

Análise de Métricas de Complexidade de Imagem no Contexto de Classificação

**Gabriel Antunes Santos¹, Honovan Paz Rocha², Rogerio Alves Santana²
Thales Francisco Mota Carvalho²**

¹*Universidade Federal Dos Vales Jequitinhonha e Mucuri, UFVJM- Janaúba, Brasil
(gabriel.antunes@ufvjm.edu.br)*

²*Universidade Federal Dos Vales Jequitinhonha e Mucuri, UFVJM- Janaúba, Brasil*

Resumo: Esse trabalho investiga a relação entre métricas de complexidade de imagem e o desempenho de modelos de aprendizado de máquina. A análise revela que essas métricas conseguem diferenciar as 5 bases de dados testadas e apresentam leve tendência de proporcionalidade direta com a eficácia dos classificadores. Conclui-se que as métricas são recursos analíticos muito úteis para a avaliação prévia de bases imagéticas.

Palavras-chave: Machine Learning, Classificação, Métricas de Imagens, Visão Computacional;

INTRODUÇÃO

Na era digital, com o avanço da tecnologia, os dispositivos de captura de imagens evoluíram significativamente, resultando em maior qualidade e facilidade de acesso. Como consequência, as imagens digitais se tornaram uma das fontes de informação mais ubíquas e ricas da atualidade, impulsionando avanços que vão desde os diagnósticos médicos por imagem, transformando a precisão na identificação de patologias (Najjar, 2023), até o monitoramento ambiental por meio do sensoriamento remoto, essencial para a gestão sustentável dos recursos naturais (Chuvieco, 2020). O volume de dados imagéticos gerados e armazenados globalmente tem aumentado de forma exponencial, impulsionado por uma ampla variedade de fontes (Vassakis et al., 2017).

Essa vasta disponibilidade de dados visuais representa uma oportunidade imensa e, simultaneamente, um desafio computacional significativo que demanda grande capacidade de processamento. As imagens digitais carregam, por sua natureza, uma grande densidade de informações, muitas vezes com ruídos, variações de iluminação, oclusões e uma diversidade de objetos e texturas que dificultam a interpretação automática (Gonzalez, 2009). Nesse contexto, a Aprendizagem de Máquina (AM) tem se mostrado uma abordagem promissora e flexível, oferecendo ferramentas capazes de aprender diretamente com dados imagéticos e realizar tarefas como classificação e segmentação (Szeliski, 2022).

Embora os algoritmos de AM tenham ampla aplicabilidade, escolher a técnica mais adequada para uma tarefa específica ainda representa um desafio considerável. Fatores como o volume de dados e as características intrínsecas das próprias imagens,

como sua complexidade textural e estrutural, influenciam diretamente o desempenho. Na ausência de diretrizes bem definidas, a escolha da técnica geralmente depende de um processo empírico de tentativa e erro, o qual demanda tempo e altos recursos computacionais de forma significativa. Dessa forma, emerge uma lacuna metodológica expressiva: a carência de abordagens sistemáticas que possam orientar a seleção de algoritmos de AM com base em uma análise prévia e quantitativa das características dos dados imagéticos.

Uma abordagem promissora para preencher essa lacuna é quantificar precocemente a complexidade dos dados, pois conjuntos mais complexos afetam de maneira direta o resultado do modelo durante o treinamento (Feng et al., 2022). Essa complexidade é determinada não apenas pelas características intrínsecas de imagens individuais, como riqueza de detalhes, formas geométricas, texturas e níveis de ruído, mas de forma crucial, pelas inter-relações dessas características em todo o conjunto, como a variabilidade intraclasse e a separabilidade entre classes.

Na literatura recente, diversas métricas vêm sendo utilizadas nas áreas de extração de características e segmentação semântica. Estudos demonstram a eficácia de abordagens como a dimensão fractal e redes complexas para analisar a rugosidade geométrica e o nível de ocupação espacial (Backes, 2010). Outras pesquisas validam o uso da Entropia de Shannon (Rahane e Subramanian, 2020) e de medidas baseadas no domínio da frequência para quantificar a desordem (Khan et al., 2022), assim como medidas de compressão (Durmus, 2020). Trabalhos adicionais têm explorado autocodificadores convolucionais (Vatandoust, 2024),



agrupamento espectral (Branchaud-Charron et al., 2019) e aprendizagem contrastiva (Liu et al., 2024) para antever a complexidade inerente a bases de dados estruturadas ou não estruturadas. Aplicações práticas também revelam que imagens mais irregulares, como as de tecidos patológicos, apresentam maior dimensão fractal (Silva et al., 2020).

Apesar da evolução metodológica, o desafio de mapear de forma direta a complexidade de um conjunto de imagens para o desempenho empírico de um espectro variado de classificadores ainda carece de consolidação. O presente estudo, portanto, tem como objetivo central elencar e analisar as métricas de complexidade de imagens propostas na literatura, avaliando sua correlação estatística com o desempenho de técnicas simples e avançadas de aprendizado de máquina para classificação de imagens.

MATERIAL E MÉTODOS

O desenvolvimento deste estudo fundamentou-se em um planejamento experimental estruturado em etapas sequenciais: seleção das bases de dados, pré-processamento e extração de características, cálculo matemático das métricas de complexidade, modelagem dos classificadores e validação cruzada seguida de análise estatística.

SELEÇÃO DE DADOS: Para abranger diferentes níveis de dificuldade, foram utilizadas cinco bases de dados representativas:

MNIST: Introduzida por LeCun et al. (2002), composta por 70.000 imagens em tons de cinza (28x28 pixels) de dígitos manuscritos (0 a 9), sendo amplamente utilizada como referência por sua estrutura padronizada.

CIFAR-10: Proposta por Krizhevsky et al. (2009), contém 60.000 imagens coloridas (32x32 pixels) divididas igualmente em 10 classes, introduzindo desafios de iluminação, perspectiva e fundos variados.

Infrared Solar Modules (ISM): Apresentada por Millendorf et al. (2020), é um conjunto real de 20.000 imagens termográficas (40x24 pixels) de painéis solares distribuídas em 12 classes (11 de falhas e uma nominal).

Lesser Mouse Deer and Rabbit (LMDR): Utilizada por Oliveira et al. (2024), é uma base binária composta por 144 imagens de animais nativos capturados em cenários heterogêneos de natureza.

Vehicle Type Recognition (VTR): Também empregada por Oliveira et al. (2024), conta com 400 imagens de veículos distribuídas em quatro classes

(ônibus, carro, caminhão e motocicleta), capturadas em variados cenários.

PRÉ-PROCESSAMENTO E EXTRAÇÃO DE CARACTERÍSTICAS: Inicialmente, todas as imagens foram padronizadas para garantir uniformidade entre as bases de dados. As etapas de pré-processamento incluíram redimensionamento, conversão para tons de cinza quando necessário, aplicação de filtro Gaussiano para redução de ruídos e normalização dos valores dos pixels.

Para os classificadores tradicionais (KNN, SVM e MLP), foram empregadas técnicas de extração de características visando reduzir a dimensionalidade e preservar informações relevantes das imagens. Foram utilizados o Local Binary Patterns (LBP) para representação de textura, o Histogram of Oriented Gradients (HOG) para descrição de formas e contornos, e o detector de bordas de Canny, responsável por destacar estruturas geométricas relevantes. As características extraídas foram posteriormente normalizadas antes da etapa de treinamento.

No caso das Redes Neurais Convolucionais (CNN), as imagens foram utilizadas diretamente, dispensando a etapa de extração manual de características, uma vez que essas arquiteturas realizam esse processo automaticamente durante o treinamento.

MÉTRICAS DE COMPLEXIDADE: A complexidade das bases de dados foi avaliada por meio de cinco métricas encontradas na literatura, selecionadas por representarem diferentes propriedades espaciais, estatísticas e espectrais das imagens:

Entropia de Shannon, utilizada para quantificar o grau de informação presente na distribuição dos níveis de cinza (Shannon, 1948). Para uma imagem em escala de cinza, a entropia é definida na Equação 1:

$$H = - \sum_{i=0}^{n-1} p_i \log p_i \quad (1)$$

em que n representa o número de níveis de cinza e p_i é a probabilidade de um pixel possuir o nível de cinza i . Ela mede a incerteza de um sistema com base na distribuição de probabilidade de uma variável, como a intensidade de pixels em uma imagem.

Entropia da Matriz de Coocorrência de Níveis de Cinza (GLCM), empregada para caracterizar padrões de textura e relações espaciais entre pixels. Nesse contexto, a entropia da GLCM é calculada comorealizado em (RAHANE; SUBRAMANIAN, 2020):



$$H_g = - \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} p(i,j) \log p(i,j) \quad (2)$$

em que n representa o número de níveis de cinza e $p(i,j)$ é a probabilidade de que dois pixels, com intensidades i e j , estejam separados pela distância e direção especificadas no deslocamento.

Delentropia, baseada na distribuição dos vetores gradientes da imagem, permitindo capturar informações estruturais globais (Khan, Naqvi e Meijering, 2022). A equação da delentropia é parecida com a Equação 1 da entropia de Shannon, mas em vez de calcular a probabilidade entre pixels individuais, ela utiliza a função de delentropia definida na Equação 3:

$$p_{i,j} = \frac{1}{4WH} \sum_{w=0}^{W-1} \sum_{h=0}^{H-1} \delta_{i,dx_{(w,h)}} \delta_{j,dy_{(w,h)}} \quad (3)$$

em que dx e dy são kernels derivativos na direção de x e y , respectivamente; δ representa o processo de binning (organização de valores em faixas) na geração do histograma; e W e H são a largura e a altura da imagem, respectivamente. A equação completa da Delentropia é descrita por:

$$DE = - \frac{1}{2} \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} p(i,j) \log p(i,j) \quad (4)$$

Aqui, I e J representam o número de bins (células discretizadas) em cada uma das duas dimensões da função de densidade de probabilidade. O termo multiplicativo em (4) está associado à amostragem generalizada de Papoulis, a qual reduz a taxa de entropia em 50%. Uma delentropia alta indica que a imagem apresenta uma ampla variação nas intensidades dos pixels, refletindo a presença de detalhes mais complexos. Por outro lado, uma delentropia baixa sugere uma distribuição uniforme das intensidades, o que aponta para uma estrutura simples e uma imagem com menos detalhes (CAGAS et al., 2024).

Frequência Média, obtida a partir da Transformada de Fourier, utilizada para representar a distribuição da energia espectral das imagens (Gonzalez e Woods, 2009). A Frequência Média (FM) de um sinal é calculada como o somatório do produto entre a frequência e o espectro de potência do sinal, dividido pela soma total do espectro de potência. O espectro de potência é obtido a partir do módulo ao quadrado da Transformada de Fourier do sinal, conforme a Equação 5.

$$\bar{f} = \frac{\sum_{k=0}^{N-1} f_k \cdot |x[k]|^2}{\sum_{k=0}^{N-1} |x[k]|^2} \quad (5)$$

Dimensão Fractal, calculada pelo método Box-Counting, utilizada para quantificar a irregularidade geométrica e o grau de ocupação espacial dos objetos presentes nas imagens (Backes, 2010). O cálculo é realizado sobrepondo uma grade de quadrados a uma imagem e contando quantos quadrados são necessários para cobrir o objeto presente na imagem (COELHO; COSTA, 1995). A estimativa da dimensão fractal pelo método Box-Counting baseia-se na relação entre o tamanho da caixa r , usada na grade de quadrados, e o número de quadrados que coincidem com o objeto analisado $N_a(r)$.

$$DF = - \lim_{r \rightarrow 0} \frac{\log N_a(r)}{\log(r)} \quad (6)$$

Assim, a partir do gráfico de $\log N_a(r)$ (número de caixas ocupadas) por $\log(r)$, pode-se calcular a dimensão fractal através do coeficiente angular da reta na equação 6.

As métricas foram calculadas individualmente para todas as imagens das cinco bases, sendo posteriormente obtidas as médias e os desvios-padrão correspondentes a cada conjunto de dados.

ALGORITMOS DE CLASSIFICAÇÃO E PLANEJAMENTO EXPERIMENTAL: O espectro de classificadores foi escolhido para cobrir desde a aprendizagem baseada em instâncias até o aprendizado profundo:

KNN: Algoritmo simples que dispensa modelo explícito de treinamento; avaliado variando o número de vizinhos (k) e métricas de distância (Euclidiana L2, Manhattan L1 e Cosseno).

SVM: Otimiza o hiperplano de separação através da maximização de margem. O algoritmo foi adaptado para multiclasse via One-vs-One, utilizando kernel RBF e explorando empiricamente o parâmetro de regularização C .

MLP: Rede neural feedforward testada com uma e duas camadas ocultas (10, 20 ou 40 neurônios), funções de ativação ReLU e Sigmoid, otimizador SGD e perdas de Entropia Cruzada e Margem Múltipla.

CNN: Rede neural convolucional responsável pela extração autônoma de características hierárquicas (bordas, texturas) através de filtros deslizantes, funções ReLU e camadas de max-pooling, seguida de achatamento e classificação.

A robustez da avaliação preditiva foi garantida pela aplicação da validação cruzada aninhada, mitigando riscos de sobreajuste.

ANÁLISE ESTATÍSTICA: Após o cálculo das métricas de complexidade e a obtenção do



desempenho dos classificadores, foi realizada uma análise estatística para investigar a relação entre esses dois conjuntos de resultados.

Inicialmente, as bases de dados foram ordenadas de acordo com os valores médios das métricas de complexidade, permitindo estabelecer um ranking relativo de dificuldade entre os conjuntos de imagens. Em seguida, foi empregada a correlação de Spearman, adequada para avaliar relações monotônicas entre variáveis sem pressupor distribuição normal dos dados, verificando a existência de associação entre a complexidade das bases e o desempenho dos classificadores.

Todas as implementações foram desenvolvidas em Python, utilizando bibliotecas como NumPy, OpenCV, Scikit-learn, PyTorch, SciPy e Matplotlib. Os experimentos foram executados em um computador equipado com processador AMD Ryzen 5 5600, 32 GB de memória RAM e placa gráfica NVIDIA RTX 3060 de 12 GB, ambiente suficiente para o treinamento e avaliação dos modelos empregados neste estudo.

RESULTADOS E DISCUSSÃO

A investigação quantitativa iniciou-se pela distribuição das métricas de complexidade em cada base de dados. O comportamento métrico revelou de forma consistente o grau de diferença dos conjuntos imagéticos.

MÉTRICAS DE COMPLEXIDADE: A Entropia de Shannon confirmou a MNIST como a base de menor complexidade composicional (média de $1,602 \pm 0,380$), refletindo a imensa uniformidade de seus fundos pretos. O ISM figurou em um nível intermediário ($5,870 \pm 0,510$), enquanto VTR ($7,030 \pm 1,170$), CIFAR-10 ($7,260 \pm 0,487$) e LMDR ($7,370 \pm 0,760$) apresentaram taxas elevadas de entropia, com a LMDR evidenciando a maior variabilidade de pixels devido aos cenários naturais ruidosos.

Na análise textural por meio da GLCM, os padrões mantiveram-se. A MNIST exibiu ausência prática de textura ($2,53 \pm 0,60$), enquanto o ISM revelou padrões intermediária provinda das manchas térmicas ($8,69 \pm 0,81$). As bases CIFAR-10 (10,21), VTR (11,57) e LMDR (12,43) dominaram o topo do ranking, atestando aspereza visual considerável e multiplicidade de transições inter-pixels. A Delentropia, medindo os gradientes espaciais, ratificou a ordem, com MNIST no vale inferior (1,65) e a LMDR consolidada no pico da desordem (5,20).

A Frequência Média, todavia, inverteu esse panorama, indo na contramão das expectativas habituais. A base MNIST registrou o valor máximo (0,181), um fenômeno estatístico motivado pela transição abrupta e em alta frequência das bordas

brancas dos dígitos em contraste imediato com o fundo totalmente negro. Em contrapartida, bases de altíssima resolução e transições suaves entre pixels, como LMDR e VTR, obtiveram as menores médias (0,064 e 0,077, respectivamente).

Ao avaliar o espaço ocupado, a Dimensão Fractal forneceu um resultado numérico escalar preciso da irregularidade das formas. A base LMDR confirmou possuir as silhuetas mais instáveis e irregulares (média de 1,69), seguida pela VTR (1,62) e CIFAR-10 (1,59). Um ponto de destaque ocorreu entre as bases mais simples: a dimensão fractal da MNIST (1,32) superou a da ISM (1,15), marcando uma inversão de complexidade geométrica, indicando que os contornos dos dígitos manuscritos possuem uma irregularidade posicional levemente maior que as manchas térmicas dos painéis solares.

DESEMPENHO DOS CLASSIFICADORES: A análise preditiva utilizando o KNN demonstrou que na base MNIST as imagens originais combinadas com a métrica de Cosseno bastaram para alcançar excelentes 97,57% de F1-Score. Para os problemas complexos, a extração prévia provou-se obrigatória. No CIFAR-10, ISM e VTR, o extrator HOG garantiu o ápice do desempenho (embora baixos: 52,61%, 44,09% e 57,02%, respectivamente), enquanto o extrator LBP dominou amplamente a distinção das texturas de animais na base LMDR, cravando 81,54% de acerto com a métrica Euclidiana L2.

A progressão para o algoritmo SVM via estratégia One-vs-One proporcionou saltos consideráveis. Sob o Kernel RBF e regularização C ajustada (majoritariamente 5), o SVM cravou 97,47% na MNIST, evoluiu o CIFAR-10 para 62,42%, deu um expressivo salto na ISM alcançando 69,50%, solidificou 87,47% na LMDR (via LBP) e atingiu 70,00% na VTR.

A aplicação da rede MLP com otimizador SGD demonstrou comportamentos polares. Foi capaz de estabilizar incrivelmente na MNIST, atingindo 96,67% (com duas camadas de 40 neurônios, função ReLU e Entropia Cruzada). Obteve também sólidos 88,44% na LMDR e 73,10% na VTR utilizando apenas uma camada oculta. No entanto, na base termográfica ISM, a rede neural rasa fracassou estruturalmente, ruindo para a marca de 34,04%, desempenho substancialmente inferior aos classificadores estatísticos mais clássicos.

Finalmente, a CNN consolidou o domínio do aprendizado profundo, extraindo matrizes de convolução otimizadas de forma end-to-end. Cravou médias excelentes: MNIST (98,80%), LMDR (98,35%), CIFAR-10 (96,28%) e VTR (89,04%), superando amplamente os concorrentes. O único revés ocorreu na base ISM, onde a complexidade inerente e a natureza dos defeitos nas células

fotovoltaicas restringiram a arquitetura CNN a 67,12%, índice que foi superado pelo ajuste de margem suportado pela SVM (69,50%), evidenciando a não-linearidade do desafio de classificação.

ANÁLISE ESTATÍSTICA E CORRELAÇÃO: Para avaliar matematicamente as observações empíricas, foi estruturada uma robusta análise estatística. Testes preliminares (Shapiro-Wilk e Levene) revelaram ausência de normalidade e presença de heterocedasticidade nos resíduos, o que motivou a adoção do teste não-paramétrico de Kruskal-Wallis. Com a constatação de diferenças significativas ($p < 0.05$) entre as bases, aplicou-se o teste post-hoc de Dunn com correção de Bonferroni.

Os resultados desta etapa estão materializados na Figura 1 do estudo, que ilustra o ranqueamento estatístico das bases com 95% de confiança. A análise minuciosa desta figura revela agrupamentos cruciais (denotados por letras estatísticas). A base MNIST isola-se sistematicamente como o conjunto de menor valor, de forma independente nas métricas de Entropia, GLCM, Delentropia e Dimensão Fractal.

Por outro lado, a Figura 1 evidencia uma forte zona de equivalência entre as bases CIFAR10, LMDR e VTR. Na métrica de Entropia de Shannon, CIFAR-10, LMDR e VTR formam um bloco estatisticamente equivalente, atestando a similaridade de desordem informacional entre elas. A GLCM confirma o pareamento entre VTR e LMDR no espectro de maior complexidade textural. A Delentropia e a Dimensão Fractal alinham as bases CIFAR-10 e VTR, apontando semelhanças estruturais e de ocupação espacial. A métrica de Frequência Média, reforçando seu comportamento inverso, agrupa LMDR e VTR na porção inferior do ranking estatístico. Em síntese, a Figura 1 consolida a ideia de que bases de imagens provenientes de ambientes distintos podem compartilhar profundas equivalências matemáticas e estruturais, quantificáveis.

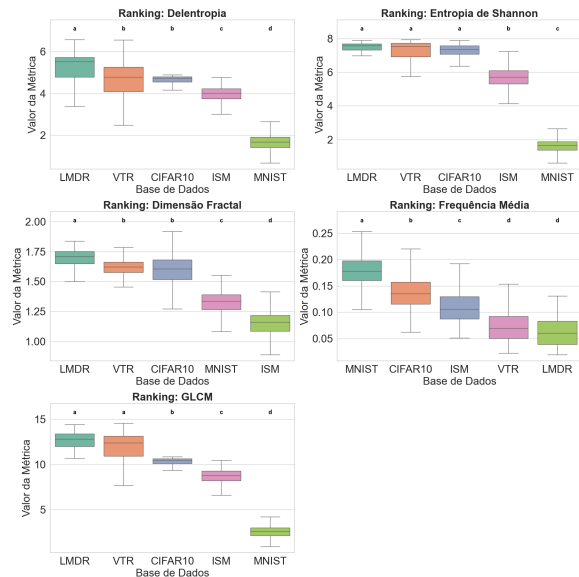


Figura 1: Ranking de Complexidade das Bases de Dados (Confiança de 95%).

Na Figura 2, à primeira vista, denota uma considerável dispersão dos dados, evidenciando alta incerteza estatística e indicando uma correlação quase inexistente. Ao analisar os valores obtidos através da correlação de Spearman fica claro que não pode-se afirmar que existe correlação entre as métricas e o desempenho da CNN por ter apresentado um alto p-valor.

Contudo, focando-se no agrupamento principal das bases de dados, ou seja retirando a base MNIST, é possível visualizar uma tendência de proporcionalidade direta para a maioria das métricas analisadas: o aumento dos valores das métricas costuma acompanhar a elevação do F1-Score da CNN. Esse comportamento, entretanto, não se confirma para a Frequência Média, que apresenta uma dinâmica de distribuição distinta.

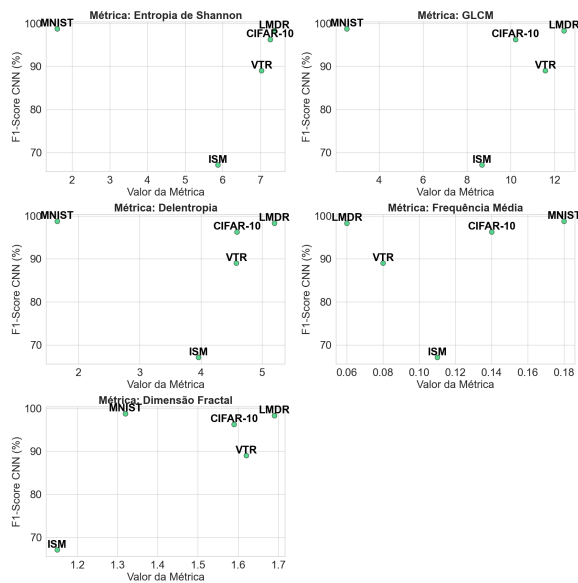


Figura 2: Análise de Dispersão dos Dados.

CONCLUSÃO

O presente trabalho teve como objetivo principal analisar a complexidade de diferentes bases de dados de imagens e investigar como essas características afetam o desempenho de algoritmos de classificação. Por meio da aplicação de métricas espaciais e de textura, nomeadamente Entropia de Shannon, GLCM, Delentropia, Frequência Média e Dimensão Fractal, em conjunto com a avaliação de classificadores tradicionais e de aprendizado profundo, compreendendo KNN, SVM, MLP e CNN, foi possível traçar um panorama claro sobre o comportamento dos dados e dos modelos.

A análise revelou que é possível quantificar e diferenciar estatisticamente a quantidade de informação, as formas e as texturas presentes nas bases de dados. A base MNIST consolidou-se com os menores valores nas métricas, o que é perfeitamente coerente com a sua natureza padronizada e estruturalmente muito homogênea de dígitos em tom branco sobre fundos pretos uniformes. Essa simplicidade resulta em baixa entropia e textura devido à ausência de ruídos ou detalhes granulares. E baixa dimensão fractal marcada por traços simples e sem complexidade espacial. Em contrapartida, o contraste abrupto entre o fundo nulo e os dígitos gera bordas muito fortes e bem definidas, o que justifica a ocorrência de uma alta frequência média.

Em contrapartida, as bases LMDR, CIFAR-10 e VTR apresentaram índices elevados, refletindo a expressiva variabilidade e riqueza de detalhes de seus cenários, além de demonstrarem similaridade estatística em diversas métricas avaliadas.

No que tange ao desempenho empírico, a arquitetura CNN confirmou sua robustez e capacidade de

extração de características, alcançando os maiores F1-Scores na maioria das bases de dados. Contudo, na base ISM, composta por termogramas de módulos solares, a CNN apresentou um desempenho inferior ao da SVM. Esse resultado evidencia a alta complexidade inerente a esse conjunto de dados e a necessidade de estudos mais aprofundados para essa tarefa específica, rompendo com a expectativa inicial de superioridade absoluta da rede neural.

Ao cruzar os dados, constatou-se que não é possível estabelecer com segurança uma correlação estatística entre as métricas de complexidade e o desempenho de classificação apoiando-se estritamente na regressão linear e no coeficiente de Spearman. Apesar disso, todas as métricas, com exceção da Frequência Média, sugerem uma tendência de proporcionalidade direta com o desempenho dos classificadores, caso a base MNIST seja desconsiderada. O comportamento da base MNIST justifica-se por sua natureza atípica; trata-se de um conjunto quase binarizado que entrega altíssimas taxas de acerto, indo na contramão da lógica mensurada pelas métricas de complexidade avaliadas.

Essa tendência observada nos demais conjuntos vai ao encontro do entendimento comum na área de aprendizado de máquina, em que bases de dados mais ricas em informações e detalhes tendem a fornecer insumos melhores para o treinamento e, conseqüentemente, para o desempenho dos algoritmos.

Em suma, conclui-se que as métricas de complexidade são ferramentas analíticas com um potencial valioso para a caracterização prévia de bases de dados, auxiliando na compreensão de suas propriedades estruturais. O sucesso da classificação, entretanto, demonstrou ser um fenômeno multifatorial. Ele depende não apenas da complexidade inerente da imagem, mas fortemente da adequação do algoritmo ao domínio do problema e de fatores não abordados neste estudo, como a sobreposição de classes e a interferência de ruídos. Por fim, a tendência de correlação identificada demanda investigações futuras com um volume maior de bases de dados, a fim de conferir maior precisão e robustez estatística, dado que uma amostra de cinco conjuntos é considerada limitada para generalizações definitivas.

Como proposta de continuidade, planeja-se aplicar a metodologia em um conjunto mais diversificado de bases de dados. Ao variar o grau de complexidade e diversidade das amostras, busca-se obter uma avaliação mais robusta e precisa sobre as tendências constatadas.

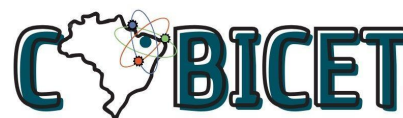


AGRADECIMENTOS

Com profunda gratidão, agradeço à minha mãe pelo apoio incondicional. Seu carinho e incentivo foram luz constante no meu caminho. À minha incrível esposa e parceira de vida, que esteve ao meu lado com conselhos sábios e apoio nos momentos mais desafiadores. Ao meu orientador e amigo, Honovan, pela generosidade em compartilhar seu conhecimento e pela paciência durante toda esta jornada. Ao admirável professor Dr. Rogério, cuja dedicação e carisma sempre me inspiraram e cuja ajuda esteve presente sempre que precisei. Por fim, ao meu coorientador, Thales, cuja sabedoria e disponibilidade foram fundamentais, por oferecer orientações valiosas e contribuir de forma efetiva para este trabalho.

REFERÊNCIAS

- BACKES, A. R. Estudos de métodos de análise de complexidade em imagens. Tese (Doutorado) - Universidade de São Paulo, 2010.
- BRANCHAUD-CHARRON, F.; ACHKAR, A.; JODOIN, P.-M. Spectral metric for dataset complexity assessment. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, p. 3215-3224, 2019.
- CAGAS, W. et al. Medical imaging complexity and its effects on gan performance. In: Proceedings of the Asian Conference on Computer Vision, p. 207-217, 2024.
- CANNY, J. A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence, Ieee, n. 6, p. 679-698, 2009.
- CHUVIECO, E. Fundamentals of satellite remote sensing: An environmental approach. CRC press, 2020.
- COELHO, R. C.; COSTA, L. d. F. The box-counting fractal dimension: does it provide an accurate subsidy for experimental shape characterization? if so, how to use it? Anais, 1995.
- DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). v. 1, p. 886-893, 2005.
- DURMUS, D. Spatial frequency and the performance of image-based visual complexity metrics. IEEE Access, IEEE, v. 8, p. 100111-100119, 2020.
- FENG, T. et al. Ic9600: A benchmark dataset for automatic image complexity assessment. IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE, v. 45, n. 7, p. 8577-8593, 2022.
- GONZALEZ, R. E. W. R. C. Processamento Digital de Imagens. 3a edição. ed. Pearson, 2009.
- HAYKIN, S. Neural networks and learning machines, 3/E. Pearson Education India, 2009.
- KHAN, T. M.; NAQVI, S. S.; MEIJERING, E. Leveraging image complexity in macro-level neural network design for medical image segmentation. Scientific Reports, Nature Publishing Group UK London, v. 12, n. 1, p. 22286, 2022.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.
- KRIZHEVSKY, A.; HINTON, G. et al. Learning multiple layers of features from tiny images. Toronto, ON, Canada, 2009.
- LECUN, Y. et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE, Ieee, v. 86, n. 11, p. 2278-2324, 2002.
- LIU, S. et al. Contrastive learning for image complexity representation. arXiv preprint arXiv:2408.03230, 2024.
- MILLENDORF, M.; OBROPTA, E.; VADHAVKAR, N. Infrared solar module dataset for anomaly detection. In: Proc. Int. Conf. Learn. Represent. v. 3, 2020.
- NAJJAR, R. Redefining radiology: a review of artificial intelligence integration in medical imaging. Diagnostics, MDPI, v. 13, n. 17, p. 2760, 2023.
- OLIVEIRA, M. P. et al. Image descriptors fusion to enhance data representation for classical classifiers applied to small datasets. In: Anais do XXVII Encontro Nacional de Modelagem Computacional e XV Encontro de Ciência e Tecnologia de Materiais. Ilhéus, Brazil: Even3, 2024.



PIETIKÄINEN, M. Local binary patterns. Scholarpedia, v. 5, n. 3, p. 9775, 2010.

RAHANE, A. A.; SUBRAMANIAN, A. Measures of complexity for large scale image datasets. In: 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC). p. 282-287, 2020.

SHANNON, C. E. A mathematical theory of communication. The Bell system technical journal, Nokia Bell Labs, v. 27, n. 3, p. 379-423, 1948.

SILVA, R. A. B. da et al. Dimensão fractal aplicada à análise de imagens térmicas do úbere de vacas saudáveis e com mastite subclínica. Brazilian Journal of Development, v. 6, n. 6, p. 37812-37822, 2020.

SZELISKI, R. Computer vision: algorithms and applications. Springer Nature, 2022.

VASSAKIS, K.; PETRAKIS, E.; KOPANAKIS, I. Big data analytics: applications, prospects and challenges. Mobile big data: A roadmap from models to technologies, Springer, p. 3-20, 2017.

VATANDOUST, M. Evaluation of the image datasets complexity using convolutional autoencoders. 2024.