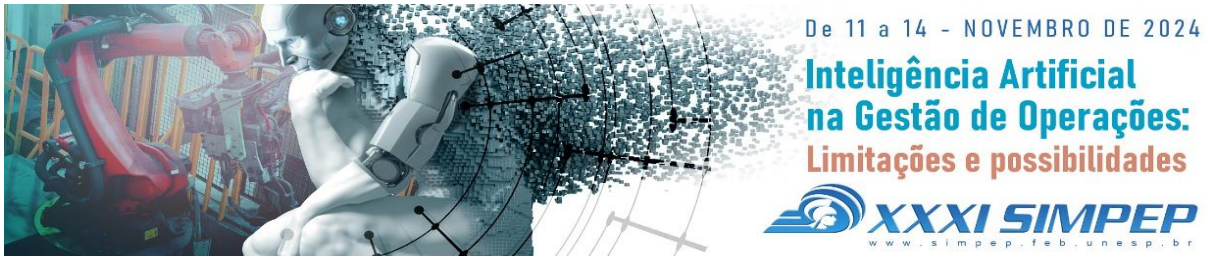


UTILIZAÇÃO DE ALGORITMOS DE APRENDIZADO DE MÁQUINA NA CLASSIFICAÇÃO E EXTRAÇÃO DE CARACTERÍSTICAS DE IMAGENS

ESTÊVÃO SANTOS CAVALCANTE – estevaos.108@gmail.com
PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS – PUC GOIÁS – GOIÂNIA - GO

MARCOS ANTONIO DE SOUSA - mas@pucgoias.edu.br
PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS – PUC GOIÁS – GOIÂNIA - GO

MARTA PEREIRA DA LUZ - marta.eng@pucgoias.edu.br
PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS – PUC GOIÁS – GOIÂNIA - GO

ÁREA: *3. PESQUISA OPERACIONAL*
SUBÁREA: *3.7 - INTELIGÊNCIA COMPUTACIONAL*

RESUMO: *ESTE TRABALHO ABORDA A UTILIZAÇÃO DE ALGORITMOS DE APRENDIZADO DE MÁQUINA PARA A CLASSIFICAÇÃO E EXTRAÇÃO DE CARACTERÍSTICAS DE IMAGENS. FORAM EMPREGADOS OS MÉTODOS SUPPORT VECTOR MACHINES (SVM), K-NEAREST NEIGHBORS (KNN) E CONVOLUTIONAL NEURAL NETWORKS (CNN) PARA ANÁLISE DAS IMAGENS. O APLICATIVO ORANGE CANVAS FOI UTILIZADO COMO FERRAMENTA DE ANÁLISE. DURANTE O ESTUDO, A CNN DEMONSTROU A MELHOR QUALIDADE DE CLASSIFICAÇÃO ENTRE OS MÉTODOS AVALIADOS. A ANÁLISE ENVOLVEU A COMPARAÇÃO DE DESEMPENHO DOS ALGORITMOS, EVIDENCIANDO A SUPERIORIDADE DA CNN EM TERMOS DE PRECISÃO E EFICÁCIA NA EXTRAÇÃO DE CARACTERÍSTICAS RELEVANTES. OS RESULTADOS INDICAM QUE A APLICAÇÃO DE CNNS É ALTAMENTE EFICAZ NA CLASSIFICAÇÃO DE IMAGENS, TORNANDO-AS UMA ESCOLHA PREFERENCIAL PARA TAREFAS SEMELHANTES. OS RESULTADOS EVIDENCIAM A IMPORTÂNCIA DE SELECIONAR ALGORITMOS ADEQUADOS PARA OTIMIZAR A PERFORMANCE EM PROJETOS DE RECONHECIMENTO DE PADRÕES E PROCESSAMENTO DE IMAGENS.*

PALAVRAS-CHAVES: *CNN; KNN; SVM; CLASSIFICAÇÃO; ORANGE CANVAS.*

1. INTRODUÇÃO

A utilização de algoritmos de aprendizado de máquina, como Support Vector Machines (SVM), K-Nearest Neighbors (KNN) e Convolutional Neural Networks (CNN), tem se mostrado eficaz na classificação e detecção de imagens, conforme evidenciado por Kasinathan et al. (2020). A extração de características morfológicas das imagens, como área, perímetro e comprimento do eixo maior, tem permitido uma identificação precisa e precoce dos padrões das imagens, proporcionando informações valiosas para a tomada de decisões baseadas em dados.

Nesse projeto, utilizou-se o executável *open-source* Orange Canvas (UL, 2024), que possui diversas bibliotecas integradas voltadas para Ciência de Dados, com finalidade de demonstrar a aplicação de maneira mais visual.

O Orange se apresenta como um laboratório virtual completo, reunindo em um único ambiente intuitivo uma vasta gama de ferramentas e bibliotecas voltadas para análise de dados. Através de *widgets* e fluxos de trabalho visuais, o software permite que usuários explorem e visualizem seus dados de forma intuitiva e interativa, desvendando padrões, identificando tendências e descobrindo *insights* valiosos.

O teste conduzido no Orange com os três algoritmos distintos - SVM, KNN e CNN - utilizou um conjunto idêntico de imagens para treinamento e validação, garantindo uma base de comparação justa. As diferenças observadas entre os modelos foram decorrentes exclusivamente dos ajustes de hiperparâmetros e das configurações específicas de treinamento empregadas para cada algoritmo. Esses ajustes incluíram parâmetros como a taxa de aprendizado, o número de camadas nas redes neurais, os critérios de parada, entre outros. Essa abordagem permitiu avaliar com precisão o impacto das configurações de treinamento na performance de cada algoritmo, destacando a importância de uma otimização adequada dos hiperparâmetros para alcançar resultados superiores na classificação de imagens.

2. FUNDAMENTAÇÃO TEÓRICA

A Inteligência Artificial (IA) é um ramo da ciência da computação que busca desenvolver mecanismos e dispositivos capazes de simular a capacidade humana de pensar e resolver problemas, ou seja, de demonstrar inteligência. Uma solução de IA integra diversas tecnologias, incluindo redes neurais artificiais, algoritmos e sistemas de aprendizado, que juntos conseguem replicar capacidades humanas relacionadas à inteligência, como raciocínio,

percepção ambiental e habilidade de análise para tomada de decisões (De Castro et al., 2016).

Há um imenso potencial benéfico para organizações que adotam a tecnologia de Inteligência Artificial (IA), independentemente do setor em que operam. A IA transcende a automação mecânica tradicional ao incorporar processos cognitivos que permitem aos sistemas aprender e se adaptar. Isso significa que uma solução de IA pode executar não apenas tarefas repetitivas, numerosas e manuais, mas também aquelas que exigem análise complexa e tomada de decisão informada. Por exemplo, na área de saúde, a IA pode auxiliar no diagnóstico de doenças através da análise de imagens médicas. No setor financeiro, pode otimizar portfólios de investimentos e detectar fraudes. Na manufatura, a IA pode prever falhas de máquinas e otimizar a produção. Essa capacidade de combinar automação com inteligência cognitiva oferece às organizações uma vantagem competitiva significativa, aumentando a eficiência, reduzindo erros e permitindo uma tomada de decisão mais rápida e precisa. Ao aprender continuamente com novos dados, os sistemas de IA se tornam cada vez mais eficazes, contribuindo para a inovação e a melhoria contínua dos processos organizacionais.

2.1 Técnicas de análise de imagem

No caso particular de atividades de análise de imagens, as técnicas de IA se tornaram imprescindíveis. Na aplicação de classificação de imagens, essas tecnologias oferecem precisão e eficiência excepcionais. Por meio de algoritmos avançados de aprendizado de máquina, é possível identificar padrões e características com uma velocidade e exatidão que superam as capacidades humanas.

As diversas técnicas de aplicação da Inteligência Artificial (IA) variam em termos de funcionamento, lógica aplicada e, principalmente, nos resultados esperados dos algoritmos desenvolvidos. Neste contexto, a aprendizagem de máquina (*machine learning*) se destaca como um conjunto de técnicas de IA que visa reconhecer padrões em bases de dados. A aprendizagem de máquina melhora a qualidade na geração de informações e conhecimentos, auxiliando gestores e aprimorando os processos de tomada de decisão nas empresas.

Entre as principais tarefas que podem ser executadas com técnicas de aprendizagem de máquina estão: análise descritiva de dados, predição (classificação e estimação), análise de grupos, associação e detecção de anomalias (Martins et al., 2019). Este trabalho foca no estudo de problemas de predição, com ênfase no processo de classificação de imagens. Os principais métodos de aprendizagem de máquina testados incluem: SVM, KNN e a CNN.

2.1.1 Support Vector Machines (SVM)

O Support Vector Machine (SVM) é um algoritmo de aprendizado de máquina supervisionado utilizado para classificação e regressão. Ele funciona criando um hiperplano que separa os dados em classes distintas em um espaço dimensional com base nas características dos dados, de acordo com Kasinathan et al. (2020). Um hiperplano é uma linha (em duas dimensões), um plano (em três dimensões) ou um subespaço em dimensões superiores, que divide o espaço de dados de modo a maximizar a margem entre as classes. A margem é a distância entre o hiperplano e os pontos de dados mais próximos de cada classe, conhecidos como vetores de suporte.

O SVM é muito utilizado em problemas de classificação, como a identificação de padrões em dados complexos. Ele requer a definição de parâmetros, incluindo o tipo de kernel (linear, polinomial, RBF), o parâmetro de regularização C e o coeficiente do kernel. A escolha adequada desses parâmetros é crucial para o desempenho do modelo.

O kernel é uma função que transforma os dados originais em um espaço dimensional mais alto, onde se torna mais fácil separar as classes com um hiperplano. Tipos comuns de kernel incluem o linear, que não altera o espaço dos dados; o polinomial, que transforma os dados em um espaço polinomial; e o RBF (Radial Basis Function), que mapeia os dados em um espaço de maior dimensão com base na distância dos pontos de dados. A seleção do kernel adequado e o ajuste dos hiperparâmetros associados são fundamentais para otimizar a eficácia do SVM.

Uma das principais vantagens do SVM é a sua capacidade de lidar com conjuntos de dados de alta dimensionalidade e separar classes não lineares. O SVM é eficaz tanto em problemas de classificação binária quanto multiclasse, além de ser robusto contra *overfitting*, ou seja, adaptação excessiva aos dados de treinamento, levando a um modelo que não consegue fazer previsões ou conclusões precisas com outros dados que não sejam os de treinamento, desde que os parâmetros sejam ajustados corretamente (Kasinathan et al., 2020).

No entanto, o SVM pode ser computacionalmente intensivo, especialmente em conjuntos de dados muito grandes. A escolha adequada do kernel e dos parâmetros de ajuste pode ser desafiadora e exige conhecimento especializado. Além disso, o SVM pode não ser tão eficaz em conjuntos de dados com ruído ou sobreposição de classes, necessitando de técnicas adicionais de pré-processamento de dados para melhorar o desempenho.

2.1.2 K-Nearest Neighbors (KNN)

O KNN é um algoritmo de aprendizado de máquina do tipo "*lazy learning*" que não requer uma fase de treinamento explícita. Ele classifica os dados com base na proximidade com os vizinhos mais próximos no espaço de características. O KNN é utilizado para classificação e regressão, onde a classe de um novo ponto de dados é determinada pela maioria dos k vizinhos mais próximos (Kasinathan et al., 2020).

O KNN é comumente utilizado em problemas de classificação e recomendação, onde a similaridade entre os dados é crucial. Para configurar o KNN, é necessário definir o valor de k (número de vizinhos), a métrica de distância (Euclidiana, Manhattan etc.) e possíveis pesos para os vizinhos. A escolha adequada de k e da métrica de distância é essencial para o desempenho do modelo.

Uma das principais vantagens do KNN é a simplicidade e facilidade de implementação. Ele é eficaz em lidar com conjuntos de dados com ruído e não exige suposições sobre a distribuição dos dados. Além disso, esse modelo de *machine learning* é um algoritmo não paramétrico, o que significa que pode se adaptar a diferentes formas de distribuição dos dados.

Por outro lado, o KNN pode ser computacionalmente caro, especialmente em conjuntos de dados grandes, devido à necessidade de calcular a distância entre todos os pontos. O desempenho do KNN também pode ser afetado pela escolha inadequada de k e da métrica de distância. Além disso, o KNN é sensível a *outliers*, ou seja, pontos que se diferenciam drasticamente dos outros, e requer um pré-processamento cuidadoso dos dados para obter resultados precisos. Essas limitações devem ser consideradas ao aplicar o KNN em problemas de classificação e regressão.

2.1.3 Convolutional Neural Networks (CNN)

As Redes Neurais Convolucionais (CNNs) são um tipo de modelo de aprendizado profundo projetado para processar dados de imagem e realizar reconhecimento visual. Elas consistem em várias camadas convolucionais que extraem características das imagens, seguidas por camadas de *pooling*, que reduzem a dimensionalidade dos dados e preservam as características mais importantes (Lopez et al., 2021). As camadas de *pooling*, como *max pooling* e *average pooling*, funcionam resumindo a presença de características em regiões subamostralizadas da imagem, o que ajuda a diminuir a quantidade de parâmetros e a prevenir o *overfitting*.

As camadas convolucionais são responsáveis por extrair características das imagens, enquanto as camadas de *pooling* reduzem a dimensionalidade dos dados, preservando as informações mais importantes. Após essas camadas, as camadas totalmente conectadas realizam a classificação final. A escolha adequada da arquitetura da rede e dos hiperparâmetros é fundamental para otimizar o desempenho do modelo, garantindo que a CNN capture os padrões relevantes nos dados de entrada e produza resultados precisos nas tarefas de visão computacional.

Uma das principais vantagens das Redes Neurais Convolucionais (CNNs) é a capacidade de aprender automaticamente características hierárquicas das imagens, sem a necessidade de extração manual de recursos. Elas são altamente eficazes no reconhecimento de padrões em imagens complexas e podem lidar com grandes conjuntos de dados (Kasinathan et al., 2020). Além disso, as CNNs são robustas a variações de escala, rotação e iluminação nas imagens.

No entanto, as CNNs podem ser computacionalmente intensivas, especialmente durante o treinamento em conjuntos de dados extensos. A interpretabilidade das CNNs pode ser um desafio devido à complexidade da arquitetura e ao grande número de parâmetros. Além disso, o ajuste fino dos hiperparâmetros pode ser trabalhoso e exigir experimentação extensiva para obter o melhor desempenho do modelo. Essas limitações devem ser consideradas ao aplicar CNNs em projetos de visão computacional.

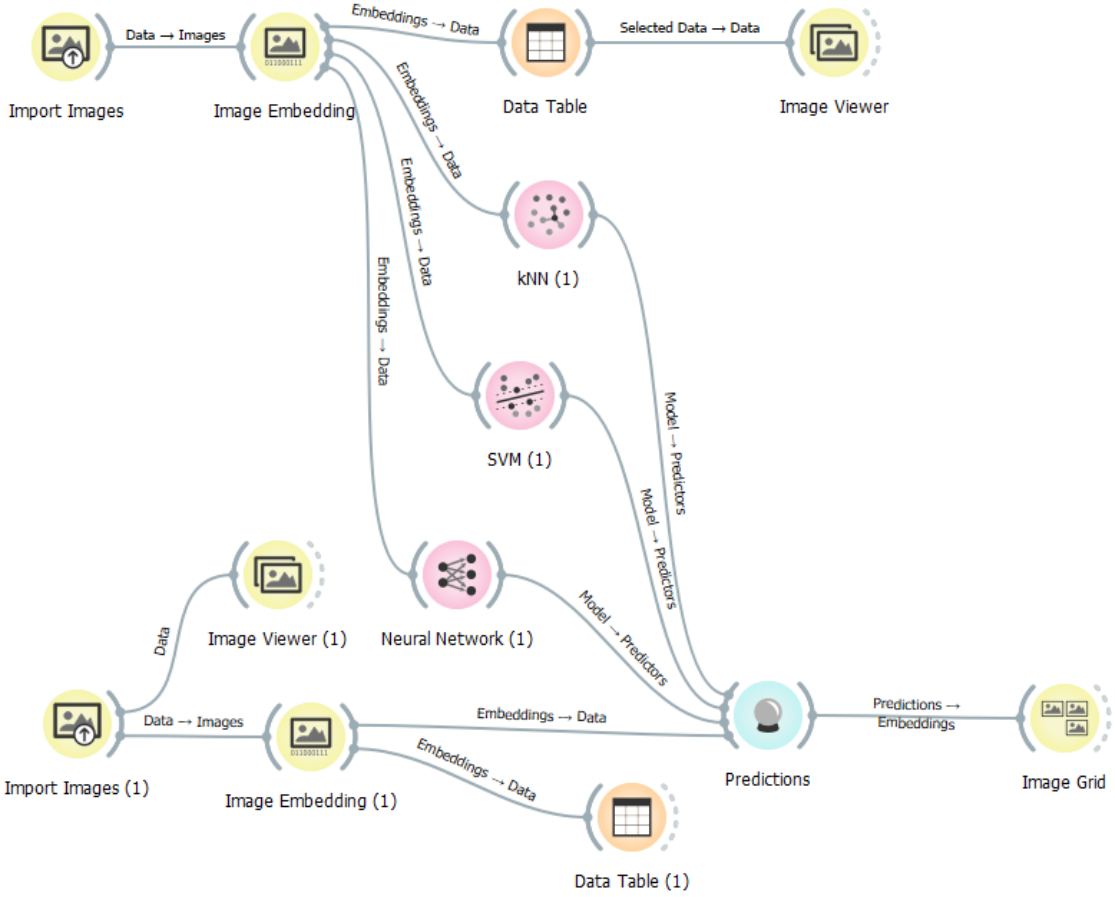
3. METODOLOGIA

O experimento realizado no *software* Orange Canvas (FIGURA 1) com os três algoritmos distintos - SVM, KNN e CNN - utilizou o mesmo conjunto de imagens para treinamento e validação, assegurando uma base de comparação equitativa. As variações nos resultados dos modelos foram exclusivamente atribuídas aos ajustes dos hiperparâmetros e às configurações específicas de treinamento de cada algoritmo.

A FIGURA 1 apresenta informações essenciais para o treinamento das IAs e a verificação de qual modelo é o mais eficaz na classificação de imagens. O "*Import Images*" é o local onde se inserem as imagens de treinamento, totalizando 1600 imagens divididas igualmente em duas categorias, com 800 imagens em cada. O conteúdo específico das imagens é irrelevante, pois as mesmas imagens são utilizadas para os três modelos. Em "*Import Images (1)*", foram inseridas 10 imagens aleatórias e distintas das imagens de treinamento, para que as IAs tentem categorizar automaticamente essas novas imagens com

base nas categorias do conjunto de treinamento.

FIGURA 1 – Treinamento dos modelos no Orange.



Fonte: autoria própria.

Após a inserção de todas as imagens de treinamento e validação no Orange Canvas, foram configurados os parâmetros e hiperparâmetros de cada modelo. No caso do KNN, representado por "KNN (1)", o número de vizinhos foi ajustado para 5, utilizou-se a métrica Euclidiana e aplicou-se peso uniforme.

O modelo SVM, denominado "SVM (1)", foi ajustado com os seguintes parâmetros: Perda de regressão Epsilon foi definida como 0,10, que controla a margem de tolerância na regressão; o Custo de regressão foi ajustado para 1,00, influenciando a penalização dos erros de classificação; o Parâmetro g foi configurado como "auto", determinando a escala para a função kernel; o Parâmetro C foi ajustado para 1,00, regulando a margem de erro permitida; o Kernel foi definido como Sigmoid, escolhendo a função de kernel para transformar os dados; a Tolerância numérica foi fixada em 0,0010, estabelecendo o critério de parada; e o Limite de iterações foi definido em 300, controlando o número máximo de iterações no treinamento.

Por último, a CNN, representada por "Neural Network (1)", foi configurada com os seguintes parâmetros: o número de neurônios nas camadas ocultas foi definido como 100, a função de ativação escolhida foi a ReLU (Retificação Linear Unitaria), e o solver utilizado foi o SGD (Gradiente Descendente Estocástico). A regularização foi aplicada com um valor de α=0,5, para evitar overfitting. O número máximo de iterações foi ajustado para 100, determinando o limite de ciclos de treinamento da rede neural. Depois que todos os modelos foram configurados, foi realizado um teste de estimativa de classificação das imagens de validação para os três algoritmos de IA simultaneamente.

4. RESULTADOS

O teste foi representado na FIGURA 1 por "*Predictions*", um método de avaliação de modelos que determina qual técnica de aprendizado de máquina é mais adequada para a classificação das imagens. Os resultados desse teste são apresentados na TABELA 1. A comparação dos algoritmos visa identificar qual deles oferece melhor desempenho em termos de precisão e extração de características relevantes, destacando a importância de selecionar algoritmos adequados para otimizar a performance em projetos de reconhecimento de padrões e processamento de imagens.

TABELA 1– Resultados de avaliação de modelos via "*Predictions*".

Categoria	Classificação (kNN)	Erro (kNN)	Classificação (SVM)	Erro (SVM)	Classificação (CNN)	Erro (CNN)
Com doença	Com doença	0.000	Com doença	0.076	Com doença	0.022
Com doença	Com doença	0.000	Com doença	0.315	Com doença	0.087
Com doença	Sem doença	1.000	Sem doença	0.982	Sem doença	0.861
Com doença	Sem doença	0.600	Sem doença	0.772	Com doença	0.444
Com doença	Com doença	0.000	Sem doença	0.969	Sem doença	0.813
Sem doença	Sem doença	0.000	Sem doença	0.017	Sem doença	0.022
Sem doença	Sem doença	0.000	Sem doença	0.009	Sem doença	0.010
Sem doença	Sem doença	0.000	Sem doença	0.079	Sem doença	0.184
Sem doença	Sem doença	0.400	Sem doença	0.198	Sem doença	0.357
Sem doença	Sem doença	0.000	Sem doença	0.005	Sem doença	0.014

Fonte: autoria própria.

Ao observar os resultados de "Prediction" na TABELA 1, a primeira coluna "Categoria" apresenta as 10 imagens de validação inseridas no Orange, que foram igualmente

distribuídas em duas categorias: Com Doença e Sem Doença, em "*Import Images (1)*". Os modelos de IA devem prever a que categoria cada uma dessas imagens pertence.

As colunas de kNN, SVM e CNN tentam se relacionar com a coluna "Categoria". Essas três primeiras colunas têm o objetivo de prever a qual categoria uma determinada imagem pertence. Além disso, cada uma delas possui uma coluna de erro que indica se houve erro na previsão. Quanto mais próximo de zero o valor na coluna de erro, menor o erro; quanto mais próximo de um, maior o erro.

Nota-se que o kNN acertou a classificação de 8 das 10 imagens, enquanto o SVM acertou 7 das 10 e a CNN 8 das 10. No entanto, ao considerar as colunas de "Erro", que indica a precisão das previsões, é possível descartar o SVM, que teve um desempenho inferior aos outros dois modelos. Ao analisar as 3 imagens em que os outros dois modelos erraram, percebe-se que o kNN cometeu um erro de 100% na terceira imagem, enquanto a CNN não comete um erro dessa magnitude.

Além disso, outro fator necessário para avaliar qual o melhor algoritmo para a aplicação foi a coluna “Precisão”, na TABELA 2, onde foram avaliadas as precisões dos modelos. Tanto o kNN quanto a CNN apresentaram 85,70% de acurácia, tornando necessário considerar um segundo critério. Dessa vez, observou-se a coluna “AUC” (*Area under ROC curve*). A AUC varia de 0 a 1, onde 1 indica um modelo perfeito e 0,5 representa um modelo que não possui capacidade discriminativa melhor do que o acaso. Em resumo, quanto maior a AUC, melhor o modelo é em distinguir entre as classes positivas e negativas (UL, 2024).

TABELA 2– Resultados da avaliação dos algoritmos.

Modelo	AUC	CA	F1	Precisão	Recuperação
kNN	0.860	0.800	0.792	0.857	0.800
SVM	0.840	0.700	0.670	0.812	0.700
CNN	0.880	0.800	0.792	0.857	0.800

Fonte: autoria própria.

Dessa forma, os resultados mostraram que o kNN obteve uma AUC de 0,86, enquanto a CNN alcançou uma AUC de 0,88. Além disso, ao analisar as colunas de "Erro" na TABELA 1, nota-se que a CNN apresenta menores variações de erro. Com base nesses critérios, concluiu-se que a CNN foi o modelo superior para a classificação das imagens de validação, apresentando melhor desempenho tanto em termos de AUC quanto na consistência dos erros.

5. CONCLUSÃO

O modelo de inteligência artificial escolhido foi a CNN, pois conseguiu classificar as imagens com mais precisão do que o kNN e a SVM, além de cometer menos erros e apresentar maior consistência nas suas predições. No entanto, é importante destacar que essa avaliação foi realizada exclusivamente no software Orange Canvas, onde os parâmetros e hiperparâmetros foram ajustados da melhor forma possível para obter os melhores resultados, conforme descrito em "2.2 Determinação da escolha do algoritmo". Apesar disso, podem ter ocorrido erros humanos durante o processo. Testar mais variações de taxa de aprendizagem, número de neurônios, número de vizinhos e iterações pode potencialmente aumentar a acurácia de um determinado modelo.

Ainda assim, foram realizados diversos testes de configurações dos modelos no aplicativo, e as mencionadas em "2.1 Técnicas de análise de imagem" apresentaram as melhores precisões e acertos para cada um dos três modelos. Vale ressaltar, no entanto, que o Orange oferece menos liberdade de configuração e criação de parâmetros em comparação com linguagens de programação como Python ou R. O software foi utilizado devido à sua facilidade de uso e rapidez para testes iniciais, sendo particularmente útil para a criação de protótipos de modelos de *machine learning*.

É importante mencionar que o estudo está limitado aos dados analisados e aos métodos de avaliação utilizados. Embora a utilização do mesmo conjunto de dados para os diferentes modelos garanta uma base comparativa justa, os resultados obtidos podem variar significativamente quando aplicados a outros conjuntos de dados ou contextos. A eficácia dos modelos de IA, como CNN, KNN e SVM, na classificação de imagens pode ser influenciada por fatores como a qualidade e a diversidade das figuras, o pré-processamento dos dados e os parâmetros específicos dos algoritmos. Portanto, embora as conclusões deste estudo sejam válidas para os dados e métodos utilizados, recomenda-se cautela ao generalizar os resultados para outras situações sem uma análise adicional.

REFERÊNCIAS

DE CASTRO, L.N; FERRARI, D.G. **Introdução à mineração de dados: conceitos básicos, algoritmos e aplicações.** São Paulo. Saraiva, 2016.

KASINATHAN, T.; SINGARAJU, D.; UYYALA, S. R. **Insect classification and detection in field crops using modern machine learning techniques.** Information Processing in Agriculture, v. 8, n. 3, p. 446-457, 2020. DOI: 10.1016/j.inpa.2020.09.006. Disponível em: https://doi.org/10.1016/j.inpa.2020.09.006. Acesso em: 25 abr. 2024.

LOPEZ, Martin Andreoni; MATTOS, Diogo M. F. Resumo de Grandes Volumes de Dados

com Filtro de Bloom: **Uma Abordagem Eficiente para Aprendizado Profundo com Redes Neurais Convolucionais em Fluxos de Rede**. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS (SBRC), 39., 2021, Uberlândia. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2021. p. 532-545. ISSN 2177-9384. DOI: https://doi.org/10.5753/sbrc.2021.16745. Acesso em: 06 jun. 2024.

MARTINS, Lucimara; MURAMATSU JÚNIOR, Mario; SERAPIÃO, Adriane. **Classificação de imagens de ideogramas Kuzushiji com Redes Neurais Convolucionais.** In: ENCONTRO NACIONAL DE INTELIGÊNCIA ARTIFICIAL E COMPUTACIONAL (ENIAC), 16., 2019, Salvador. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2019. p. 309-320. ISSN 2763-9061. DOI: https://doi.org/10.5753/eniac.2019.9293. Acesso em: 04 jun. 2024.

UL - University of Ljubljana. **Orange Data Minning (.exe).** Disponível em: https://orangedatamining.com. Acesso em: 25 abr. 2024.