

Aplicação do Algoritmo KNN para Classificação de Imagens de Cristalização de Açúcar

Felipe Akira Kimura Gama
 Curso de Engenharia Elétrica
 Universidade Federal de São Carlos,
 Brasil
 felipegama@estudante.ufscar.br

Celso Ap.de França
 Departamento de Engenharia Elétrica
 Universidade Federal de São Carlos,
 Brasil
 celsofr@ufscar.br

Resumo – O aumento da competitividade mundial no ramo sucroalcooleiro impõe desafios às empresas do setor, demandando a modernização e controle preciso dos processos industriais na fabricação de açúcar. A automação e o uso de inteligência artificial permitem a integração de análise e controle de variáveis do processo. Nesse contexto, a visão computacional se mostra uma ferramenta para auxílio de procedimentos. Com o uso de algoritmos de aprendizado de máquina, classificações visuais podem ser realizadas automaticamente. Esse trabalho aborda a aplicação do algoritmo k -vizinhos mais próximos para a classificação de zonas de supersaturação a partir de imagens de cristais de açúcar. A automatização dessa tarefa contribui ao controle de qualidade e à diminuição do erro humano, promovendo avanço desse setor de grande importância econômica e social. A metodologia utilizou-se de técnicas de processamento de imagens, como equalização de histograma e filtro mediana. Foram aplicados métodos de detecção de bordas e operações morfológicas para destaque de atributos. As características extraídas foram aplicadas ao algoritmo k -vizinhos mais próximos, que foi treinado e validado com validação cruzada e diferentes valores de k . O modelo atingiu uma acurácia máxima de 83,1% com o valor 5 para k , mostrando um bom desempenho na distinção das diferentes zonas de cristalização. Apesar disso, houve sobreposição de amostras nas regiões limítrofes das classes, o que sugere a necessidade de novas abordagens para a separação entre as zonas.

Palavras-chave: classificação de imagens; cristalização do açúcar; k -vizinhos mais próximos; visão computacional.

1. INTRODUÇÃO

O ramo sucroalcooleiro ocupa um papel de destaque na economia global, sendo o Brasil o maior produtor de açúcar e etanol com cerca de 44,44% do mercado mundial de cana em 2022 [1]. A indústria nacional tem uma posição singular pois, além da sua capacidade alta de produção, possui uma diversificação expressiva do uso da cana de açúcar, como a produção de biocombustíveis e outros derivados [2]. Entretanto, com o aumento da competitividade global, o mercado impõe desafios às empresas do setor, as quais precisam adotar inovações tecnológicas para manter sua eficiência e competitividade. Em resposta a essa demanda, a modernização e o controle preciso dos processos industriais são necessários para a redução de custos e aumento de produção.

Nesse contexto, a automação e o uso de inteligência artificial (IA) surgem como soluções para o controle de produção e qualidade, contribuindo para a melhoria da precisão e eficiência dos processos. A automação possibilita o monitoramento constante de variáveis críticas que afetam

diretamente a qualidade do produto final. A união da automação com IA permite, por exemplo, analisar grandes volumes de dados em tempo real coletados de vários sensores [3].

A visão computacional tem se mostrado uma boa ferramenta na automação de tarefas tradicionalmente realizadas por humanos [4]. Ao emular a capacidade de realizar interpretações de imagens, essa tecnologia permite a análise de aspectos visuais de produtos [5]. Uma possível aplicação no setor sucroalcooleiro é a identificação das zonas de cristalização do açúcar durante o processo de cristalização. Com o uso de algoritmos de aprendizado de máquina, como o k -vizinhos mais próximos (KNN), essa classificação pode ser automatizada com base nas características específicas de imagens dos cristais.

O algoritmo KNN é utilizado em classificações supervisionadas devido à sua eficácia e simplicidade de implementação em diversas aplicações, incluindo a classificação de imagens [6]. Com base em suas características especificadas, o KNN classifica uma amostra de acordo com a sua vizinhança [7]. Através de imagens capturadas durante a operação fabril, ele é capaz de categorizá-las a fim de contribuir com o monitoramento do processo e a tomada de decisão automática.

O objetivo desse trabalho foi desenvolver um modelo de classificação baseado no algoritmo KNN para diferenciar as zonas da cristalização do açúcar através de imagens. Considerando o desenvolvimento do trabalho, foram estabelecidos os objetivos específicos: estudar as características da etapa de cristalização para entender e interpretar as imagens do banco de dados; implementar técnicas de processamento de imagem, como equalização de histograma e remoção de ruído, para ressaltar informações específicas das imagens de entrada; utilizar técnicas que permitam detectar bordas e fundos; validar o modelo KNN utilizando acurácia e validação cruzada, assegurando a eficácia na classificação; verificar a influência do número de vizinhos no desempenho do algoritmo KNN na tarefa de classificação; e desenvolver algoritmos de visualização dos dados para facilitar a compreensão das etapas realizadas.

A relevância do tópico se dá no papel da indústria açucareira na economia e na sociedade. O açúcar possui importância estratégica porque sua produção envolve amplos setores da cadeia agroindustrial, gerando emprego e renda para milhões de pessoas [8]. Otimizar o processo de produção de açúcar pode diminuir custos, aumentar a eficiência e melhorar a competitividade no mercado global. Revitalizar o setor por meio da utilização de novas tecnologias corresponde a políticas de inovação e sustentabilidade e, simultaneamente, atende ao requisito sobre

aumento de eficiência e controle de qualidade na produção de alimentos.

O interesse pelo tema decorreu da experiência adquirida em estágio superior na área de elétrica, instrumentação e automação na empresa Raízen, onde foi possível perceber a importância da automação no controle de processos industriais. Durante o estágio, observou-se que os processos envolvidos na fabricação do açúcar possuem métodos tradicionais de monitoramento limitados. A aplicação de técnicas de IA, como o algoritmo KNN, visa contribuir para o avanço tecnológico da indústria sucroalcooleira, automatizando etapas que tradicionalmente dependem de avaliação visual humana.

2. TRABALHOS RELACIONADOS

Os estudos apresentados na seção de trabalhos relacionados aplicaram técnicas de aprendizado de máquina à classificação de imagens em diversos contextos. Ao demonstrar como diferentes algoritmos, como redes neurais convolucionais (CNN) e o algoritmo KNN, podem ser utilizados para a análise de imagens, esses trabalhos ofereceram uma base importante para o estudo proposto.

No trabalho de classificação de imagens de cristais de açúcar com aprendizado profundo [9], foram investigados modelos de aprendizado profundo para classificar imagens de cristais de açúcar. Para isso, foi criada uma base de dados de imagens de cristais em diferentes zonas de cristalização. A partir dessa base, foram implementadas quatro modelos de redes neurais convolucionais (CNN). A primeira foi uma CNN simples, com filtros de dimensões 3x3 e 15 épocas. As outras três foram aplicações de aprendizado por transferência, técnica que reaproveita outros modelos pré-treinados. Os modelos implementados foram *Inception*, *ResNet50* e *VGG16*. Os resultados obtidos estão presentes na Tabela 1.

Tabela 1 – Resultados da acurácia dos quatro modelos aplicados no trabalho de classificação de cristais

Modelo	Acurácia
CNN Simples	50,5%
Inception-v3-tf	78,5%
ResNet50-v2-tf	77,5%
VGG16-tf	85,0%

No trabalho de análise do melhor algoritmo de aprendizado de máquina para a classificação de doenças de plantas [7], os desempenhos de diferentes técnicas foram avaliados na identificação de fitopatologias por meio de imagens, sendo KNN uma delas. A partir dos atributos cor, textura e formato, e acurácia como métrica, a ferramenta apresentou acurácia média de 82%.

3. FUNDAMENTAÇÃO TEÓRICA

Nesta seção, é apresentado o processo de produção de açúcar e os conceitos relevantes envolvidos no processo de cristalização. Além disso, são explorados aspectos do processamento digital de imagens e aplicações de técnicas para manipulação e extração de características. Por fim, é

apresentado o funcionamento do algoritmo KNN, a distância euclidiana e introduzidas as técnicas de normalização e validação cruzada.

3.1 O Processo da Cana e Cristalização

Segundo [10], o processo de produção do açúcar em usinas é realizado em oito etapas:

- **Recepção da matéria-prima:** descarregamento da cana diretamente na mesa alimentadora da moenda ou em pátios de armazenagem.
- **Moagem:** a cana é lavada por limpeza a seco e é conduzida em esteiras rolantes que possuem picadores e desfibradores até os ternos em que ocorre o esmagamento da matéria-prima por rolos, resultando no caldo.
- **Tratamento do caldo:** etapa em que ocorre o aquecimento do caldo, tratamentos químicos, decantação e peneiramento para a limpeza da matéria resultando em um líquido transparente.
- **Evaporação do caldo:** A água do líquido resultante do tratamento passa por evaporação, transformando-o em um xarope grosso e amarelado.
- **Cozimento e cristalização:** o xarope passa por um processo de cozimento feito em cristalizadores para concentrá-lo e formar os cristais de açúcar e melaço.
- **Centrifugação:** o melaço é separado dos cristais formados pela ação da força centrífuga e o açúcar formado é lavado com vapor.
- **Secagem:** redução da umidade para que a deterioração do açúcar não seja acelerada.
- **Armazenamento:** estocagem do açúcar produzido em sacas.

A manipulação das características físicas e químicas do processo culmina em tipos de açúcares diferentes e com padrões distintos de qualidade. Assim, o processo exige constantes interferências rápidas dos profissionais para que o resultado final seja o desejado.

Define-se cristalização como a conversão de uma substância de seu estado original para o estado cristalino [11]. Na produção de açúcar, essa transformação acontece devido ao aumento da concentração do meio. A cristalização acontece ao evaporar a água, diminuindo a quantidade de solvente com a qual a molécula de sacarose está ligada por pontes de hidrogênio. Quando não há moléculas de água suficiente para essas ligações, as próprias moléculas de sacarose podem realizar essa ligação entre si. Esse agrupamento de moléculas de sacarose é denominado núcleo cristalino e o processo de formação é chamado de nucleação [12].

Uma solução cuja quantidade de soluto é menor que a quantidade máxima de soluto que pode ser dissolvido no solvente é denominada insaturada. Se a quantidade de soluto for igual à quantidade máxima possível de dissolução, a solução se chama saturada. Já quando a quantidade de soluto estiver acima dessa quantidade máxima, a solução será denominada como supersaturada.

O coeficiente de solubilidade ($q_{saturada,pura}$) é a quantidade de massa de sacarose ($m_{sacarose}$) presente numa solução saturada em água pura ($m_{água}$) em determinada temperatura. O cálculo do coeficiente de solubilidade está presente na Equação (1).

$$q_{saturada,pura} = \left(\frac{m_{sacarose}}{m_{\text{água}}} \right)_{saturada,pura} \quad (1)$$

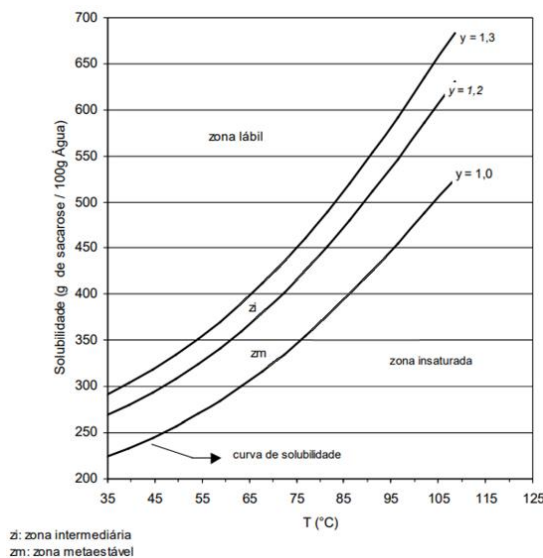
O coeficiente de supersaturação (y_i) de uma solução impura de sacarose é a fração do coeficiente de solubilidade da solução supersaturada em relação ao coeficiente de solubilidade da solução saturada em água pura. A partir do coeficiente de solubilidade de uma solução supersaturada impura (q_i), o cálculo do coeficiente de supersaturação é apresentado na Equação (2).

$$y_i = \frac{q_i}{q_{saturada,pura}} \quad (2)$$

Curva de solubilidade é o gráfico que relaciona a variação do coeficiente de solubilidade em função da temperatura. Com as curvas de solubilidade para saturação ($y_i = 1$) e supersaturação ($y_i = 1,2$ e $y_i = 1,3$) para o processo de cristalização de sacarose numa solução impura, apresentadas na Figura 1, tem-se quatro regiões de cristalização distintas:

- Zona Insaturada: região na qual há excesso de água, o qual dissolve os cristais existentes.
- Zona Metaestável: região em que a solução ainda possui uma quantidade presente de água que impede a estabilização de novos núcleos, mas ideal para a aderência de moléculas de sacarose a cristais já formados (crescimento do cristal).
- Zona Intermediária: região na qual há o crescimento de cristais pré-formados e novos núcleos são formados.
- Zona Lábil: região com formação de cristais novos e indesejados por conta da nucleação espontânea e acentuada.

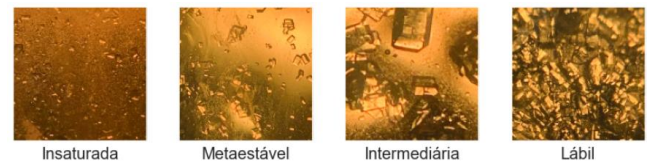
Figura 1 – Curvas de solubilidade da sacarose



Fonte: Adaptado de [13]

Na Figura 2, apresentam-se imagens de cristais de açúcar de cada zona. A partir da identificação da zona, são empregadas técnicas específicas de aquecimento e resfriamento para destinar a solução à zona desejada.

Figura 2 – Cristais de açúcar em zonas distintas



Fonte: Elaborado pelo autor

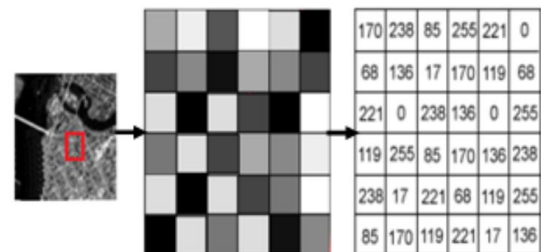
3.2 Imagem Digital

3.2.1 Estrutura da Informação

Uma imagem digital é composta por um conjunto de pontos chamados *picture elements* (pixels), que é a menor unidade visível dos dados. Duas características relevantes de uma imagem digital são suas dimensões e resolução. A dimensão é a quantidade de pixels presente no arquivo, normalmente expressa por altura e largura. Já a resolução se refere a quantidade de pixels presente em cada unidade de medida da imagem, sendo pixels por polegada (PPI) a medida aplicada a imagens digitais [14].

Cada pixel apresenta uma localização definida e uma intensidade luminosa codificada em binário. A junção de vários bits possibilita que diferentes graus de intensidade sejam representados no pixel. O tamanho comum é de oito bits, que corresponde a 2^8 ou 256 possibilidades. Na Figura 3, observa-se uma região da imagem original segmentada por pixels e os valores de intensidade luminosa para cada unidade da região específica na representação matricial.

Figura 3 – Pixels e valores de intensidade luminosa

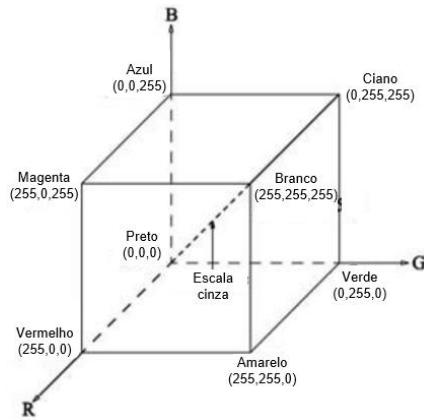


Fonte: [15]

3.2.2 Modelo de cores RGB

O modelo *Red, Green and Blue* (RGB) é formado pelos canais Vermelho (R), Verde (V) e Azul (B). A partir da soma desses componentes, formam-se as outras cores do espectro, conforme mostrado na Figura 4. Se todas as três cores estiverem com seu valor máximo, o resultado será a cor branca. Na Figura 4, percebe-se a componente diagonal formada pela reta que liga a cor branca e preta, na qual estão os valores dos níveis de cinza.

Figura 4 – Cubo RGB



Fonte: Adaptado de [16]

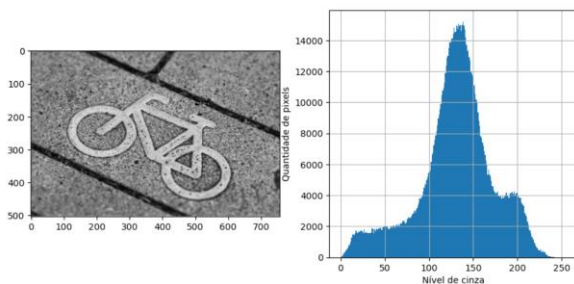
3.2.3 Histograma

O histograma de uma imagem digital é definido como a contagem de pixels de cada nível de cinza [17]. De forma probabilística, considerando uma imagem com níveis de cinza, p a distribuição de probabilidade de um nível de cinza específico numa imagem (p), o k -ésimo nível de cinza (r_k), o número de pixels na imagem com o respectivo nível de cinza (n_k) e o número de pixels da imagem (n), o histograma é definido pela função discreta apresentada na Equação (3) [17].

$$p(r_k) = \frac{n_k}{n} \quad (3)$$

Na Figura 5, tem-se um exemplo de uma imagem em escala cinza e seu histograma com a contagem de pixels para cada nível de cinza.

Figura 5 – Imagem e seu histograma

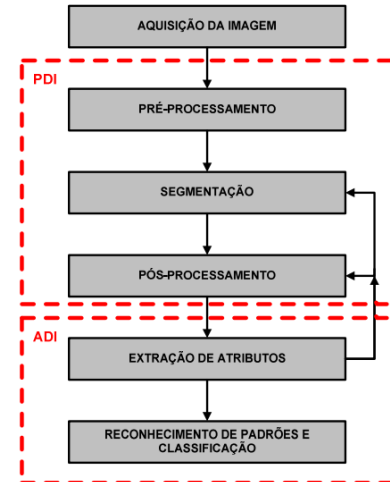


Fonte: Elaborado pelo autor

3.2.4 Processamento e Análise Digital de Imagens

O processamento de imagens consiste na adequação de características da imagem, através de operações matemáticas que alteram os valores dos pixels, para melhorar sua visualização e adaptação para extração, tratamento de dados quantitativo e avaliação [18]. A análise se refere à observação quantitativa dos dados através das identificações realizadas na imagem. Na Figura 6, tem-se o fluxograma da sequência padrão, no qual PDI corresponde ao Processamento Digital de Imagens e ADI a Análise Digital de Imagens.

Figura 6 – Fluxograma padrão de processamento e análise de imagens



Fonte: Adaptado de [19]

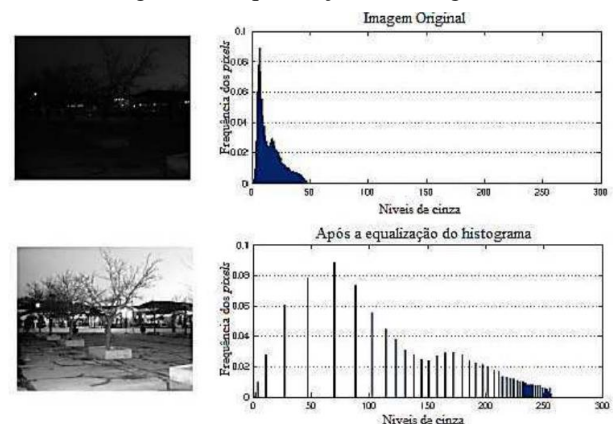
O procedimento da Figura 6 indica que, após o pré-processamento aplicado à todas as imagens para facilitar suas análises subsequentes, a segmentação, pós-processamento e extração de atributos podem ser repetidas e retiradas a depender da necessidade de aplicação. Por fim, a partir dos atributos extraídos, o reconhecimento e classificação dos padrões assimilados são realizados por dados quantitativos.

3.2.5 Equalização de Histogramas

A equalização de histograma gera uma imagem com a distribuição dos níveis de cinza de forma uniforme. A transformação realça as diferenças sutis entre níveis próximos de cinza, o que aumenta a qualidade de detalhes perceptíveis [17].

A Figura 7 exemplifica as mudanças tanto na imagem quanto no histograma ao realizar a equalização.

Figura 7 – Equalização de histograma



Fonte: Adaptado de [20]

3.2.6 Filtros

Afim de melhorar o desempenho do processo de segmentação, técnicas de filtragem modificam a imagem pixel a pixel levando em consideração os pixels vizinhos. A partir de uma máscara de

tamanho variável, são utilizadas funções matemáticas aplicadas para modificar o pixel central.

3.2.6.1 Filtro Mediana

Utilizado para suavização de ruídos do tipo impulsivo, o filtro da mediana produz uma imagem transformada de mesma dimensão que a imagem de entrada da transformação definindo o valor do pixel central como a mediana entre o próprio centro e seus vizinhos. Os vizinhos a serem considerados são definidos pela matriz estruturante [21]. A Figura 8 exemplifica o funcionamento do filtro mediana 3x3 aplicada sobre uma matriz 5x5. Um pixel central da matriz original na Figura 8(a) está pintado com fundo claro e seus respectivos vizinhos a serem considerados estão dentro da região delimitada pela área com contorno de mesma cor. Na Figura 8(b), tem-se a matriz resultante da operação realizada pela transformação, na qual o pixel está destacado de acordo com a cor do exemplo indicado na matriz original. Quando a vizinhança não está bem definida, usa-se zero como resultado.

Figura 8 – Exemplo de funcionamento do filtro mediana

9	4	5	0	8	0	0	0	0	0
10	3	2	1	7	0	5	3	5	0
9	1	6	3	15	0	3	3	6	0
0	3	8	10	1	0	6	8	8	0
1	16	9	12	7	0	0	0	0	0

(a)

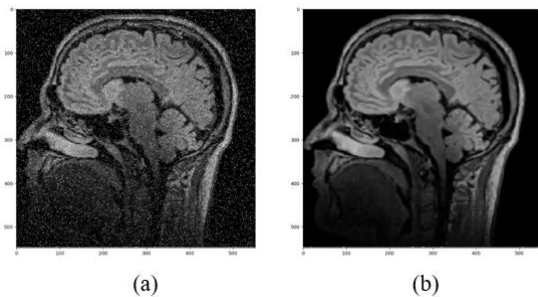
0	0	0	0	0	0	0	0	0	0
0	5	3	5	0	0	5	3	5	0
0	3	3	6	0	0	3	3	6	0
0	6	8	8	0	0	6	8	8	0
0	0	0	0	0	0	0	0	0	0

(b)

Fonte: Elaborado pelo autor

Na Figura 9, tem-se o antes e depois da aplicação do filtro mediana 5x5, no qual a Figura 9(a) é a imagem original e a Figura 9(b) a matriz transformada.

Figura 9 – Exemplo de aplicação de filtro mediana



Fonte: Elaborado pelo autor

3.2.6.2 Erosão, Dilatação e Gradiente Morfológico

A morfologia matemática compreende o estudo e análise de imagens usando operadores não lineares [22]. A partir de um elemento estruturante pré-definido, a transformação acontece baseando-se em dois operadores básicos: erosão e dilatação.

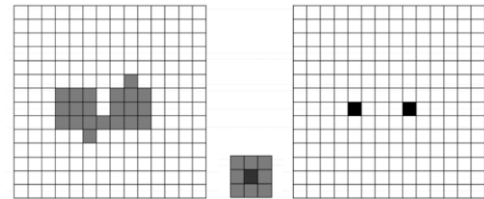
A erosão (ε) de uma imagem (f) de elementos (x) por um elemento estruturante (B) de elementos (b) é definida como a Equação (4) [23]. O resultado da operação em imagens de escala cinza é a redução dos padrões claros e alargamento das

regiões escuras.

$$[\varepsilon_B(f)](x) = \min\{f(x + b) - B(b)\}, b \in B \quad (4)$$

Considerando imagens binarizadas, os padrões assimilados são reduzidos. A Figura 10 exemplifica o funcionamento da erosão, sendo os padrões assimilados em cinza, e centro do elemento estruturante e correspondência na imagem final em preto.

Figura 10 – Imagem original, elemento estruturante e imagem resultante no funcionamento de erosão



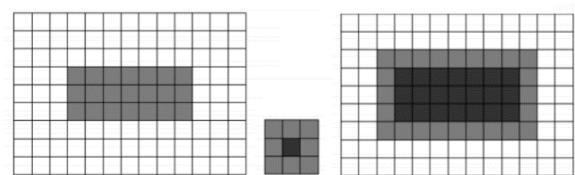
Fonte: Adaptado de [24]

Já a dilatação (δ) de uma imagem (f) de elementos (x) por um elemento estruturante (B) de elementos (b) é definida como a equação (5) [23]. O resultado da operação em imagens de escala cinza é a redução dos padrões escuros e alargamento das regiões claras.

$$[\delta_B(f)](x) = \max\{f(x + b)\}, b \in B \quad (5)$$

Considerando imagens binarizadas, os padrões assimilados são aumentados. A Figura 11 exemplifica o funcionamento da dilatação, sendo os padrões assimilados em cinza, e centro do elemento estruturante e correspondência na imagem final em preto.

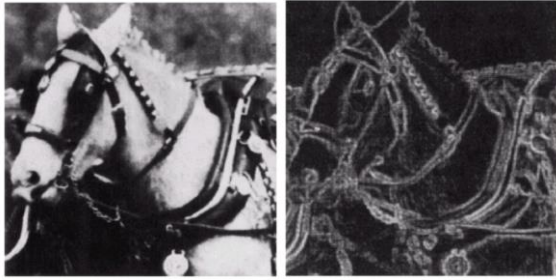
Figura 11 – Imagem original, elemento estruturante e imagem resultante no funcionamento de dilatação



Fonte: Adaptado de [24]

Por fim, o gradiente morfológico é a diferença aritmética entre a dilatação e a erosão de uma imagem por um mesmo elemento estruturante. A imagem resultante enfatiza as transições e são marcadas em níveis de cinza, sendo um bom filtro para obter contornos [23]. Na Figura 12, tem-se um exemplo da aplicação de gradiente morfológico.

Figura 12 – Exemplo de aplicação de gradiente morfológico



Fonte: Adaptado de [24]

3.2.6.3 Detector Canny de Bordas

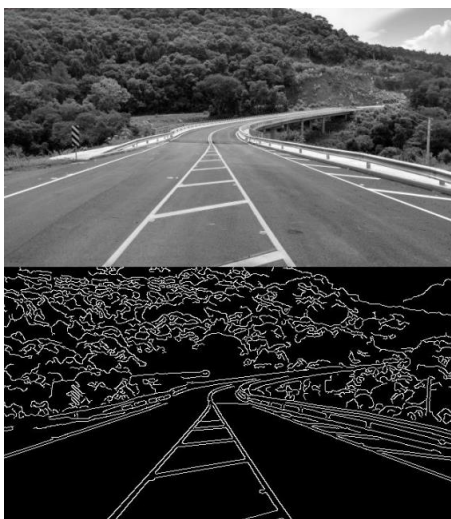
O detector de bordas proposto por Canny em 1986 é uma das primeiras técnicas criadas na área de visão computacional e continua sendo utilizado atualmente. Ele é considerado uma referência padrão nos estudos comparativos de bordas [25].

Canny definiu três requisitos para o detector: ele deve evitar a detecção de bordas falsas; bordas identificadas devem estar próximas de bordas reais; e o algoritmo deve fornecer uma resposta clara e precisa. Para isso, Canny introduziu a supressão de não-máximos e a limiarização dupla para a seleção.

A supressão de não-máximos elimina os pontos que não são máximos locais em relação ao gradiente e, portanto, não representam o pico de uma borda, mantendo apenas pontos fortes. Já a limiarização dupla utiliza dois valores de limiar: um alto para detectar bordas fortes e outro baixo para identificar bordas mais fracas que estão conectadas às bordas fortes. Esses dois limiares são ajustados experimentalmente para cada conjunto de imagem e permitem a diferenciação de bordas reais de ruídos e variações sutis das intensidades de cinza. Assim, o modelo usa o gradiente da imagem para destacar regiões com derivadas espaciais altas e rastreia essas regiões para deletar pixels que não estão em sua intensidade máxima [26].

A Figura 13 apresenta um exemplo da aplicação do detector de bordas formalizado por Canny.

Figura 13 – Exemplo de aplicação do detector de bordas de Canny



Fonte: Adaptado de [27]

3.2.6.4 Tabela de Consulta

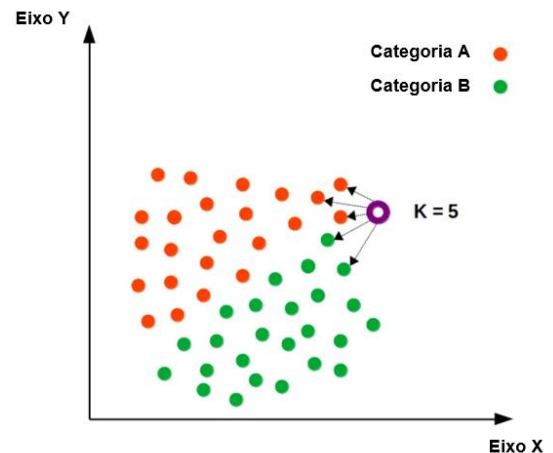
Uma tabela de consulta (LUT, *lookup table*) é uma lista de chaves e seus respectivos valores pré-calculados atribuídos. Em imagens digitais, a partir do valor original do pixel, é definido um novo valor pela LUT. Ela é usada principalmente para ajuste de contraste ou intensificação de características [28].

3.3 Algoritmo KNN

O KNN é um algoritmo de aprendizado de máquina supervisionado no qual os objetos são categorizados baseado na distância entre a mostra de teste e as amostras de treinamento no espaço de características [29].

A partir do vetor de características de entrada, o algoritmo calcula cada distância entre a amostra a ser predita e cada amostra de treino. A categoria da nova amostra é predita a partir da categoria dominante entre os k vizinhos mais próximos. São fornecidos como parâmetros ao algoritmo o número k de vizinhos a serem considerados e a métrica utilizada para o cálculo de distância. Na Figura 14, observa-se um exemplo do funcionamento do algoritmo. Considerando dados de características de duas dimensões, $k = 5$ e as categorias A e B para a classificação das amostras, o sistema predirá que a categoria da nova amostra é Categoria A pois, entre os 5 vizinhos mais próximos, prevalecem os pontos vermelhos.

Figura 14 – Exemplo de funcionamento do algoritmo KNN



Fonte: Adaptado de [30]

O valor de k escolhido interfere na eficiência do algoritmo. Ele deve ser grande o suficiente para diminuir a sensibilidade a ruído e pequeno o suficiente para que o número de vizinhos a serem considerados não ultrapasse o domínio de outras classes [31].

A métrica de proximidade mais utilizada no algoritmo KNN é a Distância Euclidiana [32], que calcula a distância entre dois pontos em um espaço euclidiano de qualquer dimensão. Essa distância (d) entre dois vetores (x e x_i) é definida pela Equação (6). Na Equação (7), tem-se um exemplo da distância euclidiana calculada entre dois vetores de três dimensões $A = (x_1, y_1, z_1)$ e $B = (x_2, y_2, z_2)$.

$$d(x_i, x) = |x_i - x| \quad (6)$$

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (7)$$

Como as características podem ter domínios de variação diferentes, é considerada uma boa prática a realização de normalização antes da entrada dos vetores em algoritmos de treinamento, como o KNN. Para dados de muitas dimensões, a normalização por *MinMaxScaler* é vantajosa [33]. O valor de x normalizado ($x_{normalizado}$) é obtido pela Equação (8) a partir do valor a ser normalizado (x), do valor máximo (x_{max}) do item no vetor e do valor mínimo (x_{min}) do item no vetor [34].

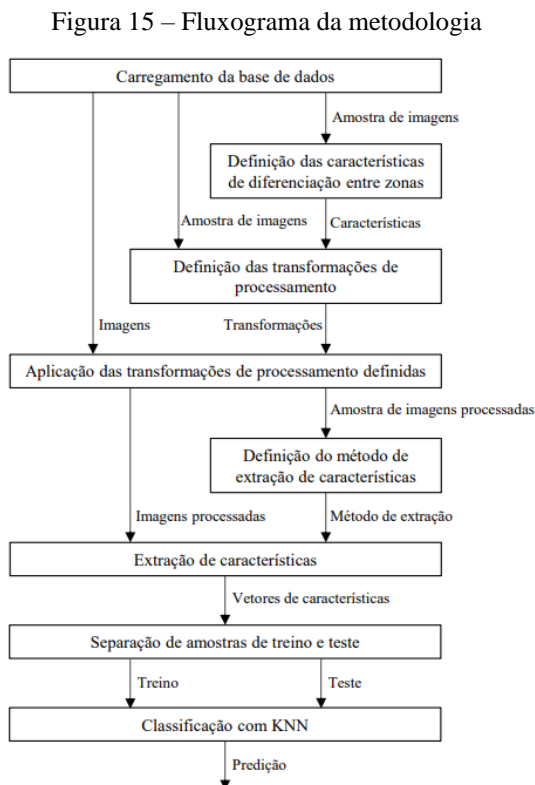
$$x_{normalizado} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (8)$$

Para avaliar a generalização de um modelo de aprendizado de máquina, utiliza-se a técnica de validação cruzada. Ela consiste em dividir as amostras em partes iguais e, em cada iteração, uma parte diferente é utilizada como teste e as outras como treino [35]. O processo é repetido até que todas as partes tenham sido usadas como teste.

A fim de avaliar o modelo, empregam-se métricas como a acurácia. Ela é calculada a partir da divisão do número de previsões corretas pelo total de previsões realizadas [36].

4. METODOLOGIA

A Figura 15 apresenta o fluxograma da metodologia utilizada.



Fonte: Elaborado pelo autor

A partir das imagens carregadas da base de dados, realizaram-se testes de transformações nas amostras para a obtenção de características desejadas. Determinado o processamento completo, aplicaram-se as transformações definidas em toda a base de dados. Com as imagens processadas, avaliou-se a métrica adequada para a extração das características obtidas. Definida a métrica, extraíram-se os vetores de características e foram atribuídos ao algoritmo KNN. Em seguida, definiram-se os parâmetros do modelo e realizou-se a predição das amostras.

Toda a metodologia foi realizada com a linguagem *Python* no *Jupyter Notebook*. Para as técnicas de aprendizado de máquina, foi utilizada a biblioteca *Scikit-learn*.

4.1 Carregamento da Base de Dados

A base de dados utilizada é o *Dataset of Sugar Crystal Thailand* [37], desenvolvida e usada no estudo de aprendizado profundo aplicado na classificação de cristais de açúcar [9]. Ela é composta por 1000 imagens coloridas de tamanho 1000x1000 pixels e pré-separadas entre pastas que diferem as classes e reserva 20% de cada zona como amostras de teste, conforme a Tabela 2.

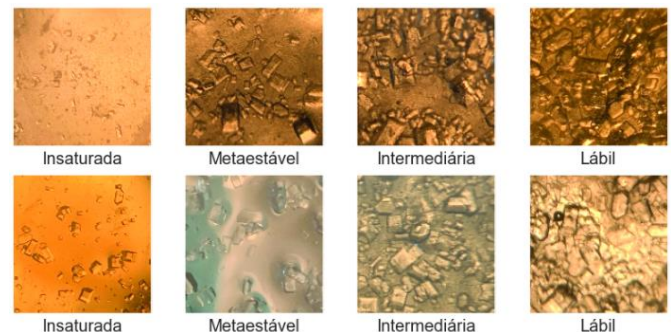
Tabela 2 – Separação de amostras na base de dados

Classe	Treino	Teste
Insaturada	200	50
Metaestável	200	50
Intermediária	200	50
Lábil	200	50

Apesar da pré-separação entre amostras de treino e teste, as pastas foram unificadas para a execução de testes com diferentes conjuntos de amostras inicialmente.

Na Figura 16, encontram-se alguns exemplos das imagens da base de dados.

Figura 16 – Exemplos de imagens da base de dados



Fonte: Elaborado pelo autor

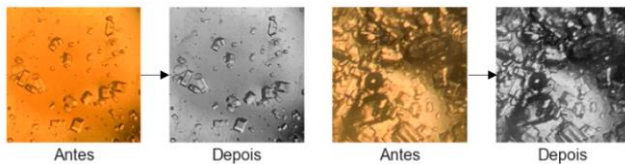
Ao realizar o carregamento da base unificada, o único pré-processamento aplicado foi o redimensionamento das imagens para 150x150 pixels.

4.2 Definição das Transformações de Processamento

Para a definição das transformações de processamento de imagens, escolheram-se duas características que diferem os cristais de uma zona à outra: quantidade de bordas presentes e quantidade de fundo visível.

Quanto à quantidade de bordas presentes, optou-se por manter as características totais das imagens coloridas como mostrado na Figura 16. Assim, as imagens foram convertidas para escala cinza para os dados ficarem com apenas um canal. A Figura 17 apresenta dois exemplos do resultado obtido dessa transformação.

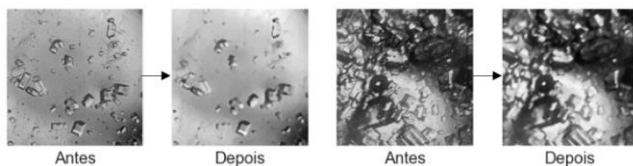
Figura 17 – Aplicação da conversão para escala cinza



Fonte: Elaborado pelo autor

Em seguida, escolheu-se aplicar o detector Canny para a identificação de bordas. Para suavizar a imagem e evitar a detecção de bordas falsas, aplicou-se um filtro mediana 5x5 antes de empregar o detector. A Figura 18 apresenta dois exemplos de resultados obtidos.

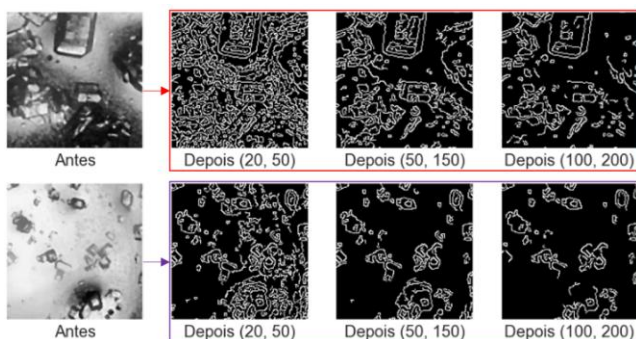
Figura 18 – Aplicação do filtro mediana



Fonte: Elaborado pelo autor

Foram testados diferentes parâmetros no detector e, como foi observado muito ruído, optou-se por utilizar valores elevados para priorizar bordas bem definidas. A Figura 19 apresenta dois exemplos para os limiares (20, 50), (50, 150) e (100, 200), dentre os quais o terceiro foi o escolhido.

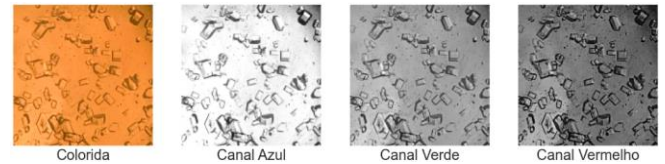
Figura 19 – Aplicação de diferentes limiares para o detector de bordas Canny



Fonte: Elaborado pelo autor

Como a abordagem realizada priorizou bordas bem definidas, decidiu-se por adotar mais uma medida para bordas, utilizando o gradiente morfológico. Para isso, foram separados os canais RGB da imagem original a fim de se verificar se um dos canais apresenta melhor visualização das bordas. A Figura 20 apresenta uma amostra dividida pelos canais, na qual o canal azul apresentou melhora de visualização.

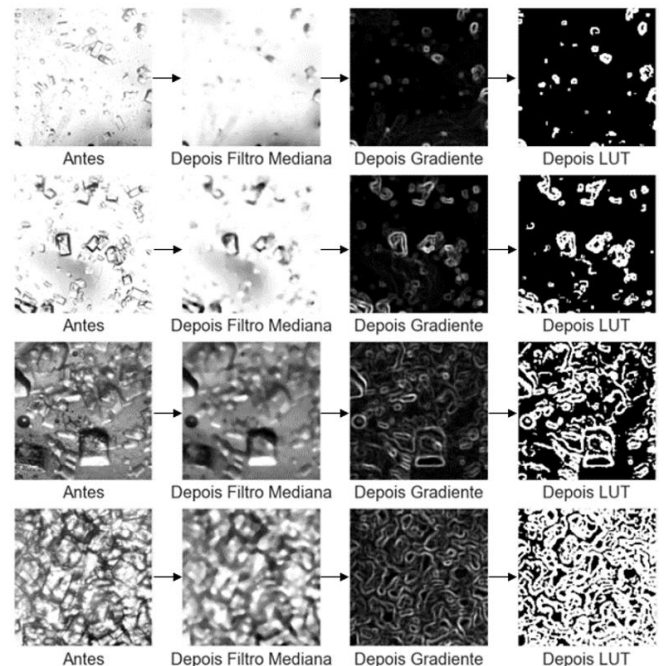
Figura 20 – Amostra dividida entre os canais RGB



Fonte: Elaborado pelo autor

Em seguida, aplicou-se um filtro mediana 5x5 seguida da transformação por gradiente morfológico, cujo elemento estruturante utilizado foi uma matriz convolucional 3x3 quadrada unitária. Às respostas obtidas, utilizou-se uma LUT para destacar arestas cinzas claras, cuja intensidade de divisão foi 20. Então, caso a intensidade do pixel seja menor que 20, a saída será nula; caso contrário, será máxima. Exemplos dos resultados desses passos são observado na Figura 21.

Figura 21 – Aplicação de filtro mediana, gradiente morfológico e LUT



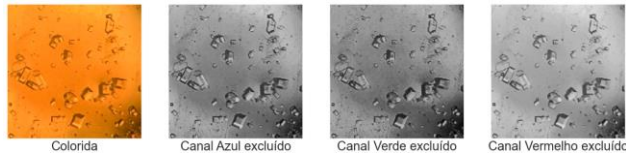
Fonte: Elaborado pelo autor

Por fim, para a quantidade de fundo visível, optou-se pela estratégia de realizar a diferenciação entre fundo e bordas por intensidade luminosa e, por conta da translucidez dos cristais, realizar a erosão das regiões identificadas para minimizar essa propriedade.

Realizou-se a exclusão de cada canal RGB e observou-se que, com a exclusão do canal vermelho, a intensidade luminosa das

bordas eram diminuídas, o contraste entre bordas e sombras era melhorado e as características de fundo eram mantidas, mas suavizadas. A Figura 22 apresenta uma amostra com cada canal RGB excluído.

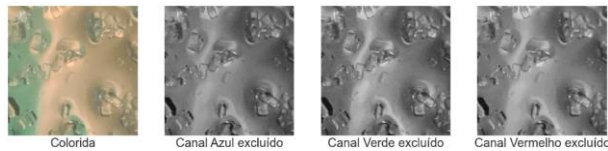
Figura 22 – Amostra com cada canal RGB excluído



Fonte: Elaborado pelo autor

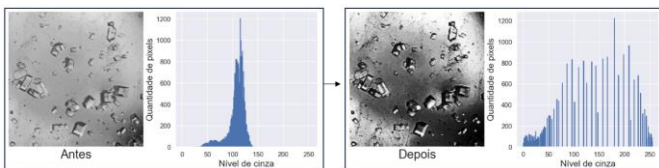
Entretanto, para o trabalho com intensidade luminosa, a transformação não foi suficiente em imagens com variação de iluminação do fundo entre grandes regiões. Um exemplo desse caso é apresentado na Figura 23. Dessa forma, optou-se por realizar a identificação de fundos claros e escuros separadamente. Assim, as imagens com o canal vermelho excluído foram convertidas para escala cinza, foi aplicado equalização para contrastar os níveis de cinza presentes e aplicado um filtro mediana 5x5. Um exemplo do resultado obtido na equalização é apresentado na Figura 24.

Figura 23 – Amostra com cada canal RGB excluído



Fonte: Elaborado pelo autor

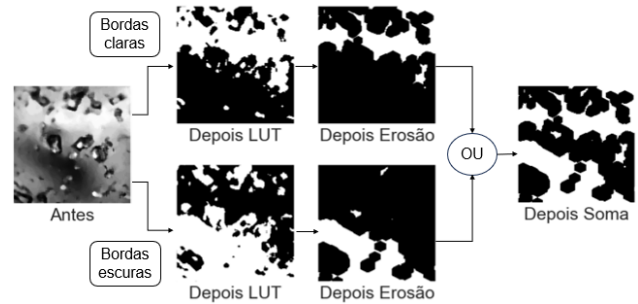
Figura 24 – Resultado de equalização



Fonte: Elaborado pelo autor

Para fundos claros, foi utilizada uma LUT com intensidade de divisão de 150. Em seguida, foram aplicadas duas iterações de erosão com elemento estruturante elíptico 5x5 para reduzir as falsas identificações de fundo devido à transparência dos cristais. Já para fundos escuros, a LUT utilizou intensidade de divisão de 130 e invertia a seleção. Para a redução de falsa identificação nos casos de sombra entre cristais e bordas concentradas, foram aplicadas três iterações de erosão de mesmo elemento estruturante utilizado aos fundos claros. Por fim, os dois resultados obtidos foram somados utilizando o operador OU bit-a-bit. A Figura 25 apresenta resultados obtidos dessas transformações a partir de uma amostra.

Figura 25 – Exemplo de processamento aplicado no reconhecimento de fundos claros e escuros



Fonte: Elaborado pelo autor

4.3 Extração de Características

Uma vez que as localizações dos pixels identificados nos resultados dos três processamentos definidos não são relevantes, a extração das características foi realizada a partir do histograma. Assim, a informação que caracterizou cada atributo foi a quantidade de pixel de cada nível de cinza.

Como as imagens resultantes são binarizadas, o vetor do histograma de cada imagem teve tamanho 256 e com valores zerados da segunda até a penúltima posição. A quantidade de pixels pretos e brancos eram armazenados na primeira e última posição respectivamente.

Para a visualização dos resultados obtidos dos três processamentos de imagens definidos, foi criado um gráfico de dispersão tridimensional com as quantidades de pixels brancos normalizadas por *MinMaxScaler*.

4.4 Classificação com KNN

Utilizando a função *StratifiedKFold* da biblioteca do *Scikit-Learn*, os vetores de características foram divididos em cinco grupos para a realização de validação cruzada. O atributo de embaralhamento foi ativado e o estado-aleatório utilizado foi 12. O embaralhamento foi utilizado pois as amostras, na pasta do banco de dados, são ordenadas por grupos visualmente semelhantes e o estado-aleatório garante replicabilidade do estudo.

Para cada grupo de amostra, os dados foram normalizados aos valores do grupo de teste utilizando a função *MinMaxScaler*. Para os testes com diferentes valores de k , foi definida a faixa de 1 à 30.

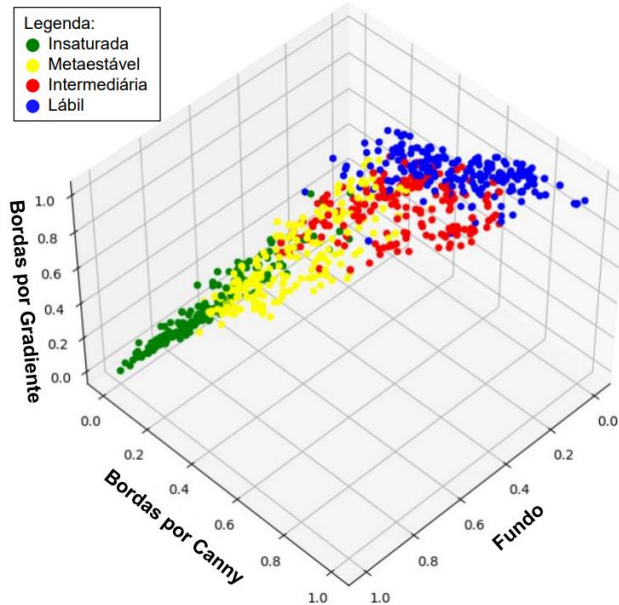
Em seguida, realizou-se a implementação do modelo KNN utilizando a métrica euclidiana, seguido do treinamento, predição dos dados de teste e obtenção da acurácia do modelo para cada valor de k e cada grupo de teste. A acurácia final do modelo para um número de vizinhos próximos considerado foi a média aritmética das acurácias obtidas com os diferentes subgrupos de teste e treinamento. Assim, definiu-se o melhor valor de k e foi obtida a matriz de confusão para a pior separação de amostras nesse valor.

Por último, foi realizada a divisão entre teste e treinamento implementado no trabalho de CNN's [9]. A partir da acurácia, foi definido o melhor valor de k para essa separação e obtida a matriz de confusão do sistema.

5. RESULTADOS E DISCUSSÕES

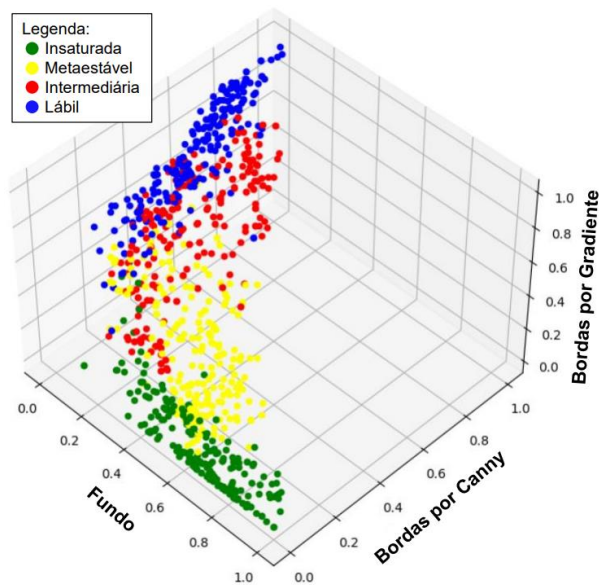
Nas Figuras 26 e 27, apresenta-se o gráfico de dispersão obtido a partir da aplicação dos processamentos definidos na metodologia por dois ângulos diferentes para todas as amostras do banco de dados.

Figura 26 – Gráfico de dispersão das amostras do banco de dados (visão 1)



Fonte: Elaborado pelo autor

Figura 27 – Gráfico de dispersão das amostras do banco de dados (visão 2)



Fonte: Elaborado pelo autor

Constatou-se que as estratégias aplicadas categorizaram as zonas de cristalização. No entanto, nas regiões de intersecção entre classes há sobreposição de amostras, o que torna o sistema mais sensível ao distinguir amostras localizadas nos limiares. Além disso, no espaço ao redor do ponto (Fundo =

0,3; Bordas por Canny=0,3; Bordas por Gradiente=0,4) tem-se uma mistura das quatro classificações, mas prevalecendo amostras das zonas intermediária e metaestável. Já na região onde há muito fundo identificado e poucas bordas, houve boa separação das amostras da zona instaurada; enquanto na região onde há pouco fundo e muitas bordas identificadas houve boa separação de amostras da zona lábil.

A acurácia final, calculada a partir da média ponderada das acurácias por grupo de teste e treinamento, é apresentada no gráfico da Figura 28.

Figura 28 – Acurácias finais para diferentes valores de k



Fonte: Elaborado pelo autor

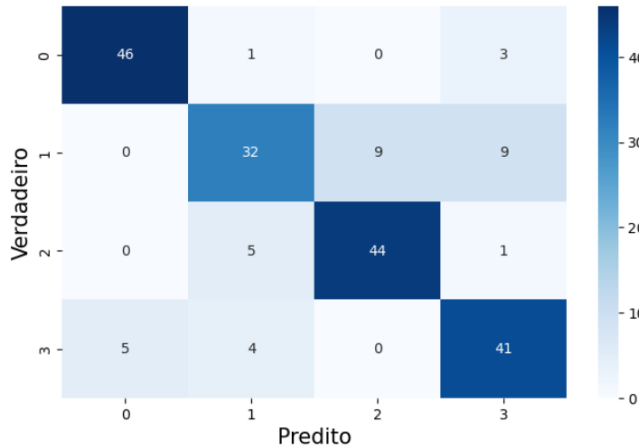
Apesar da maior acurácia final ter sido obtida com $k = 1$, optou-se por seguir com um valor diferente pois o algoritmo estaria suscetível a identificar uma amostra de treinamento atípica como o vizinho mais próximo. Além disso, estaria propenso a sobreajuste. Dessa forma, pela Figura 28, obteve-se a maior acurácia de 83,1% com $k = 5$. Os valores de acurácia de cada grupo de teste para esse número de vizinhos foram apresentados na Tabela 3, pela qual constatou-se que os grupos 3 e 4 apresentaram as piores acurácias da divisão da validação cruzada.

Tabela 3 – Acurácias de cada divisão do grupo de teste para $k = 5$

k	G1	G2	G3	G4	G5	Média
5	84,5%	83,0%	81,5%	81,5%	85,0%	83,1%

Optou-se por seguir com o grupo 3 para a realização da matriz de confusão, apresentada na Figura 29. O dicionário das classes da Figura 29 está presente na Tabela 4.

Figura 29 – Matriz de confusão do grupo 3 da validação cruzada



Fonte: Elaborado pelo autor

Tabela 4 – Dicionário da matriz de confusão

Valor	Classificação
0	Zona Insaturada
1	Zona Intermediária
2	Zona Lábil
3	Zona Metaestável

Pela Figura 29, concluiu-se que o modelo apresentou piores resultados ao prever imagens da zona intermediária, tendo confundido 36% de suas amostras igualmente entre as zonas metaestável e lábil.

Com a separação de teste e treinamento realizado em [9], implementou-se o modelo proposto e obtiveram-se os valores de acurácia para cada valor de k . Os resultados foram apresentados na Figura 30 e o valor máximo de acurácia obtido foi de 72% com $k = 5$ e $k = 6$.

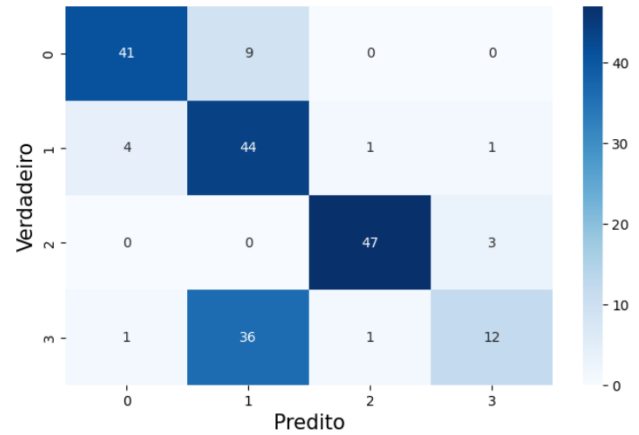
Figura 30 – Acurácia para diferentes valores de k



Fonte: Elaborado pelo autor

Para a matriz de confusão, seguiu-se com $k = 5$. O resultado obtido foi apresentado na Figura 31, na qual constatou-se que a divisão de treino e teste realizada apresentou melhora de 37,5% na predição de imagens da zona intermediária, mas uma piora de 70,7% na predição de amostras da zona metaestável. O dicionário da matriz de confusão é o mesmo encontrado na Tabela 4.

Figura 31 – Matriz de confusão da divisão realizada em [9]



Fonte: Elaborado pelo autor

A acurácia de 72% do modelo KNN proposto superou a de 50,5% do modelo simples baseado em CNN implementado em [9]. Um dos possíveis fatores para o qual a CNN obteve acurácia menor do que a KNN foi o uso de apenas máscaras 3x3, as quais são pequenas quando comparadas ao tamanho da imagem, perdendo a captura de padrões globais.

6. CONCLUSÕES

O modelo KNN proposto apresentou uma acurácia de 83,1% na utilização do estado-aleatório 12 na técnica de validação cruzada. Embora essa tenha sido a maior acurácia, com $k = 5$, todas as acurácias finais calculadas com k variando de 1 a 30 foram maiores que 80%.

Quando comparado ao modelo de CNN em [9], o modelo KNN apresentou 72% de acurácia enquanto a rede convolucional atingiu 50,5%.

Embora tenha obtido acurácia melhor que a rede convolucional simples, o modelo apresentou pontos de intersecção no gráfico de dispersão das características das amostras. Assim, sugere-se a avaliação de novos atributos para melhor diferenciação entre zonas e aplicação de outras técnicas nas definições de processamento de imagens. Por conta da translucidez dos cristais, a abordagem de identificação de fundo foi prejudicada em contextos com alta variabilidade de iluminação. Recomenda-se a implementação de técnicas avançadas de equalização, como a Equalização de Histograma Adaptativa com Limite de Contraste (CAHLE) que aplica a transformação de forma separada para cada parte da imagem.

7. REFERÊNCIAS

- [1] FAO – FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. Statistics Division, 2022. Disponível em: <http://www.fao.org/faostat/en/#data/QC>. Acesso em: 14 ago. 2024.
- [2] PLACIOS-BERECHE, M. C.; PALACIOS-BERECHE, R.; ENSINAS, A. V.; GALLEGO, A. G.; MODESTO, M.; NEBRA, S. A. Brazilian sugar cane industry – A survey on future improvements in the process energy management. Energy, v. 259, 2022.
- [3] RIBEIRO, J.; LIMA, R.; ECKHARDT, T.; PAIVA, S. Robotic Process Automation and Artificial Intelligence in Industry 4.0 – A Literature Review. Procedia Computer Science, v. 181, p. 51-58, 2021.
- [4] MARENGONI, M.; STRINGHINI, R. R. Tutorial: Introdução à Visão Computacional usando OpenCV. Revista Eletrônica de Computação e Tecnologia da Informação, v. 16, n. 1, p. 125-160, 2010.

- [5] NETO, E. P. F. Visão Computacional para Identificação de Cores em Tempo Real com OpenCV e Python. Trabalho de Conclusão de Curso (Graduação), Engenharia de Computação, Centro Universitário de Brasília, Brasília, 2020.
- [6] RICHARDS, J. A. Remote Sensing Digital Image Analysis: An Introduction, 5th ed.; Germany, 2013; p. 494.
- [7] NEELAKANTAN, P. Analyzing the best machine learning algorithm for plant disease classification. *Materials Today: Proceedings*, v. 80, p. 3668-3671, 2023.
- [8] EDUARDO, A. S.; RODRIGUES, W. O. P.; FRAINER, D. M. Setor sucroenergético do estado de Mato Grosso do Sul: as relações intersetoriais a partir da matriz insumo-produto. *Escola Superior de Sustentabilidade*, v. 13, n. 6, 2022.
- [9] CHAYATUMMAGOON, S.; CHONGSTITVATANA, P. Image classification of sugar crystal with deep learning. 13th International Conference on Knowledge and Smart Technology, Chonburi, Tailândia, 2021.
- [10] MACHADO, S. S.; Tecnologia da Fabricação do Açúcar, 1a. ed., IFG (Santa Maria), Inhumas, 2012.
- [11] COSTA, C. B. B.; GIULIETTI, M. Introdução à Cristalização: princípios e aplicações. São Carlos: EdUFSCar, 2010.
- [12] FRANCISCO, B. Cristalização da sacarose. LinkedIn, 2023. Disponível em: <https://www.linkedin.com/pulse/cristaliza%C3%A7%C3%A3o-da-sacarose-bruno-francisco/>. Acesso em: 15 ago. 2024.
- [13] MANTELATTO, P. E. Estudo do processo de cristalização de soluções impuras de sacarose de cana-de-açúcar por resfriamento. 2005. Dissertação (Mestrado em Engenharia Química) – Universidade Federal de São Carlos, Centro de Ciências Exatas e Tecnologia, São Carlos, 2005.
- [14] CARMO, C. C. Detecção e identificação de doenças em folhas utilizando redes neurais. [2023]. Trabalho de Conclusão de Curso (Graduação em Engenharia Elétrica) – Centro de Ciências Exatas e de Tecnologia, Universidade Federal de São Carlos, São Carlos, 2023.
- [15] GÜLGÜN, O. D.; EROL, H. Classification performance comparisons of deep learning models in pneumonia diagnosis using chest X-ray images. *Turkish Journal of Engineering*, v. 4, n. 3, 2019.
- [16] CORONADO, M.; SANTOS, A. G. The influence of water repellent products on the chromatic modifications of the ceramic brick. 8th International Masonry Conference. Dresde, Alemanha, 2010.
- [17] OLIVEIRA, R. G. D. Equalização e segmentação de imagens termográficas aplicadas à metodologia de inspeção de isoladores poliméricos. Dissertação (Mestrado em Engenharia Elétrica) – Centro de Engenharia Elétrica e Informática, Universidade Federal de Campina Grande, Campina Grande, 2017.
- [18] GOMES, O. F. M. Processamento e Análise de Imagens Aplicada à Caracterização Automática de Materiais. Dissertação de Mestrado – Departamento de Ciência dos Materiais e Metalurgia, Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro. p. 141. 2001.
- [19] GOMES, O. F. M. Microscopia co-localizada: novas possibilidades na caracterização de minérios. Tese (Doutorado em Engenharia Metalúrgica e de Materiais) – Departamento de Ciência dos Materiais e Metalurgia, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2007.
- [20] JANANI, V.; DINAKARAN, M. Infrared Image Enhancement Techniques – A Review. 2nd International Conference on Current Trends in Engineering and Technology, ICCTET'14, 2014.
- [21] PROJETO MAC MULTIMÍDIA. Material didático para disciplinas de Introdução à Computação. MAC 115 - Introdução à Computação para Ciências Exatas e Tecnologia. Prof. Francisco Reverbel. Universidade de São Paulo, Instituto de Matemática e Estatística - IME, Departamento de Ciência da Computação, 2008.
- [22] NETO, R. A. V. Introdução à morfologia matemática binária e em tons de cinza. Relatório de Iniciação Científica – Departamento de Matemática, Centro de Ciências Exatas, Universidade Estadual de Londrina, Londrina, 2010.
- [23] MEDEIROS, N. G.; SILVA, E. A.; NOGUEIRA, J. R. Segmentação morfológica de imagens utilizando o gradiente morfológico multi-escala. *Revista Brasileira de Cartografia*, n. 54, Universidade Estadual Paulista, Faculdade de Ciências e Tecnologia, Presidente Prudente, 2009.
- [24] REN, T. I. Processamento de imagem morfológica (morfologia matemática). Material Didático – Centro de Informática, Universidade Federal de Pernambuco, Recife, 2010.
- [25] BOBYR, M.; ARKHIPOV, A.; EMEL'YANOV, S.; MILOSTNAYA, N. A method for creating a depth map based on a three-level fuzzy model. *Engineering Applications of Artificial Intelligence*, v. 117, parte B, 2023.
- [26] MENG, B. C. C.; DAMANHURI, N. S.; OTHMAN, N. A. Smart traffic light control system using image processing. *IOP Conference Series: Materials Science and Engineering*, v. 1088, 2021.
- [27] ANTONELLO, R. Introdução a Visão Computacional com Python e OpenCV. Versão 0.8 – Não corrigida. 2017. Disponível em: <https://professor.luzerna.ifc.edu.br/ricardo-antonello/wp-content/uploads/sites/8/2017/02/Livro-Introdu%C3%A7%C3%A3o-a-Vis%C3%A3o-Computacional-com-Python-e-OpenCV-3.pdf>. Acesso em: 02 set. 2024.
- [28] BEERAVOLU, A. R.; AZAM, S.; JONKMAN, M.; SHANMUGAM, B.; KANNOORPATTI, K.; ANWAR, A. Preprocessing of Breast Cancer Images to Create Datasets for Deep-CNN. *IEEE Access*, v. 9, p. 36682-36692, 2021.
- [29] LIM, K. S.; LEE, L. H.; SIM, Y. A review of machine learning algorithms for fraud detection in credit card transactions. *IJCSNS International Journal of Computer Science and Network Security*, v. 21, n. 9, p. 31-36, 2021.
- [30] ASHISH MEHTA. K-Nearest Neighbors (KNN) Algorithm For Machine Learning. Disponível em: <https://ai.plainenglish.io/k-nearest-neighbors-knn-algorithm-for-machine-learning-ab6f17df4a7f>. Acesso em: 4 de junho de 2024.
- [31] WEBB, A. R.; COPSEY, K. D. Statistical pattern recognition. Third edition. Malvern, UK: John Wiley & Sons, 2011.
- [32] LIMA, D. C. Métodos para estimativa de imagens NIR a partir de imagens de câmeras RGB. Dissertação (Mestrado em Ciência da Computação) – Centro de Ciências Exatas e de Tecnologia, Universidade Federal de São Carlos, São Carlos, 2020.
- [33] DEEPA, B.; RAMESH, K. Epileptic seizure detection using deep learning through min max scaler normalization. *International Journal of Health Sciences*, v. 6, n. S1, p. 10981-10996, 2022.
- [34] SHAHEEN, H.; AGARWAL, S.; RANJAN, P. MinMaxScaler Binary PSO for Feature Selection. First International Conference on Sustainable Technologies for Computational Intelligence. *Advances in Intelligent Systems and Computing*, v. 1045. Springer, Singapore, 2020.
- [35] KRSTAJIC, D.; BUTUROVIC, L. J.; LEAHY, D. E. Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics*, v. 6, n. 10, 2014.
- [36] SHRIKANT, S. Understanding the Accuracy Score Metric's Limitations in the Data Science Classification Problems. LinkedIn, 26 mai. 2023. Disponível em: <https://www.linkedin.com/pulse/understanding-accuracy-score-metrics-limitations-data-akshay-w/>. Acesso em: 10 set. 2024.
- [37] CHAYATUMMAGOON, S. [dataset]. 2021. Dataset of Sugar Crystal Thailand.