

# **CURRICULUM LEARNING E OUTRAS ESTRATÉGIAS DE TREINAMENTO PARA REMOÇÃO DE RUÍDOS EM IMAGENS**

**João Gabriel Sasseron Roberto Amorim**

**Fernando Pereira dos Santos**

Universidade do Estado de São Paulo

jgsasseron@usp.br fernando\_persan@alumni.usp.br

## **Objetivos**

Elaborar uma arquitetura eficaz de Denoising Autoencoder (DAE) é o primeiro passo no desenvolvimento de modelos de aprendizado de máquina robustos. Este tipo de rede neural tem como objetivo remover ruídos dos dados de entrada através de uma função de reconstrução, comparando os dados de saída com os dados originais sem ruídos. Ademais, é crucial selecionar hiperparâmetros e camadas da rede com precisão para garantir que o modelo aprenda representações significativas dos dados.

Além do treinamento convencional com todos os dados disponíveis, o uso do Curriculum Learning (CL) pode melhorar o desempenho do DAE, permitindo à rede desenvolver uma compreensão gradual e refinada dos dados. No entanto, é essencial comparar essa abordagem para avaliar seu impacto no modelo. Uma análise rigorosa das métricas de desempenho pode revelar os benefícios do CL. Se demonstrar vantagens significativas sobre os métodos tradicionais, podemos concluir que o treinamento com CL é benéfico para os DAE.

## **Métodos e Procedimentos**

Um primeiro passo é determinar a arquitetura base para o estudo de remoção de ruídos. Como resultado, foi selecionada uma arquitetura de rede neural, com 8 camadas,

cada uma com 8 filtros de tamanho 3x3 como modelo a ser empregado. Na sequência, implementações relacionadas ao CL são necessárias. Este tipo de técnica suplementar ao treinamento convencional é composta por duas funções: pacing e scoring. A primeira função determina o ritmo de incorporação dos dados ao treinamento, começando com uma parcela das imagens e aumentando gradativamente até ter todo o conjunto. A função de scoring objetiva pontuar o nível de dificuldade da instância, sendo assim, exemplos considerados mais fáceis são fornecidos desde o início do treinamento e os mais difíceis gradativamente. Assim, temos uma simulação direta da diretriz de estudos em uma disciplina, por exemplo. Em nosso estudo, aplicamos diferentes funções de pacing (linear, log, ladder e ladderlog), assim como scoring: RMSE e CHISC (dos Santos and Ponti, 2019).

Por fim, foi realizado o treinamento do modelo para cada nível de intensidade de ruído e para cada scoring function, utilizando o processo de otimização Adam (Kingma and Ba, 2014). Então, o treinamento foi conduzido ao longo de 300 épocas, cada uma composta por 2000 iterações, com uma taxa de aprendizado de 0.001.

## **Resultados**

Nas Tabelas de 1 a 4 podemos ver os resultados obtidos por meio do treinamento convencional (constante) e por diversas funções de pacing. Assim, como as diferentes

performances considerando diferentes níveis de ruídos e scoring functions.

Métrica	Constante	Linear	Log	Ladder	Ladder Log
RMSE	0,1855	0,1748	<b>0,1713</b>	0,1762	0,1741
PSNR	62,951	63,427	<b>63,610</b>	63,367	63,465
SSIM	0,5360	0,5590	<b>0,5734</b>	0,5487	0,5557

Tabela 1: Aplicação com 10 de ruído, scoring function RMSE.

Métrica	Constante	Linear	Log	Ladder	Ladder Log
RMSE	0,1855	<b>0,1735</b>	0,1803	0,1768	0,1737
PSNR	62,951	<b>63,500</b>	63,177	63,316	63,492
SSIM	0,5360	<b>0,5593</b>	0,5485	0,5503	0,5558

Tabela 2: Aplicação com 10 de ruído, scoring function CHISC.

Podemos notar que a aplicação de CL influenciou de maneira positiva o aprendizado. À medida que houve a diminuição dos valores obtidos pela métrica RMSE, com o melhor cenário obtido pela scoring function RMSE com uma diminuição de 0.014155 na tabela 1 e o pior cenário obtido pela CHISC com uma diminuição de 0,011966 na tabela 2.

Métrica	Constante	Linear	Log	Ladder	Ladder Log
RMSE	0,2110	0,2118	0,2125	0,2159	<b>0,2104</b>
PSNR	61,760	61,731	61,707	61,579	<b>61,793</b>
SSIM	0,4693	0,4641	0,4555	0,4508	<b>0,4663</b>

Tabela 3: Aplicação com 30 de ruído, scoring function RMSE.

Métrica	Constante	Linear	Log	Ladder	Ladder Log
RMSE	0,2110	<b>0,2099</b>	0,2119	0,2138	<b>0,2099</b>
PSNR	61,760	61,813	61,731	61,658	<b>61,807</b>
SSIM	0,4693	0,4550	0,4639	0,4579	<b>0,4716</b>

Tabela 4: Aplicação com 30 de ruído, scoring function CHISC.

A técnica CL obteve resultados satisfatórios para remoção do ruído gaussiano para todas as scoring functions RMSE e CHISC. Pois, ficou evidente que os resultados obtidos superaram aqueles alcançados pelo método constante. Além disso, notamos que a métrica RMSE demonstrou seu desempenho mais destacado em cenários com níveis de ruído mais baixos, como ilustrado na Tabela 1. Por outro lado, em situações com taxas de ruído mais elevadas, a métrica CHISC mostrou resultados mais satisfatórios. Isso fica claro na análise da Tabela 4, na qual observamos uma diminuição de 0,001147 no RMSE, um aumento de 0,05345 no PSNR e um aumento de 0,014311 no SSIM. Estes resultados consolidam a comprovação do aprimoramento no desempenho.

## Conclusões

Com base nestes resultados, concluímos que o método de treinamento de redes neurais CL é aplicável aos DAEs, com resultados iguais ou superiores ao método tradicional. Além disso, observamos que o CL é sensível aos hiperparâmetros utilizados, com desempenho superior ao usar a métrica CHISC para níveis mais elevados de ruído e a métrica RMSE para índices de ruído inferiores. Outro fator de destaque é que o uso da metodologia oferece redução de tempo de treinamento, pois os dados são incorporados gradativamente, fazendo com que a quantidade de imagens das primeiras épocas seja bem reduzida.

## Referências

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

dos Santos, F. P. and Ponti, M. A. Homogeneity index as stopping criterion for anisotropic diffusion filter. In Computer Analysis of Images and Patterns: 18th International Conference, CAIP 2019, Salerno, Italy, September 3–5, 2019, Proceedings, Part II 18, pages 269–280. Springer, 2019.