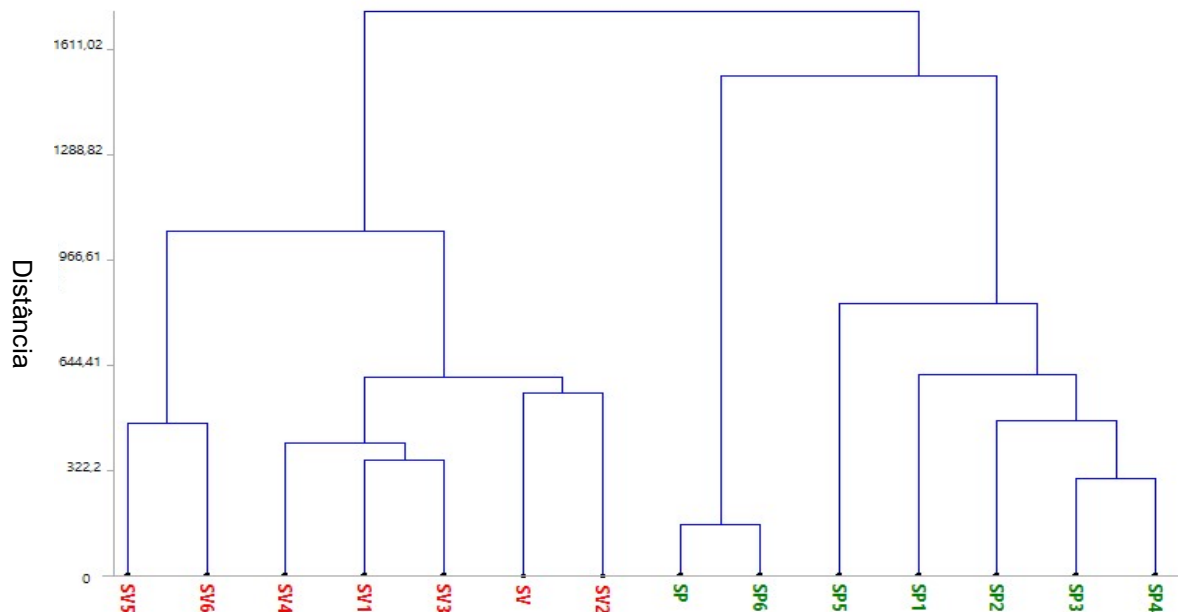


Gráfico 21: HCA Teste 8 - Average Histograma (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Average Linkage. Distância Euclidiana (Autoria Própria).

O teste de número 9, o Método Auto Single Channels foi avaliado. PCA: Auto Single Channels R, G, B; HCA: Average Single Channels RGB, Complete Single Channels RGB e Single Single Channels RGB.

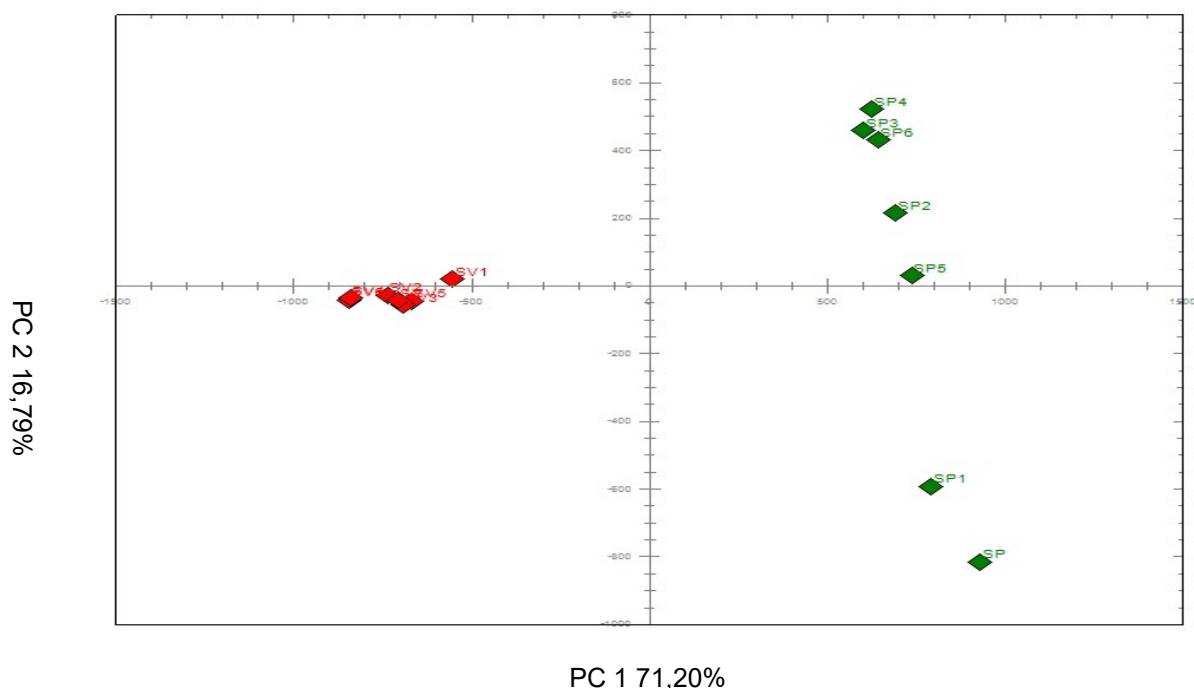


Gráfico 22: PCA Teste 9 - PCA: Auto Single Channels R, G, (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

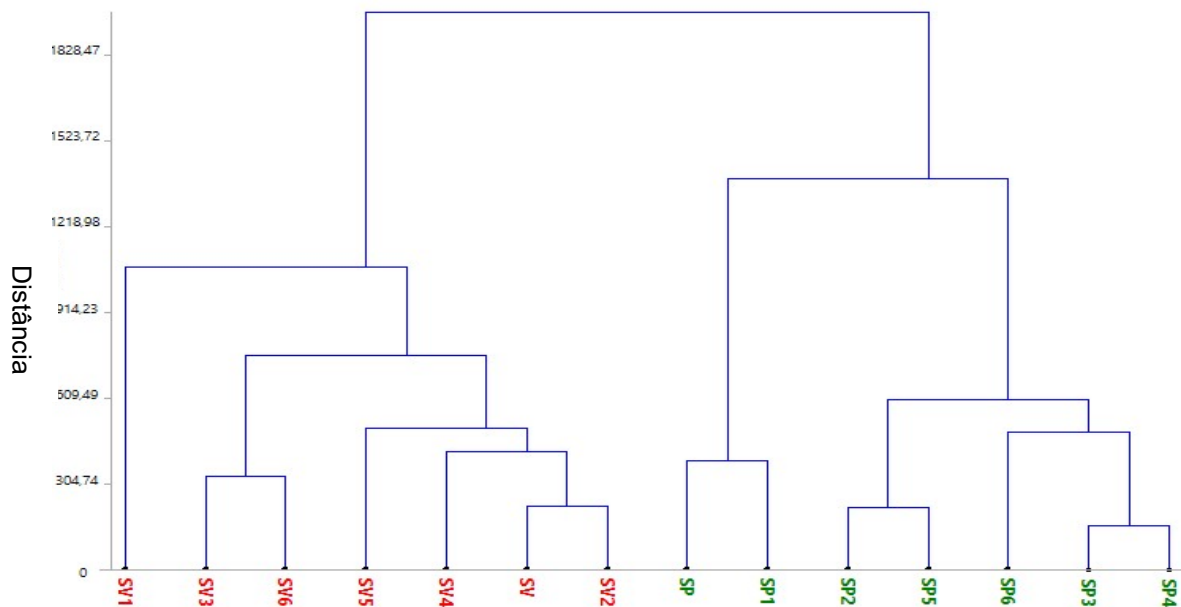


Gráfico 23: HCA Teste 9 - Complete Single Channels RGB ((SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana (Autoria Própria).

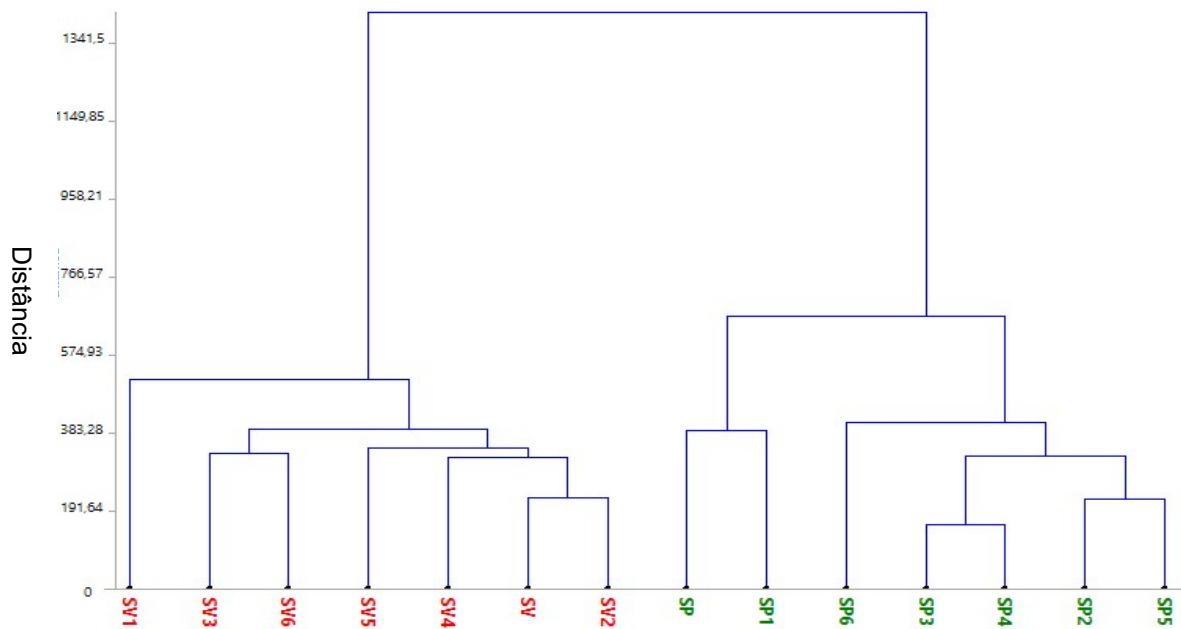


Gráfico 24: HCA Teste 9 - Single Single Channels RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Single Linkage. Distância Euclidiana (Autoria Própria).

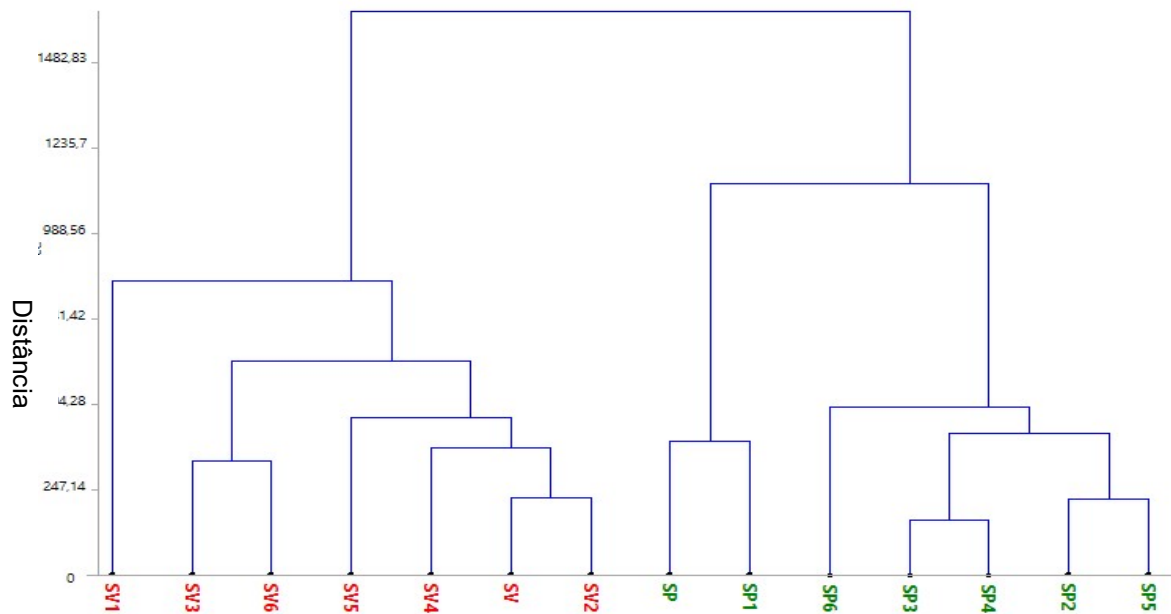


Gráfico 25: HCA Teste 9 - Average Single Channels RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Average Linkage. Distância Euclidiana (Autoria Própria).

O décimo teste foi feito no Método Mean Single Channels. PCA: Mean Single Channels R, G, B; HCA: Average Single Channels RGB, Complete Single Channels RGB e Single Single Channels RGB foram os tipos de análises exploratórias avaliadas.

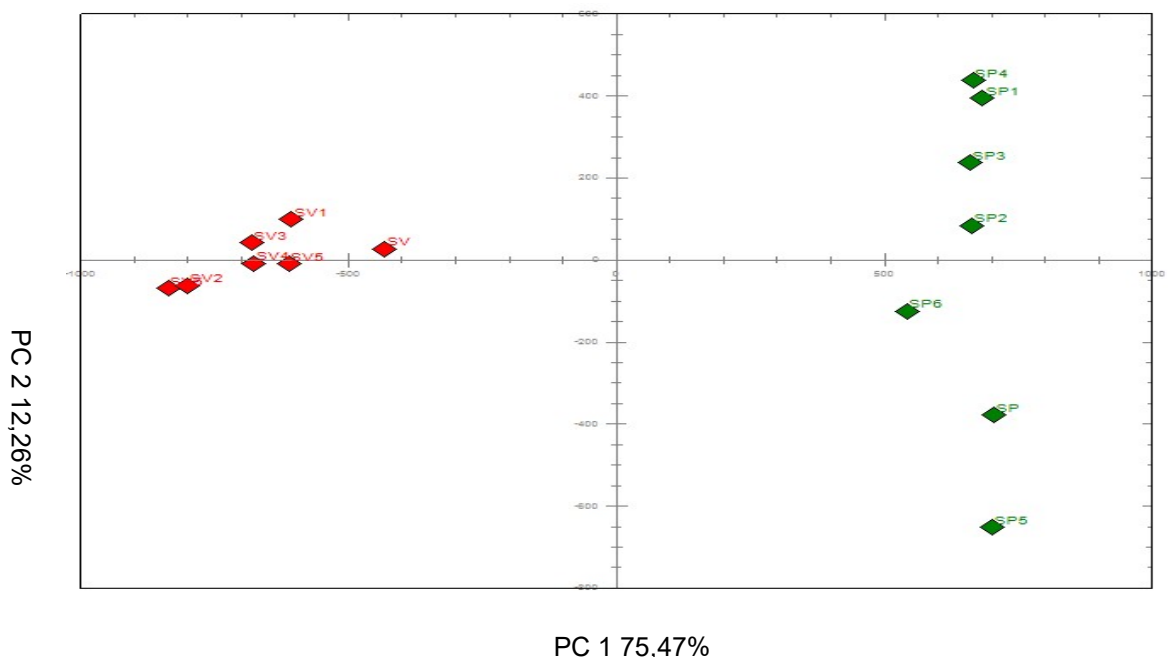


Gráfico 26: PCA Teste 10 - PCA: Mean Single Channels R, G, (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

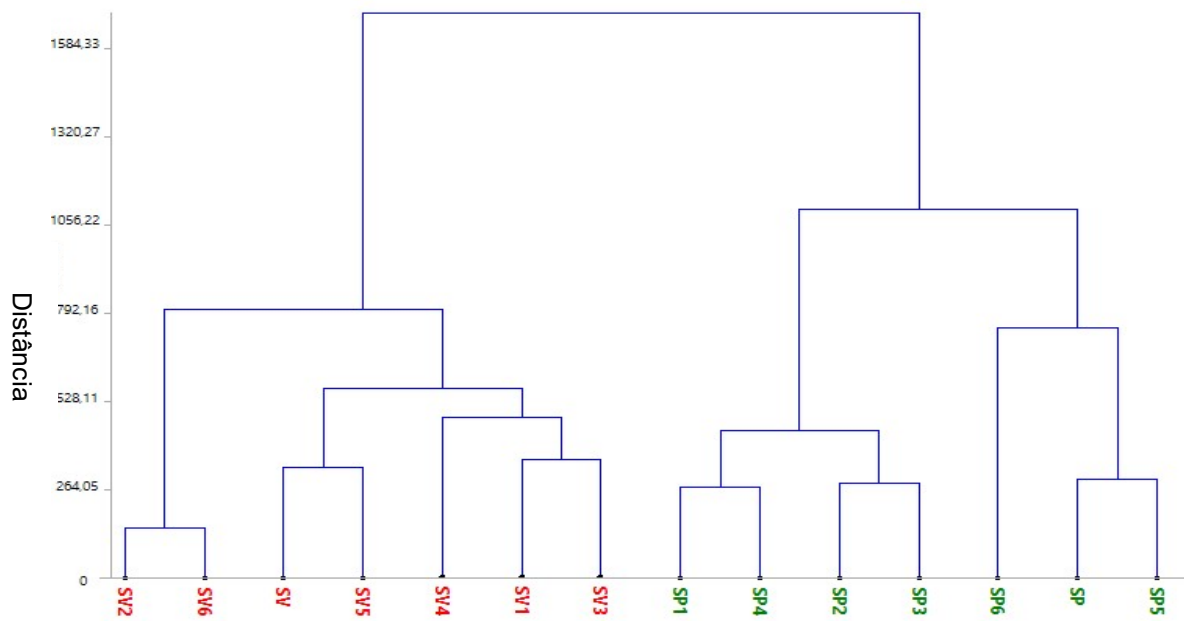


Gráfico 27: HCA Teste 10 - Complete Single Channels RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana (Autoria Própria).

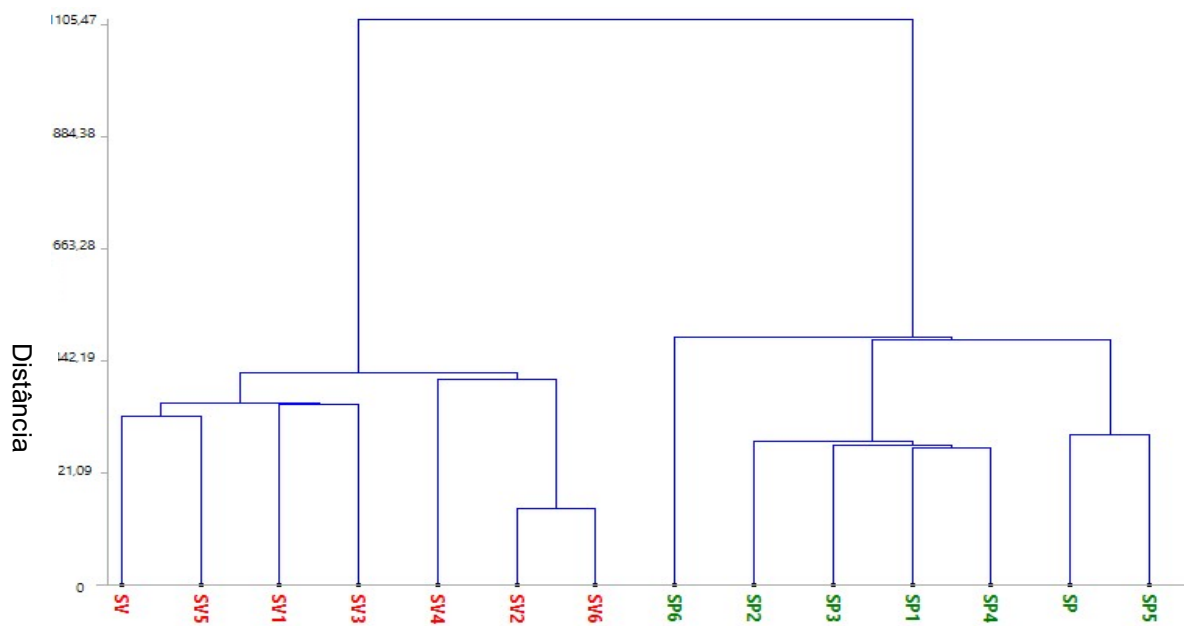


Gráfico 28: HCA Teste 10 - Single Single Channels RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Single Linkage. Distância Euclidiana (Autoria Própria).

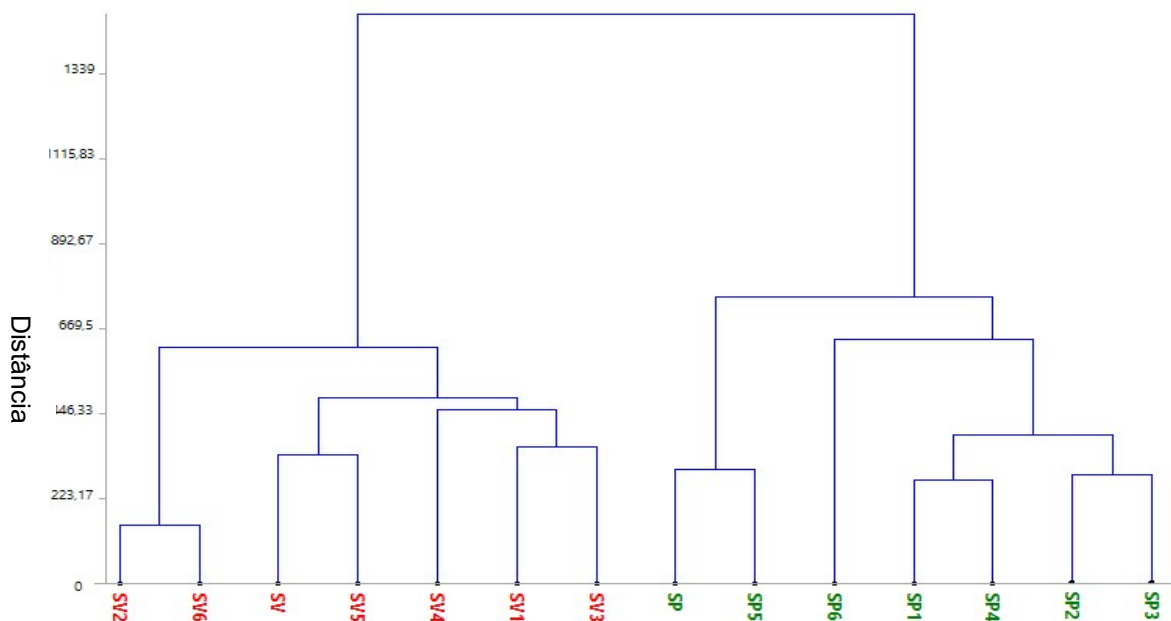


Gráfico 29: HCA Teste 10 - Average Single Channels RGB (SP em verde amostra 0% e SV em vermelho amostra adulterada em 25% m/m) Método Average Linkage. Distância Euclidiana (Autoria Própria).

A partir da análise de todos estes gráficos de PCA e HCA, observou-se que todos os métodos de aglomeração e de agrupamentos hierárquicos apresentaram boa classificação das amostras. Assim, o método escolhido foi o Mean Single Channel.

5.3 PCA, HCA e CNN.

Diante das melhores condições para captação das imagens digitais e métodos aglomerativos dos dados foram definidos e iniciaram-se as análises.

Em um primeiro momento, 80 amostras foram analisadas com apenas 03 adulterações, com massa de 02 (duas) gramas cada, sendo elas: 20 amostras 0% (SP), 30 amostras adulteradas em 10% m/m (SD) e 30 amostras adulteradas em 25% m/m (SV).

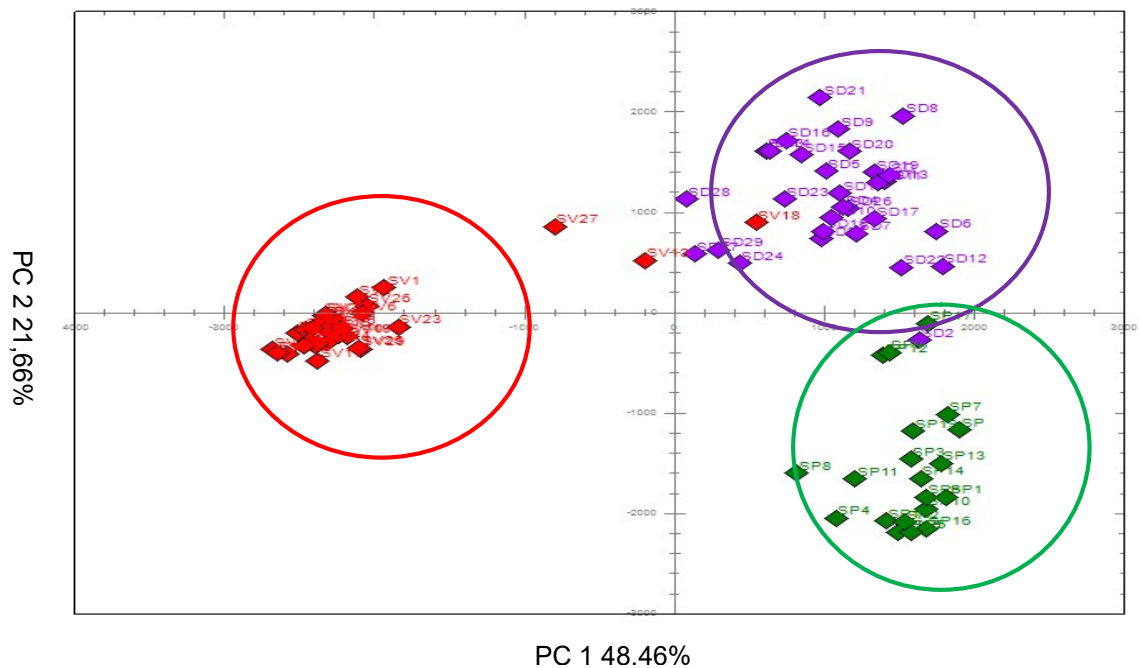


Gráfico 30: PCA - 80 amostras (SP em verde amostra 0%, SD em roxo amostra adulterada em 10% m/m e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

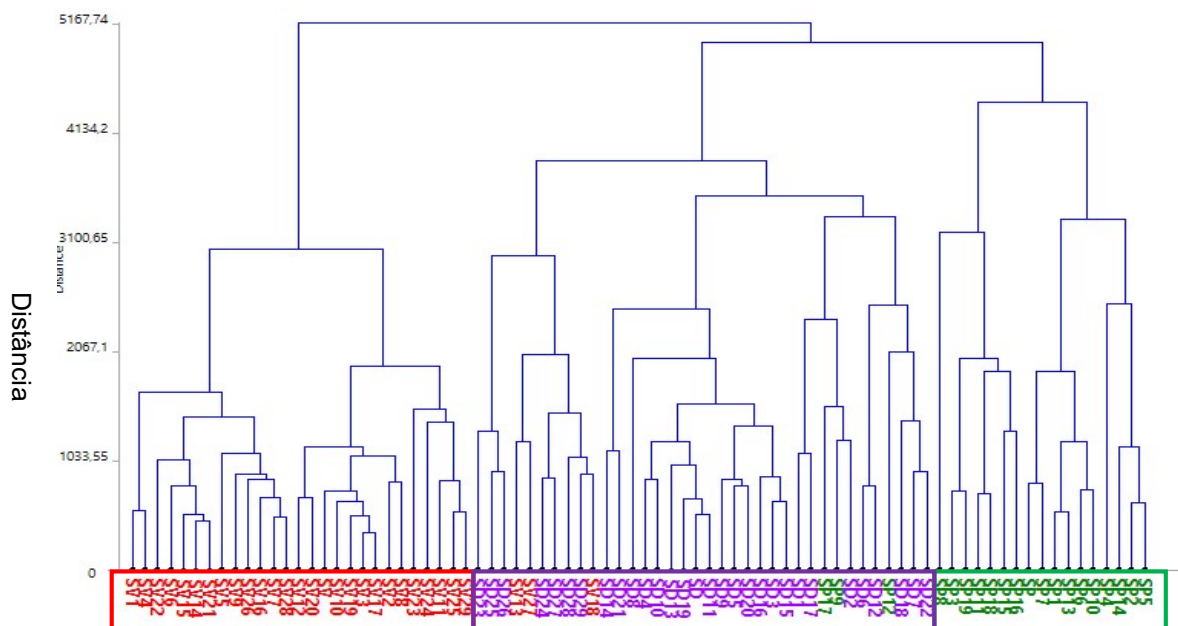


Gráfico 31: HCA - 80 amostras (SP em verde amostra 0%, SD em roxo amostra adulterada em 10% m/m e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana (Autoria Própria).

A partir da análise dos gráficos 28 e 29 observou-se que houve a separação das amostras, com pequenos erros. As amostras de 0% e 10% m/m foram as que mais apresentaram esses erros, já as amostras adulteradas em

25% m/m apresentaram uma separação mais efetiva, pois as amostras 0% e 10% m/m apresentaram características mais próximas do que as amostras de 25% m/m de adulteração. Têm-se também no gráfico PCA, uma PC1 explicando 48,46% da variância e a PC2 21,66%.

Em um segundo momento, 120 amostras foram utilizadas, pesando 02 (dois) gramas cada uma, e 04 adulterações, sendo 0%, 10%, 15% e 25% m/m de bicarbonato de sódio. O objetivo de realizar tantas adulterações foi avaliar se haveria diferença nos métodos aglomerativos a partir do aumento da quantidade de amostras. As captações foram feitas em triplicata, totalizando 360 captações. Foi nessa última etapa que a Inteligência Artificial foi utilizada, através das Redes Neurais Convolucionais.

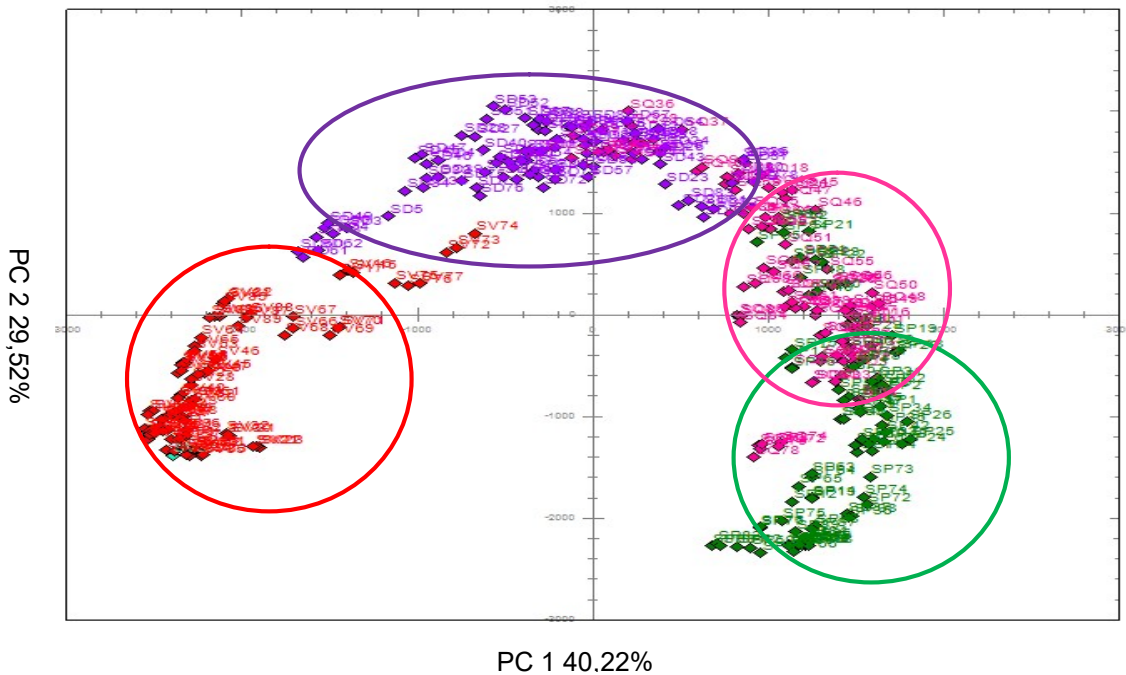


Gráfico 32: PCA – 360 captações (SP em verde amostra 0%, SD em roxo amostra adulterada em 10% m/m, SQ em rosa amostra adulterada em 15% m/m e SV em vermelho amostra adulterada em 25% m/m) (Autoria Própria).

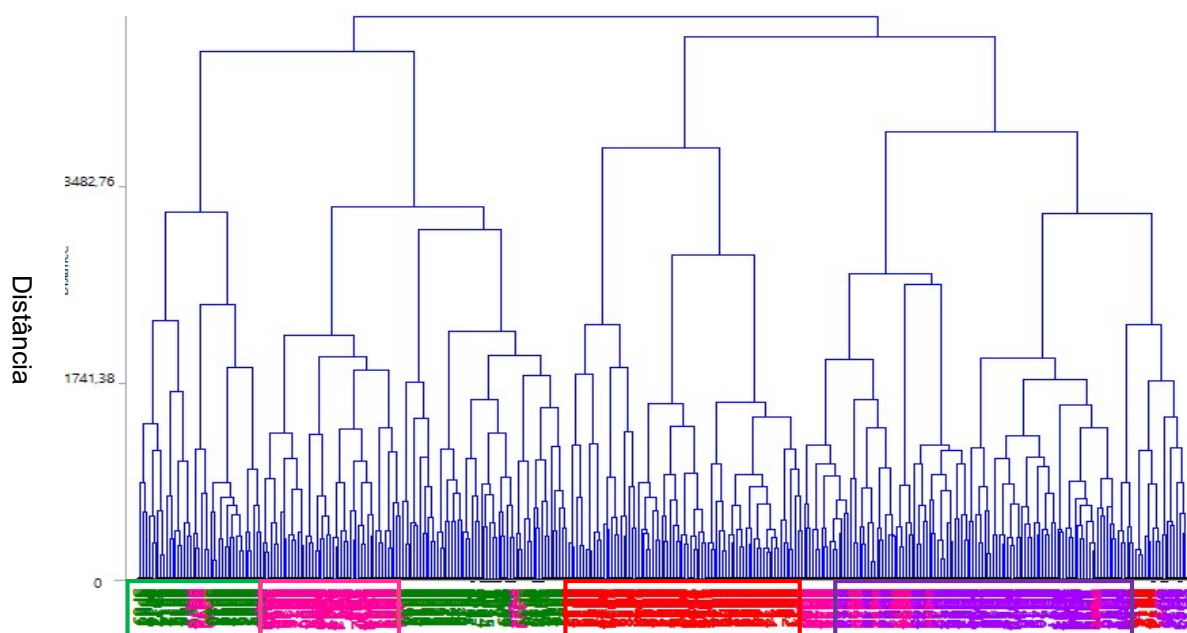


Gráfico 33: HCA – 360 captações (SP em verde amostra 0%, SD em roxo amostra adulterada em 10% m/m, SQ em rosa amostra adulterada em 15% m/m e SV em vermelho amostra adulterada em 25% m/m) Método Complete Linkage. Distância Euclidiana (Autoria Própria).

Observou-se que a quantidade de amostras e a qualidade dos gráficos de PCA e HCA cresceram de forma inversamente proporcional, pois à medida que o número de amostras aumentou houve uma perda significativa na capacidade do algoritmo em classificar as amostras. Nos gráficos 30 e 31, os dados fornecidos estão tão aglomerados que esse tipo de análise exploratória, não é eficiente na distinção das diferentes adulterações. Assim, esses métodos de aglomeração de dados apresentam funcionalidade quando estudadas 80 amostras (Gráficos 28 e 29), contudo quando 360 amostras são utilizadas há um aumento no erro de probabilidade. Além disso, PCA e HCA não são bons algoritmos de análise discriminatória.

Diante de todas essas limitações supracitadas, a IA através das CNN foi utilizada neste estudo. A vantagem do uso da CNN em Química, em detrimento das demais técnicas de IA, é que a CNN é capaz de realizar distinção de imagens digitais em classes (SANTOS *et. al.*, 2010), analisar um volume grande de dados com redução dos erros nas análises. Assim, o uso da CNN propicia análises rápidas, eficientes e minimizam erros analíticos (LECUN *et al.*, 2015). A CNN processa dados presentes em uma imagem nos três canais

de cores: RGB (LECUN *et al.* 2015) e fornece informações confiáveis em um curto período de tempo.

As CNN utilizam algoritmos de aprendizado profundo para realizar a caracterização das imagens. Já existem diversos algoritmos atualmente que fazem esse trabalho: AlexNet (KRIZHEVSKY *et al.*, 2012), VGG (SIMONYAN *et al.*, 2014), Inception (CHEN, *et al.*, 2017), GoogleNet (LIN *et al.*, 2013) e ResNet (HE *et al.*, 2016).

Redes neurais residuais - ResNet-18 e ResNet-50 - com 18 camadas e 50 camadas, respectivamente, foram utilizados neste estudo. Na figura 44, têm-se a estrutura desses algoritmos. Na primeira linha observa-se uma estrutura com uma camada convolucional inicial (conv1), nas linhas abaixo quatro conjuntos de blocos residuais e uma camada totalmente conectada no final.

Nome Camada	Tamanho da Saída	18 Camadas (ResNet18)	50 Camadas (ResNet50)
conv1	112 x 112	<i>7 x 7, 64, Passo 2</i> 3 x 3, Max Agrupamento, Passo 2	
conv2_x	56 x 56	$\begin{pmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{pmatrix} \times 2$	$\begin{pmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 128 \end{pmatrix} \times 3$
conv3_x	28 x 28	$\begin{pmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{pmatrix} \times 2$	$\begin{pmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 256 \end{pmatrix} \times 4$
conv4_x	14 x 14	$\begin{pmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{pmatrix} \times 2$	$\begin{pmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1028 \end{pmatrix} \times 6$
conv5_x	7 x 7	$\begin{pmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{pmatrix} \times 2$	$\begin{pmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2028 \end{pmatrix} \times 3$
	1 x 1	Agrupamento Médio, 1000-d fc, softmax	
FLOPs (Operações de ponto flutuante por segundo)		$1,8 \times 10^9$	$3,6 \times 10^9$

Figura 17: Arquitetura dos algoritmos das redes neurais residuais ResNet18 e ResNet50 (Autoria Própria)

Quando as imagens de entrada definidas por categorias são carregadas no algoritmo, as redes são treinadas para serem invariantes às distorções nos dados da imagem. As 360 amostras foram divididas em categorias com 57

amostras de treinamento, 23 amostras de validação e 13 amostras de teste (Tabela 6).

Parâmetros	Quantidade
População Total	360
Conjunto de treinamento para categoria	54
Conjunto de validação para categorias	23
Conjunto de teste para categorias	13

Tabela 6: Principais parâmetros utilizados na CNN (Autoria Própria).

Na etapa de validação dos dados, duas matrizes de confusão foram criadas. Uma para a rede neural residual ResNet-18 (Figura 18) e outra para a ResNet-50 (Figura 19). A função da matriz de confusão é verificar os dados obtidos na avaliação de modelos de classificação, em categorias. Desse modo, é um sistema de classificação binário, do tipo SIM ou NÃO (WATZLAWICK *et al.*, 2003). Nas figuras 20 e 21 têm-se os dados obtidos pelas matrizes de confusão da ResNet-18 e ResNet-50, respectivamente.

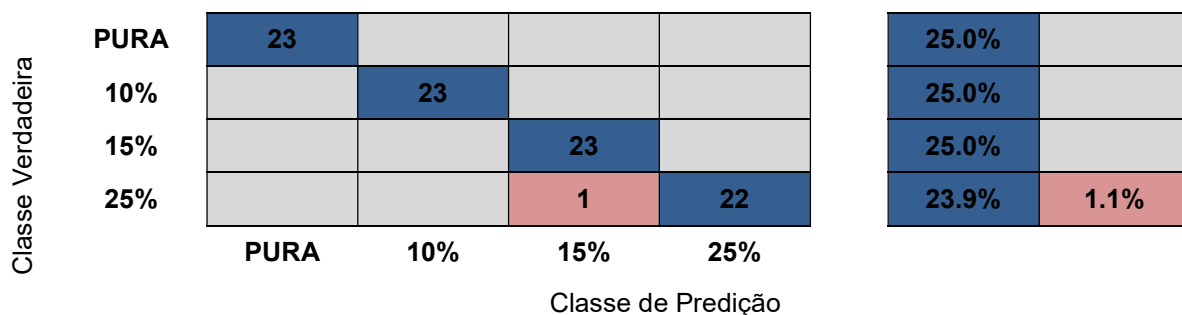


Figura 18: Matriz de Confusão ResNet18(Autoria Própria).

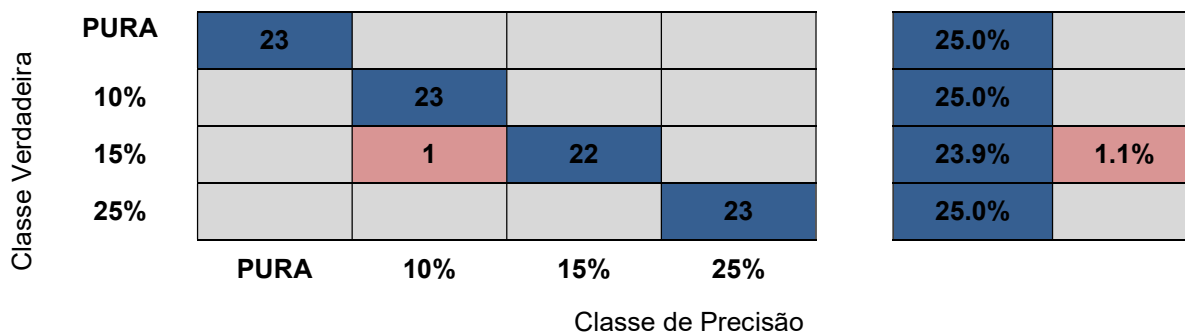


Figura 19: Matriz de Confusão ResNet50 (Autoria Própria).

Classe Saída	Pura	23 25.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	10%	0 0.0%	23 25.0%	0 0.0%	0 0.0%	100% 0.0%
	15%	0 0.0%	0 0.0%	23 25.0%	1 1.1%	95.8% 4.2%
	25%	0 0.0%	0 0.0%	0 0.0%	22 23.9%	100% 0.0%
		100% 0.0%	100% 0.0%	100% 0.0%	95.7% 4.3%	98.9% 1.1%
	Pura	10%	15%	25%		
		Classe Alvo				

Figura 20: Dados da Matriz de Confusão ResNet18 (Autoria Própria).

Classe Saída	Pura	23 25.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	10%	0 0.0%	23 25.0%	1 1.1%	0 0.0%	95.8% 4.2%
	15%	0 0.0%	0 0.0%	22 23.9%	0 0.0%	100% 0.0%
	25%	0 0.0%	0 0.0%	0 0.0%	23 25.0%	100% 0.0%
		100% 0.0%	100% 0.0%	95.7% 4.3%	100% 0.0%	98.9% 1.1%
	Pura	10%	15%	25%		
		Classe Alvo				

Figura 21: Dados da Matriz de Confusão ResNet50 (Autoria Própria).

Observou-se que, na fase de validação, tanto a ResNet-18 quanto a ResNet-50 obtiveram o mesmo percentual de acerto na classificação das amostras (98,9%). Em ambos, os algoritmos erraram apenas uma amostra no momento da classificação. Na ResNet-18 uma amostra de 25% de contaminação foi classificada como 15% e na ResNet-50 uma amostra de 15% foi classificada como 10%.

Na Tabela 7, tem-se a taxa de acerto obtida na fase de validação e na fase de testes durante a classificação, utilizando as duas arquiteturas de redes neurais residuais (ResNet-18 e ResNet-50) como extratores de características para as categorias definidas nos experimentos.

Categorias	ResNet-18		ResNet-50	
	% Validação	% Teste	% Validação	% Teste
Pura	100	100	100	100
10%	100	62	95,8	100
15%	95,8	100	100	92
25%	100	92	100	100
Média	98,9	88,5	98,9	98

Tabela 7: Comparação da precisão da validação e precisão do teste para as arquiteturas ResNet-18 e ResNet-50 entre categorias (Autoria Própria).

Os resultados apresentados na Tabela 6 mostraram a maior taxa de precisão do ResNet-50 nos testes, enquanto na validação dos resultados ambos os sistemas tiveram a mesma taxa de precisão. Além disso, os dados demonstraram a dificuldade do ResNet-18 em testar imagens com 10% de contaminação, pois obtivemos 62% de precisão, diferentemente do ResNet-50, que obteve 100%.

6 CONCLUSÃO

Utilizando redes neurais convolucionais foi possível classificar as amostras de *Spirulina* possibilitando avaliar a adulteração do produto através das imagens digitais. O método utilizado é simples, rápido e de baixo custo quando comparados com outros métodos reportados na literatura para detecção de adulteração de alimentos. As Redes Neurais Convolucionais associada a imagens digitais mostrou-se muito eficientes na classificação das amostras contaminadas com diferentes níveis de adulteração frente ao uso de PCA e HCA, ou seja, pode ser uma importante ferramenta utilizada na avaliação da qualidade da *Spirulina*. Assim, os objetivos deste estudo foram alcançados com sucesso.

Desse modo, a Inteligência Artificial, especialmente as redes neurais convolucionais são uma aliada com grande potencial de crescimento e diversificação em análises, podendo ser amplamente utilizadas nas diversas áreas dentro da Química.

Como perspectivas futuras, existem as possibilidades, a utilização da calibração multivariada; mais níveis de adulteração podem ser utilizados, como valores abaixo de 10%; e testar outros algoritmos de Inteligência Artificial.

REFERÊNCIAS

- Africa • Spirulina producers report.** Disponível em: <<https://www.spirulinaresource.com/spirulina-reports/africa/>>. Acesso em: 4 nov. 2023.
- ALCANTARA, M. Del M. G. Revisión del estado actual de la problemática y de los métodos de análisis para determinación de metales pesados en espirulina. Universidad de Jaén. (2018). Disponível em: <https://tauja.ujaen.es/bitstream/10953.1/8732/1/TFG_Guardia_Alcantara_Maria_del_Mar.pdf>.
- ARAÚJO, A.; MARINHO, W.; DE ARAÚJO GOMES, A. A fast and inexpensive chemometric-assisted method to identify adulteration in acai (*Euterpe oleracea*) using digital images. **Food analytical methods**, v. 11, n. 7, p. 1920–1926, 2018.
- AUED-PIMENTEL, S.; CASTRO, F. D.; SOUSA, R. J.; AMARAL-MELLO, M. R. P.; ABE-MATSUMOTO, L. T. Commercialised coconut oil in São Paulo City, Brazil: Evaluation of authenticity and nutritional labelling. **Journal of Agriculture and Life Sciences**, 2(1), 76–83. 2015
- AUED-PIMENTEL, S.; TAKEMOTO, E.; KUMAGAI, E. E.; CANO, C. B. Calculation of the difference between the actual and theoretical ECN 42 triacylglyceride content to detect adulteration in olive oil samples commercialized in Brazil. **Química Nova**, 31(1), 31–34, 2008. <https://doi.org/10.1590/S0100-40422008000100006>.
- BOCHIE, K.; GILBERT, M. S.; GANTERT, L.; BARBOSA, M. S. M.; MEDEIROS, D. S. V.; CAMPISTA, M. E. M. Aprendizado Profundo em Redes Desafiadoras: Conceitos e Aplicações. Em: **Minicursos do XXXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos**. [s.l.] SBC, 2020. p. 140–189.
- BÖCK, F. C.; HELFER, G. A.; DA COSTA, A. B.; DESSUY, M. B.; FERRÃO, M. FPhotoMetrix and colorimetric image analysis using smartphones. **Journal of chemometrics**, v. 34, n. 12, 2020.
- BOTELHO, B. G.; DANTAS, K. C. F.; SENA, M. M. Determination of allura red dye in hard candies by using digital images obtained with a mobile phone and N-PLS. **Chemometrics and intelligent laboratory systems: an international journal sponsored by the Chemometrics Society**, v. 167, p. 44–49, 2017.
- CALDEIRA, L. A.; JUNIOR, V. R. R.; FONSECA, C. M.; MELO, L. M. de.; CRUZ, A. G.; OLIVEIRA, L. L. dos S. Characterization of milk commercialized in Janaúba–MG. **Alimentos e Nutrição**. Araraquara, 21(2), 191–196. 2010

CEBI, N.; DOGAN, C. E.; DEVELIOGLU, A.; YAYLA, M. E. A.; SAGDIC, O. Detection of L-Cysteine in wheat flour by Raman microspectroscopy combined chemometrics of HCA and PCA. **Food chemistry**, v. 228, p. 116–124, 2017.

CHOODUM, A. ; KANATHARANA, P. ; WONGNIRAMAİKUL, W. ; NIC, D. Using the iPhone as a device for a rapid quantitative analysis of trinitrotoluene in soil. **Talanta**, v. 115, p. 143–149, 2013.

CIFERRI, O. **Spirulina, the Edible Microorganism**. Microbiological Reviews, Dec. 1983, p. 551-578.

COSTA, J. A. V.; FREITAS, B. C. B.; ROSA, G. M.; MORAES, L.; MORAIS, M. G.; MITCHELL, B. G. Operational and economic aspects of Spirulina-based biorefinery. **Bioresource technology**, v. 292, n. 121946, p. 121946, 2019.

COUTINHO, M. S.; MORAIS, C. L. M.; NEVES, A. C. O.; MENEZES, F. G.; LIMA, K. M. G. Colorimetric determination of ascorbic acid based on its interfering effect in the enzymatic analysis of glucose: An approach using smartphone image analysis. **Journal of the Brazilian Chemical Society**, 2017.

DAMACENO, S. S.; VASCONCELOS, R. O. Inteligência Artificial: uma breve abordagem sobre seu conceito real e o conhecimento popular. **Caderno de Graduação - Ciências Exatas e Tecnológicas** – UNIT/SERGIPE, 2018, v. 5, n.1, 11 Recuperado

de <https://periodicos.set.edu.br/cadernoexatas/article/view/5729>

DARTSCH, P. C. Antioxidant potential of selected Spirulina platensis preparations. **Phytotherapy research: PTR**, v. 22, n. 5, p. 627–633, 2008.

DELANEY, J. L. ; HOGAN, C. F. ; TIAN, J. ; SHEN, W. Electrogenated chemiluminescence detection in paper-based microfluidic sensors. **Analytical chemistry**, v. 83, n. 4, p. 1300–1306, 2011.

ESPINOZA, T.; MESA, F. R.; VALENCIA, E.; QUEVEDO, R. Types of fraud in meat and meat products: a review. **Scientia Agropecuaria 6 (3): 223 – 233 (2015)**

EVERSTINE, K.; SPINK, J.; KENNEDY, S. Economically motivated adulteration (EMA) of food: common characteristics of EMA incidents. **Journal of food protection**, v. 76, n. 4, p. 723–735, 2013.

FERRÃO, M. F.; MELLO, C.; BORIN, A.; MARETTO, D. A.; POPPI, R. J. LS-SVM: uma nova ferramenta quimiométrica para regressão multivariada. Comparação de modelos de regressão LS-SVM e PLS na quantificação de adulterantes em leite em pó empregando NIR. **Química nova**, v. 30, n. 4, p. 852–859, 2007.

FERREIRA, M. Quimiometria III – revisitando a análise exploratória dos dados multivariados. **Química nova**, 2022.

FERREIRA, M. M. *Spirulina*: uma revisão. Disponível em: <
<https://repositorio.ufu.br/bitstream/123456789/29490/1/SpirulinaUmaRevis%C3%A3o.pdf>> Acesso em 01/11/2023.

FERREIRA, M. M. C. et al. Quimiometria I: calibração multivariada, um tutorial. **Química nova**, v. 22, n. 5, p. 724–731, 1999.

FILHO, O. M.; NETO, H. V. Processamento Digital de Imagens, Rio de Janeiro: **Brasport**, 1999. ISBN 8574520098.

FRANCO, M. de O. K.; SUAREZ, W. T.; MAIA, M. V.; SANTOS, V. B. dos. Smartphone application for methanol determination in sugar cane spirits employing digital image-based method. **Food analytical methods**, v. 10, n. 6, p. 2102–2109, 2017.

FURTADO, M. I. V. Redes neurais artificiais [recurso eletrônico]: uma abordagem para sala de aula / Maria Inês Vasconcellos Furtado. – Ponta Grossa (PR): **Atena Editora**, 2019.

GALVIN-KING, P.; HAUGHEY, S. A.; ELLIOTT, C. T. Herb and spice fraud; the drivers, challenges and detection. **Food control**, v. 88, p. 85–97, 2018.

GALVIN-KING, P. ; HAUGHEY, S. A.; ELLIOTT, C. T. The Detection of Substitution Adulteration of Paprika with Spent Paprika by the Application of Molecular Spectroscopy Tools. **Foods**, 9, 944, 2020; DOI:10.3390/foods9070944.

GARCIA, D. P. ; CARASCHI, J. C. ; VENTORIM, G. ; VIEIRA, F. H. A. ; PROTÁSSIO, T. de P. Assessment of plant biomass for pellet production using multivariate statistics (PCA and HCA). **Renewable energy**, v. 139, p. 796–805, 2019.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Londres, England: MIT Press, 2016.

GRIGOLO, S.; FIOREZE, A. C. da C. L., DENARDI, S.; VACARI, J. Implicações da análise univariada e multivariada na dissimilaridade de acessos de feijão comum. **Revista de Ciências Agroveterinárias**, v. 17, n. 3, p. 351–360, 2018.

HE, X, K.; ZHANG, S. R. ; SUN, J. **Deep residual learning for image recognition**. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). **Anais...IEEE**, 2016.

HELPER, G. A. Chemostat: Desenvolvimento de *Software* para Análise Exploratória de Dados Multivariados. **Santa Cruz do Sul**, 2014.

HELPER, G. A.; BOCK, F.; MARDER, L.; FURTADO, J. C.; COSTA, A. B. da; FERRÃO, M. F. Chemostat: Exploratory multivariate data analysis software. **Química nova**, 2015.

HENRIKSON, R. Earth Food Spirulina. **Ronore Enterprises**, Inc. Sixth Printing: 2009.

JOAO, A.F. ; SQUISSATO, A.L. ; FERNANDES, G.M. ; CARDOSO, R.M. ; BATISTA, A. D. ; MUÑOZ, R. A. A. Iron (III) determination in bioethanol fuel using a smartphone-based device. **Microchemical journal, devoted to the application of microtechniques in all branches of science**, v. 146, p. 1134–1139, 2019.

JUNG, F. ; KRÜGER-GENGE, A. ; WALDECK, P. ; KÜPPER, J.H. Spirulina platensis, a super food? **Journal of cellular biotechnology**, v. 5, n. 1, p. 43–54, 2019.

KARKOS, P. D. et al. *Spirulina* in clinical practice: Evidence-based human applications. **Evidence-based complementary and alternative medicine: eCAM**, v. 2011, p. 1–4, 2011.

KÉPEŠ, E. ; VRÁBEL, J. ; BRÁZDIL, T. ; HOLUB, P. ; POŘÍZKA, P. ; KAISER, J. Interpreting convolutional neural network classifiers applied to laser-induced breakdown optical emission spectra. **Talanta**, v. 266, n. 124946, p. 124946, 2024.

KOH, A. ; KANG, D. ; XUE, Y. ; LEE, S. ; PIELAK, R. M., KIM, J. ; HWANG, T. ; MIN, S. ; BANKS, A. ; BASTIEN, P. ; MANCO, M. C. ; WANG, L. ; AMMANN, K. R. ; JANG, K-IN. ; WON, P. ; HAN, S. ; GHAFARI, R. ; PAIK, U. ; SLEPIAN, M. J. ; BALOOCH, G. ; HUANG, Y. ; ROGERS, J. A. A soft, wearable microfluidic device for the capture, storage, and colorimetric sensing of sweat. **Science translational medicine**, v. 8, n. 366, p. 366ra165, 2016.

KULSHRESHTHA, A. et al. Spirulina in health care management. **Current pharmaceutical biotechnology**, v. 9, n. 5, p. 400–405, 2008.

LAFARGA, T. ; FERNÁNDEZ-SEVILLA, J. M. ; GONZÁLEZ-LÓPEZ, C. ; ACIÉN-FERNÁNDEZ, F. G. Spirulina for the food and functional food industries. **Food research international (Ottawa, Ont.)**, v. 137, n. 109356, p. 109356, 2020.

LAMARCA, R. S. ; LUCHIARI, N. D. C. ; BONJORNO, A. F. ; PASSARETTI FILHO, J. ; CARDOSO, A. A. ; DE LIMA GOMES, P. C. F. Determination of formaldehyde in cosmetic products using gas-diffusion microextraction coupled with a smartphone reader. **Analytical methods: advancing methods and applications**, v. 11, n. 29, p. 3697–3705, 2019.

LAMARCA, R. S.; LIMA GOMES, P. C. F. DE. A low cost method for carbamazepine, ciprofloxacin and norfloxacin determination in pharmaceutical formulations based on spot-test and smartphone images. **Microchemical journal, devoted to the application of microtechniques in all branches of science**, v. 152, n. 104297, p. 104297, 2020.

LEAO, A. C. Gerenciamento de cores para imagens digitais. 2005.

LECUN, Y. ; BOTTOU, L. ; BENGIO, Y. ; HAFFNER, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE. Institute of Electrical and Electronics Engineers**, v. 86, n. 11, p. 2278–2324, 1998.

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

LIMA, C. M. De ; FERNANDES, D. D. S. ; PEREIRA, G. E. ; GOMES, A. de A. ; ARAÚJO, M. C. U. De ; DINIZ, P. H. G. D. Digital image-based tracing of geographic origin, winemaker, and grape type for red wine authentication. **Food chemistry**, v. 312, n. 126060, p. 126060, 2020.

LIMA, J. S.; SAMPAIO, A. P. P. O.; SILVA, P. S.; DUFOSSÉ, M. C. da S.; ROSA, A. M. B. P.; MORAES, C. M. de; ROSS, T. B. Avaliação de métodos de detecção da fosfatase alcalina em leite bovino, bubalino e caprino. **Brazilian journal of food technology**, v. 24, 2021.

LIU, Y.; LI, R.; XIONG, B.; ZHANG, S.; ZHANG, X.; IU, H.; FERNANDO, T. A novel vanadium redox flow battery modelling method using honey badger optimization assisted CNN-BiLSTM. **Journal of power sources**, v. 558, n. 232610, p. 232610, 2023.

LOSEVA, L. P.; DARDYNSKAYA, I. V. Spirulina-natural sorbent of radionucleides. In: **Proceedings of the 6th International Congress of Applied Algology**. 1993.

LUDERMIR, T. B. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. **Estudos Avançados**, v. 35, n. 101, p. 85–94, 2021.

LUDWIG, S. K.; ZHU, H; PHILLIPS, S.,; SHILEDAR, A.; FENG, S.; TSENG, D.; VAN GINKEL, L. A.; NIELEN, M. W.; OZCAN, A. Cellphone-based detection platform for rbST biomarker analysis in milk extracts using a microsphere fluorescence immunoassay. **Analytical and bioanalytical chemistry**, v. 406, n. 27, p. 6857–6866, 2014.

MARINHO, O. R.; LIMA, M. J. A.; ROCHA, F. R. P.; REIS, B. F.; KAMOGAWA, M. Y. A greener, fast, and cost-effective smartphone-based digital image procedure for quantification of ethanol in distilled beverages. **Microchemical journal, devoted to the application of microtechniques in all branches of science**, v. 147, p. 437–443, 2019.

MARTINEZ, A. W. ; PHILLIPS, S. T. ; CARRILHO, E. ; THOMAS, S. W. III ; SINDI, H. ; WHITESIDES, G. M. Simple telemedicine for developing regions: camera phones and paper-based microfluidic devices for real-time, off-site diagnosis. **Analytical chemistry**, v. 80, n. 10, p. 3699–3707, 2008.

MARTINS, G. B. C.; SUCUPIRA, R. R.; SUAREZ, P. A. Z. Chemistry and colors. **Revista Virtual de Química**, v. 7, n. 4, p. 1508–1534, 2015.

MARTINS, R. G.; GONCALVES, I. G.; MORAIS, M. G. de; COSTA, J. A. V. Bioprocess engineering aspects of biopolymer production by the Cyanobacterium *Spirulina* strain LEB 18. **International journal of polymer science**, v. 2014, p. 1–6, 2014.

MORENO, A. S.; BEZERRA, T. S.; BEZERRA, T. S.; MELLO, N. D. Câmeras Analógicas e Digitais: Qual Oferece Melhor Qualidade? Revista Educação - UNG-Ser, v. 11, n. 2 ESP, p. 19–33, 2016.

MOTA, C.; LIMA, A.; NASCIMENTO, A.; MIRANDA, P.; de MELLO, R. **Classificação de Páginas de Petições iniciais Utilizando Redes Neurais Convolucionais Multimodais**. Anais do Encontro Nacional de Inteligência Artificial e Computacional (ENIAC 2020). **Anais...Sociedade Brasileira de Computação - SBC**, 2020.

NETO, J.C.; TERNES, S.; SOUZA, K. X .S. de; YANO, I. H.; QUEIROS, L. R. Uso de redes neurais convolucionais para detecção de laranjas no campo. 2019. Disponível em:
<<http://www.alice.cnptia.embrapa.br/alice/handle/doc/1125722>>

PANERO, F. S.; VIEIRA, M. F. P.; CRUZ, Â. M. F.; MOURA, M. F. V.; SILVA, H. E. B. da. Aplicação da análise exploratória de dados na discriminação geográfica do quiabo do Rio Grande do Norte e Pernambuco. **Eclética Química Journal**, v. 34, n. 3, p. 33–40, 2009.

PAUL, J. A.; DOUWES, M. Two-dimensional photographic posture recording and description: a validity study. **Applied ergonomics**, v. 24, n. 2, p. 83–90, 1993.

PIERINE, G. D.; FERNANDES, D. D. S.; DINIZ, P. H. G. D.; ARAÚJO, M. C. U. de; NEZIO, M. S. DI; CENTURIÓN, M. E. A digital image-based traceability tool of the geographical origins of Argentine propolis. **Microchemical journal, devoted to the application of microtechniques in all branches of science**, v. 128, p. 62–67, 2016.

PRADO, N. B. do; ABREU, C. B. de; PINHO, C. S.; JUNIOR, M. M. de N.; SILVA, M. D.; ESPINO, M.; SILVA, M. F.; DIAS, F. de S. Application of multivariate analysis to assess stress by Cd, Pb and Al in basil (*Ocimum basilicum* L.) using caffeic acid, rosmarinic acid, total phenolics, total flavonoids and total dry mass in response. **Food chemistry**, v. 367, n. 130682, p. 130682, 2022.

QUEIROZ, J. E. R.; GOMES, H. M. Introdução ao processamento digital de imagens. **Rita**, 13(2), 11-42. 2006

RAWAT, W.; WANG, Z. Deep convolutional neural networks for image classification: A comprehensive review. **Neural computation**, v. 29, n. 9, p. 2352–2449, 2017.

REILE, C. G.; RODRÍGUEZ, M. S.; FERNANDES, D. D. de S.; GOMES, A. de A.; DINIZ, P. H. G. D.; ANIBAL, C. V. DI. Qualitative and quantitative analysis based on digital images to determine the adulteration of ketchup samples with Sudan I dye. **Food chemistry**, v. 328, n. 127101, p. 127101, 2020.

RIBEIRO, F. C. P. ; OLIVEIRA, A. S. ; ARAUJO, A. ; MARINHO, W. ; SCHNEIDER, M. P. ; PINTO, L. ; GOMES, A. A. Detection oxidative degradation in lubricating oil under storage conditions using digital images and chemometrics. **Microchemical journal, devoted to the application of microtechniques in all branches of science**, v. 147, p. 622–627, 2019.

RIBEIRO, J. C. B.; CARDOSO, C. R.; ESMERINO, L. A.; SANTOS, R. D.; DEMIATE, I. M.; NOGUEIRA, A. Physicochemical and microbiological quality of grated Parmesan-type cheese sold in Ponta Grossa, Paraná (Brazil). **Revista do Instituto de Laticínios Cândido Tostes**, v. 67, n. 387, p. 21–29, 2012.

RIBEIRO, M. V. DE M.; BORALLE, N.; PEZZA, H. R.; PEZZA, L.; TOCI, A. T. Authenticity of roasted coffee using ¹H NMR spectroscopy. **Journal of food composition and analysis: an official publication of the United Nations University, International Network of Food Data Systems**, v. 57, p. 24–30, 2017.

ROCHA, J. C. Cor luz, cor pigmento e os sistemas RGB e CMY. *Revista Belas Artes*, 3(2). 2010.

SANCHEZ, J. L.; PEREIRA, S. B. G.; LIMA, P. C. de; POSSEBON, G.; TANAMATI, A.; POPPI, R. J.; TANAMATI, A. A. C.; BONA, E. Mid-infrared spectroscopy and support vector machines applied to control the hydrogenation process of soybean oil. **European Food Research and Technology**, v. 243, n. 8, p. 1447–1457, 2017.

SANTOS, A.; AIRES, K.; VERAS, R.; UCHOA, V.; SANTOS, L. **Uma Abordagem de Classificação de Imagens Dermatoscópicas Utilizando Aprendizado Profundo com Redes Neurais Convolucionais**. Anais do Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS). **Anais...Sociedade Brasileira de Computação - SBC**, 2017.

SANTOS, M. K.; FERREIRA JUNIOR, J. R.; WADA, D. T.; TENÓRIO, A. P. M.; NOGUEIRA-BARBOSA, M. H.; AZEVEDO-MARQUES, P. M. de. Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: advances in imaging towards to precision medicine. **Radiologia brasileira**, v. 52, n. 6, p. 387–396, 2019.

SCHLEDER, G. R.; FAZZIO, A. Machine Learning na Física, Química, e Ciência de Materiais: Descoberta e Design de Materiais. **Revista Brasileira de Ensino de Física**, v. 43, n. suppl 1, 2021.

SILVA, G. O. da; DE ARAUJO, W. R.; PAIXÃO, T. R. L. C. Portable and low-cost colorimetric office paper-based device for phenacetin detection in seized cocaine samples. **Talanta**, v. 176, p. 674–678, 2018.

- SILVA, J. A. S. da; MAIRINK, C. H. P. Inteligência artificial. **LIBERTAS: Revista de Ciências Sociais Aplicadas**, v. 9, n. 2, p. 64–85, 2019.
- SHANLING, H.; WENZHENG, D.; HE, S.; PENG, X.; SHOUDONG, Z.; LONG, C.; YONG, L. Real-time rubber quality model based on CNN-LSTM deep learning theory. **Materials today. Communications**, v. 35, n. 106110, p. 106110, 2023.
- SOARES, P. L. B.; DA SILVA, J. P. Aplicação de Redes Neurais Artificiais em Conjunto com o Método Vetorial da Propagação de Feixes na Análise de um Acoplador Direcional Baseado em Fibra Ótica. **Revista Brasileira de Computação Aplicada**, v. 3, n. 2, p. 58–72, 2011.
- SOUZA, V.; ARAUJO, L.; SILVA, L.; SANTOS, A. **Análise Comparativa de Redes Neurais Convolucionais no Reconhecimento de Cenas**. Anais do XI Computer on the Beach - COTB '20. **Anais...Itajaí**: Universidade do Vale do Itajaí, 2020.
- SPINK, J. ; FORTI, N. N. D. ; MOYER, D. C. ; MIAO, H. ; WU, Y. Food Fraud prevention: Policy, strategy, and decision-making – implementation steps for a government agency or industry. **Chimia**, v. 70, n. 5, p. 320, 2016.
- SPINK, J.; MOYER, D. C. Defining the public health threat of food fraud. **Journal of food science**, v. 76, n. 9, p. R157–R163, 2011.
- SPINK, J.; MOYER, D. C.; SPEIER-PERO, C. Introducing the food fraud initial screening model (FFIS). **Food control**, v. 69, p. 306–314, 2016.
- TAVARES, L. A.; MEIRA, M. C.; AMARAL, S. F. DO. Inteligência Artificial na Educação: Survey. **Brazilian Journal of Development**, v. 6, n. 7, p. 48699–48714, 2020.
- THANAKIATKRAI, P.; YAODAM, A.; KITPIPIT, T. Age estimation of bloodstains using smartphones and digital image analysis. **Forensic science international**, v. 233, n. 1–3, p. 288–297, 2013.
- THONGPRAJUKAEW, K. ; CHOODUM, A. ; SA-E, B ; HAYEE, U. Smart phone: A popular device supports amylase activity assay in fisheries research. **Food chemistry**, v. 163, p. 87–91, 2014.
- TIBOLA, C. S. ; SILVA, S. A. Da ; DOSSA, A. A. ; PATRÍCIO, D. I. Economically motivated food fraud and adulteration in Brazil: Incidents and alternatives to minimize occurrence. **Journal of food science**, v. 83, n. 8, p. 2028–2038, 2018.
- VOGADO, L.; VERAS, R.; ARAUJO, F.; SILVA, R.; AIRES, K. **Rede Neural Convolucional para o Diagnóstico de Leucemia**. Anais do Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2019). **Anais...Sociedade Brasileira de Computação - SBC**, 2019.

WATZLAWICK, L. F.; KIRCHNER, F. F.; DE AZAMBUJA MADRUGA, P. R. Avaliação de classificação digital em imagens de vídeo multiespectral utilizando matriz confusão. **Revista Ciências Exatas e Naturais**, v. 5, n. 1, p. 47–57, 2003.

YAMADA, P. L. F. de; AQUINO, C. I. de; SILVA, A. M. da; MARCIANO, M. A. M.; MATTOS, E. C. de; NOGUEIRA, M. D. Estudo microscópico de paprica (*Capsicum annuum* L.): deteco de fraudes e materias estranhas. **Vigilncia Sanitria em Debate: Sociedade, Cincia & Tecnologia**, v. 9, n. 1, p. 123–128, 2021.

ZARGHAMPOUR, F. ; YAMINI, Y. ; BAHARFAR, M. ; JAVADIAN, G. ; FARAJI, M. On-chip electromembrane extraction followed by sensitive digital image-based colorimetry for determination of trace amounts of Cr(vi). **Analytical methods: advancing methods and applications**, v. 12, n. 4, p. 483–490, 2020.