Universidade de Pernambuco
Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

Proposta de Dissertação de Mestrado

Área: Computação Inteligente

Título: Desenvolvimento de Estratégias Combinadas de Aumento de Dados e Injeção de Ruído em Redes Neurais Profundas para Generalização Robusta e Equitativa em Cenários Fora da Distribuição

Orientador: Diego Marconi Pinheiro Ferreira Silva (dmpfs@ecomp.poli.br)
Co-Orientador: Carmelo José Albanez Bastos Filho (carmelo.filho@upe.br)

Descrição:

As Redes Neurais Profundas são, de fato, vulneráveis a ataques adversariais, destacando a necessidade urgente de estratégias para melhorar a generalização robusta dos modelos [1,2]. Estudos recentes mostram que a integração de dados com injeção de ruído durante o treinamento da rede ajuda a fortalecer a resiliência do modelo contra várias formas de corrupções [3,4,5].

Algumas preocupações surgem em relação ao overfitting [6]. Isso ocorre quando um modelo se torna excessivamente ajustado aos dados de treinamento, dificultando a generalização. A regularização explícita induzida por injeções de ruído combate o overfitting ao penalizar conteúdo de alta frequência, promovendo trajetórias de treinamento mais estáveis [7]. Essa preferência pela estabilidade ajuda a melhorar a generalização e a mitigar os riscos de overfitting.

As estratégias de aumento de dados desempenham um papel crucial na regularização do modelo. Estudos recentes mostraram que a combinação de estratégias de aumento de dados é mais eficaz do que fazer poucas modificações nos dados [1]. As estratégias comuns de aumento [8], combinadas com a injeção de ruído, ainda são negligenciadas, mas emergem como uma melhoria significativa para alcançar robustez em modelos de redes neurais. Essa combinação não apenas aumenta a diversidade dos dados de treinamento, mas também introduz uma melhor regularização, ajudando a prevenir o overfitting e a melhorar a robustez do modelo contra ataques adversariais.

O termo "fora da distribuição" (OOD) é frequentemente usado de forma ambígua, causando confusão e o risco de superestimar o progresso no tratamento de dados OOD. A OOD pode empregar, por exemplo, distribuições transformadas [9]. Essas distribuições consistem em dados com corrupções comuns de imagens do mundo real, fornecendo insights sobre a robustez do modelo contra ataques adversariais.

Além disso, estudos recentes que investigam técnicas de validação cruzada para estimar erros de previsão revelaram que intervalos de confiança padrão derivados da validação cruzada podem ter cobertura muito abaixo do nível desejado. Espera-se que os intervalos padrão de validação cruzada exibam desempenho aprimorado com o aumento da regularização [10].

Este projeto de pesquisa busca investigar a seguinte questão: Como regularizar redes neurais profundas com combinações de estratégias de aumento de dados e injeção de ruído para melhorar a generalização e a robustez para dados fora da distribuição de forma equitativa? Os objetivos do projeto são investigar redes neurais profundas regularizadas com estratégias combinadas de aumento de dados usando injeção de ruído, incorporar injeção de ruído em arquiteturas de aprendizado profundo comumente usadas e introduzir novos métodos estatísticos para analisar o

comportamento de erro de cobertura dos intervalos de confiança na captura de erros de previsão.

Referências Bibliográficas:

[1] L. Li and M. Spratling, “Data augmentation alone can improve adversarial train- ing,” arXiv preprint arXiv:2301.09879, 2023.

[2] D. Hendrycks and T. Dietterich, “Benchmarking neural network robustness to common corruptions and perturbations,” Proceedings of the International Conference on Learning Representations, 2019.

[3] T. S. Nazaré, G. B. P. da Costa, W. A. Contato, and M. Ponti, “Deep convolutional neural networks and noisy images,” in Progress in Pattern Recognition, Im- age Analysis, Computer Vision, and Applications: 22nd Iberoamerican Congress, CIARP 2017, Valparaíso, Chile, November 7–10, 2017, Proceedings 22, pp. 416– 424, Springer, 2018.

[4] U. Barros,P.Rocha,M.Oliveira,A..Ribeiro,R. Monteiro, and D. Pinheiro, “Regularizing neural networks with noise injection for classifica- tion of brain tumor in magnetic resonance imaging,” in 2023 IEEE Latin American Conference on Computational Intelligence (LA-CCI), pp. 1–6, 2023.

[5] A. Camuto, M. Willetts, U. Simsekli, S. J. Roberts, and C. C. Holmes, “Explicit regularization in gaussian noise injections,” Advances in Neural Information Pro- cessing Systems, vol. 33, pp. 16603–16614, 2020.

[6] X. Ying, “An overview of overfitting and its solutions,” in Journal of physics: Con- ference series, vol. 1168, p. 022022, IOP Publishing, 2019.

[7] C. M. Bishop, “Training with noise is equivalent to tikhonov regularization,” Neural computation, vol. 7, no. 1, pp. 108–116, 1995.

[8] Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” Journal of big data, vol. 6, no. 1, pp. 1–48, 2019.

[9] S. Farquhar and Y. Gal, “What’out-of-distribution’is and is not,” in NeurIPS ML Safety Workshop, 2022.

[10] S. Bates, T. Hastie, and R. Tibshirani, “Cross-validation: what does it estimate and how well does it do it?,” Journal of the American Statistical Association, pp. 1–12, 2023.