

# Avaliação de Classificadores de Imagem Multi-rótulo

Igor Goulart Cabral

<sup>1</sup>Programa de Pós-Graduação em Informática  
Pontifícia Universidade Católica de Minas Gerais (PUC Minas)

icabral@sga.pucminas.br

*Abstract.*

*Resumo.*

## 1. Introdução

As redes neurais compõem uma parte importante no campo da Inteligência artificial. Essas arquiteturas são essenciais para a evolução dos algoritmos de aprendizado profundo, contribuindo para o progresso importante na análise, validação e implementação de dados complexos. As metodologias utilizadas, tem sido empregadas em diversos domínios como saúde, educação, entretenimento e finanças, demonstrando sua capacidade de transformar a sociedade.

A inteligência artificial (IA) é uma área dedicada à criação de algoritmos capazes de aprender. Ele aborda diversas metodologias baseadas na lógica, algoritmos de busca e estruturas de raciocínio probabilístico. [Prince 2023]. O Machine Learning representa um ramo onde os modelos são desenvolvidos para aprender a partir de dados, ajustando parâmetros matemáticos para tomar decisões..[Prince 2023]

No campo do Machine Learning, as redes neurais profundas se destacam, e a metodologia empregada em seu desenvolvimento é chamada Deep Learning. Estes modelos têm demonstrado uma capacidade incrível de lidar com cenários complexos, sendo amplamente aplicados em tarefas como tradução automática de idiomas, reconhecimento e busca de imagens, e no desenvolvimento de assistentes virtuais e chatbots.[Prince 2023]

A avaliação de sistemas computacionais, incluindo aqueles que usam redes neurais profundas, envolvem diversas metodologias de teste white-box e black-box. O teste white-box busca pelo conhecimento da estrutura interna da arquitetura ou de seus dados de treinamento para elaborar cenários, e muitas vezes, essa abordagem é inviável devido a restrições de acesso e preocupações com a privacidade dos dados usados.[Wang et al. 2024]. Os testes de black-box, no entanto estão ganhando atenção no mundo acadêmico, uma vez que não tem a necessidade de acesso as arquiteturas ou aos dados de treino, gerando grandes estudos sobre estratégias para sua aplicação.[Wang et al. 2024]

Este trabalho tem como objetivo principal investigar a aplicabilidade do método LV-CIT, proposto por [Wang et al. 2024], que se destaca como uma abordagem eficiente para testar classificadores de imagem baseados em DNNs, em especial na avaliação da capacidade desses modelos em lidar com correlações entre múltiplos objetos. Portanto, este estudo visa replicar e estender o experimento original de [Wang et al. 2024], apresentando novos datasets e arquiteturas de redes neurais profundas para validar a acurácia e a generalização do método LV-CIT em diferentes situações.

## 2. Trabalhos Relacionados

A avaliação de sistemas baseados em Redes Neurais Profundas (DNNs) é uma área de pesquisa que tem atraído atenção, com diversos métodos de testes sendo citados na literatura. Entre eles, o Teste de Interação Combinatória (CIT), que surge como uma metodologia inovadora para avaliar como as diferentes variáveis de entrada de um modelo interagem e afetam seu desempenho. O CIT é relevante em cenários de alta dimensionalidade, onde a identificação de impacto das combinações de entrada na saída é importante. Essa abordagem não apenas revela relações e dependências existentes entre as variáveis, mas também oferece insights valiosos sobre a estrutura dos dados, o que contribui para a melhoria do desempenho e da confiabilidade do modelo. [Prince 2023]

A aplicação do CIT em testes de DNNs é ampla, abrangendo tanto abordagens white-box quanto black-box. Em testes white-box, o CIT tem sido utilizado na criação de conjuntos de dados que visam capturar interações complexas de características de entrada, como, por exemplo, na exploração de combinações de operações de transformação de imagem para diversificar imagens de treinamento. No entanto, a crescente preocupação com a segurança e a privacidade dos dados, que muitas vezes impossibilita o acesso a detalhes internos da arquitetura da rede ou seus dados de treinamento, então o foco da pesquisa tem mudado para metodologias de teste black-box [Wang et al. 2024], onde o CIT também se mostra promissor.

Um desafio encontrado nos testes de classificadores de imagens é a avaliação de modelos multi-rótulos, nos quais a detecção e o correto tratamento de correlações entre múltiplos objetos presentes em uma imagem são importantes. Para abordar essa complexidade e aprimorar a avaliação do desempenho desses sistemas, métodos específicos de teste têm sido desenvolvidos. [Wang et al. 2024]

Nesse contexto, o método LV-CIT [Wang et al. 2024] foi introduzido com o objetivo de aprimorar a avaliação do desempenho de classificadores de imagens multi-rótulos. O LV-CIT se diferencia dos demais por modelar cada rótulo como um parâmetro de entrada binário, o que permite uma análise das combinações de valores de rótulo por meio da geração de um array de cobertura de valor de rótulo t-way. Esta estratégia confere ao LV-CIT uma capacidade robusta de detecção de erros, incluindo erros de excesso, e de incompatibilidade, como destacado por seus autores[Wang et al. 2024].

Apesar da eficácia apresentada pelo LV-CIT na literatura, a validação de sua robustez em outros contextos é importante. Assim o trabalho desenvolvido nesse artigo é explorar a aplicação do LV-CIT em um ambiente black-box utilizando novos datasets e diferentes arquiteturas de redes neurais profundas. O objetivo é replicar e estender os experimentos originais de [Wang et al. 2024], comparando os resultados obtidos e buscando validar a acurácia e a aplicabilidade do método em uma quantidade maior de cenários, identificando modelos que apresentem o melhor desempenho sob essas novas condições de teste.

## 3. Experimentos

Devido a um problema com o carregamento de certas dependências e versões dos softwares necessários não especificados nos arquivos de Git ou no artigo, não foi possível até o presente momento realizar os experimentos. Algo que já foi esperado que acontecesse pelos próprios autores do artigo original.

## **4. Resultados**

## **5. Conclusões e Trabalhos Futuros**

### **References**

Prince, S. J. (2023). *Understanding deep learning*. MIT press.

Wang, P., Hu, S., Wu, H., Niu, X., Nie, C., and Chen, L. (2024). A combinatorial interaction testing method for multi-label image classifier. In *2024 IEEE 35th International Symposium on Software Reliability Engineering (ISSRE)*, pages 463–474. IEEE.