

**ARQUITETURA ÓTIMA PARA UMA REDE NEURAL DO TIPO CNN:
ESTUDO DE CASO NA ESTIMATIVA DE VIDA ÚTIL REMANESCENTE EM MOTORES
DE AERONAVES¹**

SOUZA, Pedro ²
POVEDA, Pedro Fernando³
OLIVEIRA, Mauro Machado de⁴

RESUMO

O presente estudo aborda o desafio da otimização de arquiteturas de Redes Neurais Convolucionais (CNNs) para a estimativa da Vida Útil Remanescente (RUL, do inglês *Remaining Useful Life*) em sistemas eletromecânicos. O objetivo principal foi identificar, por meio de uma investigação empírica estruturada, uma melhor configuração de CNN, que apresente um desempenho preditivo superior ao da arquitetura de referência no conjunto de dados C-MAPSS. Partindo de uma arquitetura básica, a metodologia de otimização foi conduzida para encontrar uma melhor combinação entre as configurações, resultados que foram avaliados por meio da métrica da Raiz do Erro Quadrático Médio (RMSE) e seu desvio padrão. Utilizou-se o conjunto de dados referentes à degradação de motores a jato de aeronaves obtidos por meio de simulações experimentais e disponibilizados no repositório de dados da Agência Espacial Americana (NASA). A configuração final obtida apresentou valores de RMSE inferiores e maior estabilidade em comparação com o modelo de referência inicial.

Palavras-chave: Vida Útil Remanescente, Redes neurais convolucionais, Predição de falhas, Monitoramento da condição eletromecânica real.

INTRODUÇÃO

As Redes Neurais Convolucionais (CNNs) representam uma classe de modelos de aprendizado profundo, originalmente popularizados na análise de imagens bidimensionais, uma aplicação consolidada no trabalho seminal de LeCun (1998). Sua principal característica é a capacidade de extrair hierarquias de características complexas dos dados, uma propriedade que tem se mostrado igualmente eficaz na análise de dados sequenciais, como as séries temporais. Um dos campos de aplicações em que esses modelos são particularmente relevantes é na engenharia de prognósticos e na manutenção preditiva, um componente essencial da Indústria 4.0. A estimativa da Vida Útil Remanescente (RUL, do inglês *Remaining Useful Life*) de componentes críticos é um objetivo central, estimar com precisão o instante provável de ocorrência de falha é crucial

¹ Trabalho vinculado ao projeto de Iniciação Científica "Predição de falhas e estimativa de vida útil remanescente em elementos de máquinas rotativas pelo monitoramento da condição eletromecânica real", fomentado pela INSTITUIÇÃO FEDERAL DE SÃO PAULO por meio do programa PIBIFSP.

² Graduando em Engenharia Mecânica; IFSP; São Paulo; São Paulo; p.segura@aluno.ifsp.edu.br;

^{3,4} Professor EBT no IFSP, Doutor em Tecnologia Nuclear pelo IPEN/USP; São Paulo; São Paulo

para garantir a segurança operacional, minimizar paradas não programadas e otimizar os custos de manutenção em setores de alto risco, como o aeroespacial.

A complexidade dos dados de sensores nesta área torna a tarefa um estudo de caso relevante para a otimização de arquiteturas de CNN. No entanto, o desempenho de uma rede neural está intrinsecamente ligado à sua arquitetura, a forma como as camadas são organizadas, os tipos de operações que executam e como os dados fluem nelas. A arquitetura de uma rede neural CNN é determinada por um conjunto de parâmetros configuráveis, que engloba o número e as dimensões dos filtros, as funções de ativação e as metodologias de regularização, a quantidade e a natureza das camadas, que podem ser convolucionais, responsáveis pela extração de características dos dados, camadas de *pooling* que efetuam a redução da dimensionalidade enquanto preservam informações cruciais, as camadas densas estabelecem conexões entre todos os neurônios da camada precedente para assimilar combinações complexas de características e as camadas de saída que produzem o resultado final da rede neural.

Na ausência de diretrizes universalmente estabelecidas para a concepção da arquitetura ideal para problemas de prognóstico, a investigação empírica e sistemática torna-se uma abordagem metodológica essencial. A exploração realizada por meio da modificação controlada de parâmetros e da reconfiguração da arquitetura de camadas permite mapear a relação entre as escolhas de arquitetura e o desempenho preditivo da rede. Esse processo, embora intensivo, gera informações sobre quais características arquitetônicas são mais sensíveis e impactantes para a captura dos padrões de degradação.

OBJETIVOS

O presente trabalho tem como objetivo geral investigar e identificar uma arquitetura otimizada da CNN que seja especificamente eficaz para a tarefa de estimativa da RUL no conjunto de dados C-MAPSS (*Commercial Modular Aero-Propulsion System Simulation*, cujo modelo de falha foi estabelecido nos trabalhos fundamentais de Saxena (2008). Por meio de uma série de simulações computacionais controladas. Para alcançar este objetivo principal, foram estabelecidos os seguintes objetivos específicos:

- Metodologia de Otimização Estruturada: Aplicar metodologia empírica de otimização em duas fases (análise de sensibilidade univariada e otimização combinatória) para um processo sistemático e reprodutível.
- Avaliar Sensibilidade do Modelo: Analisar o impacto dos parâmetros de arquitetura e treinamento no desempenho do modelo, especificamente para a CNN profunda e séries temporais dos conjuntos de dados.
- Propor Configuração Superior: Identificar e propor uma arquitetura final que supere a referência da MathWorks em Raiz do Erro Quadrático Médio (RMSE) e seu desvio padrão (StDev).

METODOLOGIA

A. Pré-processamento

Inicialmente todo o *dataset*, contendo os valores encontrados por cada um dos 21 sensores, foi submetido a uma sequência de passos de pré-processamento que incluem: Seleção de características baseada na análise de prognóstico, normalização dos dados dos sensores, padronização do aprendizado, por meio da limitação da RUL a um teto de 150 ciclos operacionais dos motores simulados, focando o modelo na fase de degradação mais acentuada.

B. Arquitetura do Modelo de Referência

A arquitetura base consistia em cinco blocos convolucionais com configurações crescentes de número e tamanho de filtros. A saída desses blocos era processada por uma camada densa e uma camada de saída.

C. Processo Experimental de Otimização

Dada as numerosas possibilidades de configurações, a otimização do modelo foi conduzida por meio de um processo de investigação empírica estruturado em duas fases principais. Esta abordagem foi projetada para identificar eficientemente os parâmetros de maior impacto e, subsequentemente, explorar as interações benéficas entre eles.

- **ANÁLISE DE SENSIBILIDADE UNIVARIADA:** O objetivo desta fase inicial foi isolar e quantificar o impacto de cada parâmetro individual no desempenho do modelo. Partindo da arquitetura de referência, cada parâmetro de interesse foi variado dentro de um intervalo de valores pré-definido, enquanto todos os outros eram mantidos fixos. Este processo foi aplicado a parâmetros da arquitetura e do treinamento. Ao final desta fase, foram identificados os valores que, para cada parâmetro isolado, resultaram no melhor desempenho.
- **OTIMIZAÇÃO COMBINATÓRIA E AJUSTE FINO:** Nesta etapa, o objetivo foi investigar as sinergias e interações entre os parâmetros mais promissores. Para isso, foram avaliados novos modelos que combinavam as melhores configurações identificadas. Este processo permitiu refinar o modelo e convergir para a arquitetura final.

A Tabela 1 resume a descrição das configurações utilizadas nos processos de análise de sensibilidade e otimização combinatória.

Tabela 1 – Configurações de arquitetura e treinamento

MiniBatchSize	Divide o conjunto de dados utilizados para treinamento, a fim de diminuir o custo computacional
MaxEpochs	Número de vezes em que o algoritmo irá analisar o subconjunto de dados e alterar os pesos do modelo
InitialLearnRate	Taxa de aprendizado inicial do modelo. A taxa de aprendizado influencia diretamente na intensidade da alteração dos pesos
LearnRateSchedule	Define um esquema para a manipulação da taxa de aprendizado inicial, podendo variar ao longo do treinamento
Optimizer	Determina como os resultados obtidos via Retropropagação serão utilizados na alteração dos pesos
GradientThreshold	Limita o valor dos gradientes calculado em processos de Retropropagação, impedindo que a alteração dos pesos seja abrupta
Shuffle	Funções de embaralhamento do conjunto de dados, a fim de prevenir o hiperajuste do modelo
FilterSize	Dimensiona o tamanho do filtro (em matrizes de pixels) que desliza sobre os dados de entrada. Os filtros são detectores de características que atuam por meio de operações matemáticas
NumFilters	Determina o número de filtros que irá percorrer os dados durante o funcionamento da rede neural, esses filtros não necessariamente buscam as mesmas características
Padding	Estabiliza qual estratégia será utilizada para a manutenção das dimensões dos dados. Quando atravessam camadas, na ausência de um mecanismo de Padding, a resolução dos dados diminui

Fonte: Autores

D. Métricas de Avaliação de Desempenho

Para avaliar o desempenho do modelo em cada teste, a métrica primária selecionada foi a RMSE.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

RESULTADOS E DISCUSSÃO

A metodologia de otimização sistemática permitiu identificar uma arquitetura de CNN com desempenho significativamente superior ao modelo de referência. A configuração otimizada alcançou um RMSE de 14,07 (com desvio padrão de 7,07) no conjunto de dados FD001, representando uma redução de quase 20% no erro de predição em comparação com o RMSE de 17,77 do modelo base disponibilizado pela MathWorks. Na Tabela 2 são apresentados os melhores resultados, com respectivas as configurações utilizadas, comparados com a configuração básica inicial.

Tabela 2 – Comparação entre os resultados obtidos

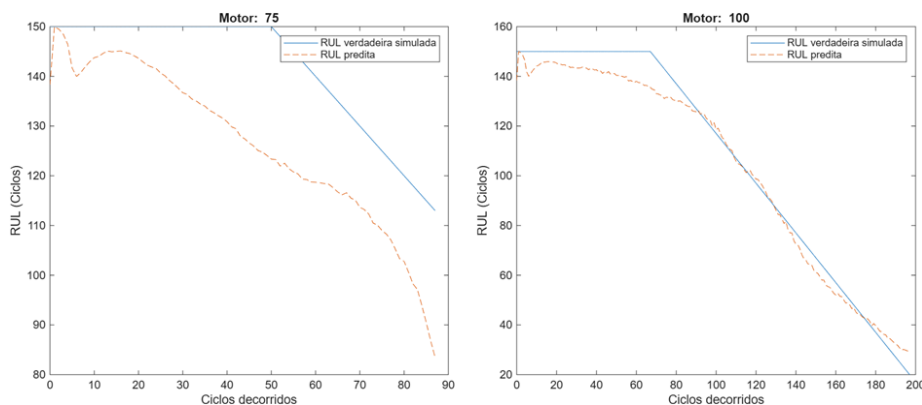
REDE (CLASSIFICAÇÃO ARQUITETURAS)	PARÂMETROS CONFIGURADOS (ARQUITETURA DA REDE)							RESULTADOS	
	Max Epochs	Mini batches	Initial LearnRate	Gradient Threshold	Shuffle	Optimizer	camada (nº filtros / tamanho / preenchimento)	RMSE (Mean)	StDev (RMSE)
Base Set	40	16	0,01	1	never	Adam	1ª camada (32 / 5 / causal) 2ª camada (64 / 7 / causal) 3ª camada (128 / 11 / causal) 4ª camada (256 / 13 / causal) 5ª camada (512 / 15 / causal)	17,77	7,2
I (melhor)	30	8	0,006	5	every-epoch	sgdm	1ª camada (64 / 10 / same) 2ª camada (100 / 21 / same) 3ª camada (2048 / 140 / causal) 4ª camada (1024 / 140 / same) 5ª camada (1024 / 7 / causal)	14,13	7,07
II (2ª melhor)							1ª camada (64 / 10 / same) 2ª camada (100 / 21 / same) 3ª camada (2048 / 140 / causal) 4ª camada (1024 / 140 / causal) 5ª camada (1024 / 15 / causal)	14,22	7,68
III (3ª melhor)							1ª camada (64 / 10 / same) 2ª camada (64 / 7 / causal) 3ª camada (2048 / 140 / causal) 4ª camada (256 / 13 / causal) 5ª camada (1024 / 15 / causal)	14,28	6,42

Parâmetro LearnRateSchedule configurado em todas as arquiteturas: Piecewise

Fonte: Autores

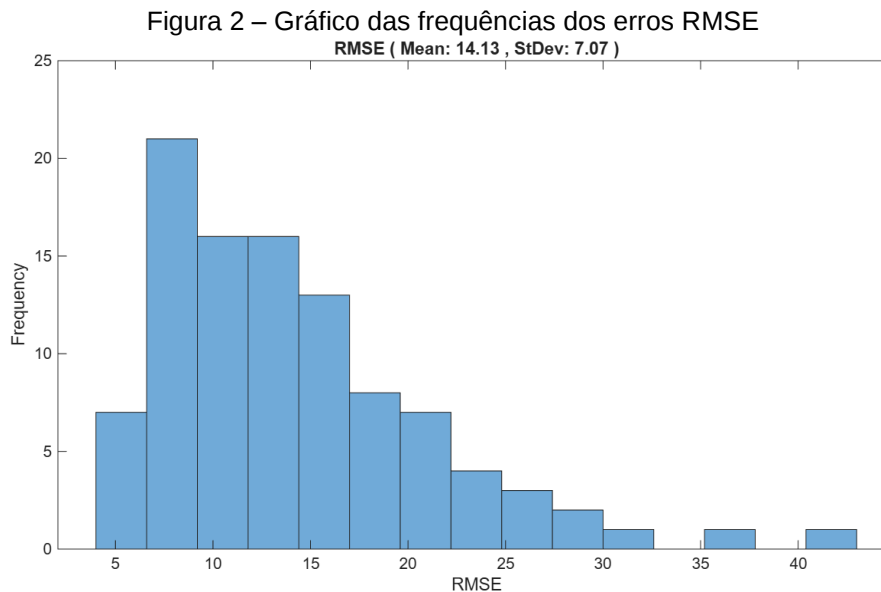
Os gráficos dos motores 75 e 100, apresentados na Figura 1, são demonstrações da predição simulada real e a obtida por meio da configuração I. Ressalta-se que, apesar de uma diminuição do RMSE, motores específicos ainda apresentam uma diferença entre o simulado e o real.

Figura 1 – Gráfico das predições dos motores 75 e 100



Fonte: Autores

O gráfico da Figura 2 mostra os resultados das estimativas para as RUL's, obtidas com a utilização da configuração I da rede neural, para as 100 unidades (motores), com as respectivas frequências, erro RMSE e desvio padrão.



Fonte: Autores

CONSIDERAÇÕES FINAIS

Este trabalho demonstrou a eficácia de uma metodologia de otimização sistemática para Redes Neurais Convolucionais aplicada à estimativa da vida útil remanescente. Ao aplicar uma abordagem estruturada de análise de sensibilidade e combinação de parâmetros, foi possível desenvolver um modelo aprimorado. Esse modelo reduziu o erro de predição em cerca de 20% no *dataset* C-MAPSS FD001, em comparação com a implementação de referência.

A principal contribuição consiste na identificação de uma arquitetura não-intuitiva e de alto desempenho, destacando a importância de explorar configurações que divergem de arquiteturas de progressão monotônica. Conclui-se que a otimização direcionada, mesmo que focada em um subconjunto de dados, é uma estratégia para extrair o potencial de modelos de aprendizado em aplicações de engenharia complexas. Como trabalhos futuros, a metodologia proposta pode ser estendida para os demais *datasets* do C-MAPSS (FD002, FD003, FD004) para validar a generalização dos resultados.

REFERÊNCIAS

- LECUN, Yann *et al.* **Gradient-Based Learning Applied to Document Recognition**. IEEE, 1998.
- MATHWORKS. **Turbofan Engine Degradation Simulation**. Natick, MA: The MathWorks, Inc., 2025. Disponível em: <https://www.mathworks.com/help/predmaint/ug/turbofan-engine-degradation-simulation.html>. Acesso em: 11 jul. 2025.
- SAXENA, Abhinav *et al.* **Damage Propagation Modeling for Aircraft: Engine Run-to-Failure Simulation**. Denver: IEEE, 2008.