



UNIVERSIDADE FEDERAL DO AMAZONAS  
FACULDADE DE TECNOLOGIA  
CURSO DE GRADUAÇÃO EM ENGENHARIA MECÂNICA

---



RODRIGO MARQUES DE ALMEIDA NASCIMENTO

**Estimação das emissões de  $O_2$  e  $CO_2$  na combustão de gás natural  
em um forno industrial através de imagens de chamas e redes  
neurais**

Manaus

2022

RODRIGO MARQUES DE ALMEIDA NASCIMENTO

**Estimação das emissões de O<sub>2</sub> e CO<sub>2</sub> na combustão de gás natural  
em um forno industrial através de imagens de chamas e redes  
neurais**

**Versão Corrigida**

Trabalho de Conclusão de Curso apresentado à Faculdade de Tecnologia da Universidade Federal do Amazonas como parte dos requisitos necessários para obtenção do grau de Engenheiro Mecânico.

Orientador: Prof. Dr. Gustavo Cunha da Silva Neto

Manaus

2022

## Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

N244e Nascimento, Rodrigo Marques de Almeida  
Estimação das emissões de O<sub>2</sub> e CO<sub>2</sub> na combustão de gás natural em um forno industrial através de imagens de chamas e redes neurais / Rodrigo Marques de Almeida Nascimento . 2022  
66 f.: il. color; 31 cm.

Orientador: Gustavo Cunha da Silva Neto  
TCC de Graduação (Engenharia Mecânica) - Universidade Federal do Amazonas.

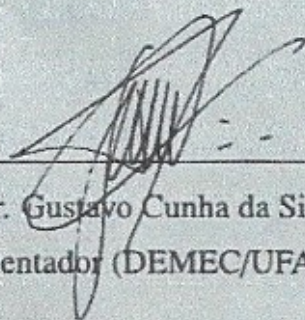
1. Combustão. 2. Gás Natural. 3. Redes Neurais. 4. Sensores Virtuais. 5. Visão Computacional. I. Silva Neto, Gustavo Cunha da. II. Universidade Federal do Amazonas III. Título

**Estimação das emissões de O<sub>2</sub> e CO<sub>2</sub> na combustão de gás natural  
em um forno industrial através de imagens de chamas e redes  
neurais**

**RODRIGO MARQUES DE ALMEIDA NASCIMENTO**

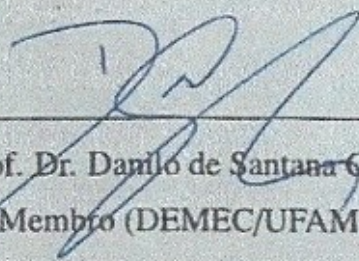
Trabalho de Conclusão de Curso apresentado à Faculdade de Tecnologia da Universidade Federal do Amazonas como parte dos requisitos necessários para obtenção do grau de Engenheiro Mecânico.

Aprovada por:



---

Prof. Dr. Gustavo Cunha da Silva Neto  
Orientador (DEMEC/UFAM)



---

Prof. Dr. Danilo de Santana Chui  
Membro (DEMEC/UFAM)



---

Prof. Dr. Nilton Pereira da Silva  
Membro (DEMEC/UFAM)

Manaus-AM, 22 de junho de 2022.

Aos meus pais, que sempre estiveram ao meu lado.

## **AGRADECIMENTOS**

A Deus, em primeiro lugar, por sempre me guiar pelo melhor caminho e permitir que eu seja tudo o que sou.

A meus pais, Wouner e Marilia, por toda a dedicação e paciência que tornaram o caminho até aqui muito mais fácil e prazeroso para mim.

À minha futura esposa, Amanda Luiza, por ser meu ponto de paz e por sempre me fazer crer que sou capaz.

A um grande amigo, Flávio Dias, pelo companheirismo durante essa longa jornada e por todas as risadas que demos juntos.

Ao meu orientador, Prof. Dr. Gustavo Neto, pela confiança depositada em mim; pela oportunidade de trabalharmos juntos e por todos os conselhos.

A todos que contribuíram direta ou indiretamente para o fim desta jornada.

*“It is so shocking to find out how many people do not believe that they can learn, and how many more believe learning to be difficult.”*

*(Frank Herbert)*

## RESUMO

NASCIMENTO, Rodrigo Marques de Almeida. **Monitoramento da combustão de gás natural em fornos de refino por meio de processamento digital de imagens e aprendizado de máquina. 2022. 66f.** Trabalho de Conclusão de Curso (Engenharia Mecânica) - Universidade Federal do Amazonas, Manaus, 2022.

A combustão é atualmente o principal método utilizado para produzir energia na matriz energética mundial. Através da queima de petróleo, carvão mineral e gás natural a humanidade supre suas necessidades energéticas enquanto migra equilibradamente para fontes de energias renováveis. Deriva-se desse contexto então a necessidade de um controle eficiente da combustão, um processo-chave na redução das consequências do efeito estufa, e, por procedência, de uma medição precisa das emissões de  $O_2$  e  $CO_2$  em processos de combustão. Devido às condições hostis de processo como a alta temperatura, o difícil acesso à área controlada e o risco de explosão, a utilização de sensores diretamente na chama muitas vezes é inviabilizada e o monitoramento por imagens tem se mostrado uma alternativa conveniente. Neste trabalho, um dispositivo de inferência baseado em redes neurais artificiais que utiliza imagens de chamas monocromáticas capturadas por uma câmera CCD (*charged-coupled device* - dispositivo de carga acoplada) é projetado para estimar as emissões de  $O_2$  e  $CO_2$  em um forno a gás natural. O método proposto demonstra ser promissor, pois estima sem atraso, diferente de analisadores de gases tradicionais, e com boa precisão nas razões de equivalência testadas (raiz do erro quadrático médio abaixo de 2%).

**Palavras-chave:** Combustão; Gás Natural; Redes Neurais; Sensores Virtuais; Visão Computacional.



# ABSTRACT

NASCIMENTO, Rodrigo Marques de Almeida. **Natural gas combustion monitoring in refining furnaces by digital image processing and machine learning. 2022. 66f.** Undergraduate thesis (Mechanical Engineering) - Federal University of Amazonas, Manaus, 2022.

Combustion is currently the main method to produce energy in the world energy matrix. Through the burning of oil, coal and natural gas, mankind supplies its energy needs while migrating in a balanced way to renewable energy sources. From this context comes the need for efficient control of combustion, a key process in reducing the consequences of the greenhouse effect, and hence an accurate measurement of CO<sub>2</sub> and O<sub>2</sub> emissions in combustion processes. Due to the hostile process conditions such as high temperature, difficult access to the controlled area and the risk of explosion, the use of sensors on the flame is often impractical and image monitoring has proven to be a convenient alternative. In this paper, an inference device based on artificial neural networks that uses monochromatic flame images captured by a CCD (charged-coupled device) camera is designed to estimate CO<sub>2</sub> and O<sub>2</sub> emissions in a natural gas furnace. The proposed method seems promising: it estimates without delay, unlike traditional gas analyzers, and with good accuracy on the tested equivalence ratios (root mean square error below 2%)

**Keywords:** Combustion; Computer Vision; Natural Gas; Neural Networks; Virtual Sensors.

## LISTA DE ILUSTRAÇÕES

Figura 1	Arquitetura proposta por Liu, Fan e Chen (2017). . . . .	17
Figura 2	Processo de discretização de uma imagem. . . . .	20
Figura 3	Imagens digitais após o processo de amostragem. . . . .	21
Figura 4	Imagem digital após os processos de Amostragem e Quantização. .	22
Figura 5	Imagem digital após os processos de Amostragem e Quantização. .	23
Figura 6	Relação entre inteligência artificial, aprendizado de máquina e aprendizado profundo. . . . .	27
Figura 7	Histórico CASP de primeiros colocados na precisão média de mo- delagem livre. . . . .	28
Figura 8	Aprendizado de máquina <i>vs</i> <i>softwares</i> tradicionais. . . . .	29
Figura 9	Representação visual de camadas ocultas. . . . .	29
Figura 10	O processo de aprendizado de máquina. . . . .	31
Figura 11	Exemplo de melhoria nas representações de uma rede neural. . . .	33
Figura 12	Exemplo da operação de convolução. . . . .	35
Figura 13	Exemplo de hierarquia de padrões. . . . .	36
Figura 14	Exemplo da operação de <i>pooling</i> . . . . .	36
Figura 15	Diagrama do método proposto. . . . .	38
Figura 16	Fornalha horizontal utilizada nos experimentos. . . . .	39
Figura 17	Emissões de O <sub>2</sub> e CO <sub>2</sub> obtidas durante o ensaio sob diferentes ra- zões de equivalência. . . . .	40
Figura 18	Emissões de O <sub>2</sub> e CO <sub>2</sub> ao longo do experimento, com tempo morto. .	42
Figura 19	Emissões de O <sub>2</sub> e CO <sub>2</sub> ao longo do experimento, atrasados em 114s. .	42
Figura 20	Ilustração do processo de recorte das imagens. . . . .	43
Figura 21	Ilustração do filtro de média móvel com 30 amostras. . . . .	44

Figura 22	Diagrama do macroprocesso da arquitetura proposta. . . . .	46
Figura 23	Entrada escolhida para gerar exemplos de representações intermediárias da rede, imagem de chama no instante $t = 419,8s$ . . . . .	48
Figura 24	Representações intermediárias das camadas convolutivas pré-treinamento, coloridas artificialmente para facilitar visualização. . . . .	49
Figura 25	Representações intermediárias das camadas convolutivas pós-treinamento, coloridas artificialmente para facilitar visualização. . . . .	50
Figura 26	Representações intermediárias das camadas convolutivas pós-treinamento, não manipuladas. . . . .	51
Figura 27	Representações intermediárias da rede completa. . . . .	52
Figura 28	Evolução do erro quadrático médio ao longo do período de treinamento. . . . .	54
Figura 29	Conjunto de aprendizado antes (esquerda) e depois (direita) da remoção de dados para compor o conjunto de testes. . . . .	55
Figura 30	Teor de $O_2$ e $CO_2$ amostrado vs estimado ao longo do experimento. . . . .	56
Figura 31	Histograma da raiz do erro quadrático médio amostrado vs estimado no conjunto de testes. . . . .	56
Figura 32	Divisão do conjunto de dados entre aprendizado e testes após nova segmentação. . . . .	57
Figura 33	Divisão do conjunto de dados entre aprendizado e testes para o segundo cenário. . . . .	58
Figura 34	Teor de $O_2$ e $CO_2$ amostrado vs estimado no cenário II. . . . .	59
Figura 35	Divisão do conjunto de dados entre aprendizado e validação para o terceiro cenário. . . . .	59
Figura 36	Teor de $O_2$ e $CO_2$ amostrado vs estimado ao longo do experimento. . . . .	60
Figura 37	Teor de $O_2$ e $CO_2$ amostrado vs estimado ao longo do experimento. . . . .	60

## LISTA DE ABREVIATURAS E SIGLAS

CASP	<i>Critical Assessment of Structure Prediction</i> , ou avaliação crítica de predição de estruturas
CCD	<i>Charge-coupled device</i> , ou dispositivo de carga acoplada, é um sensor para captação de imagens
CO <sub>2</sub>	Dióxido de carbono, comumente chamado de gás carbônico
CO	Monóxido de carbono
DAE	<i>Denoising Autoencoder</i> , ou autocodificadores com filtragem de ruídos
GAN	<i>Generative Adversarial Network</i> , ou redes neurais adversariais generativas
GB	<i>Gigabytes</i>
GPC	<i>Gaussian Process Classifier</i> , ou classificação por processos gaussianos
GPU	<i>Graphics Processing Unit</i> , ou placa de vídeo
IPT	Instituto de Pesquisas Tecnológicas do Estado de São Paulo
MAE	<i>Mean absolute error</i> , ou erro absoluto médio
MB	<i>Megabytes</i>
NO <sub>x</sub>	Óxidos de nitrogênio
O <sub>2</sub>	Dioxigênio, comumente chamado de oxigênio
RMSE	<i>Root mean squared error</i> , ou raiz do erro quadrático médio
SGD	<i>Stochastic gradient descent</i> , ou gradiente descendente estocástico
SO <sub>2</sub>	Dióxido de enxofre
SSAE	<i>Stacked Sparse Autoencoder</i> , ou autocodificadores esparsos empilhados
SVM	<i>Support Vector Machine</i> , ou máquina de vetores de suporte
VGG16	Arquitetura popular de rede neural convolucional com 16 camadas

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
1.1	Motivação . . . . .	13
1.2	Objetivos . . . . .	15
1.2.1	Objetivo geral . . . . .	15
1.2.2	Objetivos específicos . . . . .	15
1.3	Descrição do trabalho . . . . .	15
<b>2</b>	<b>REVISÃO DA LITERATURA</b>	<b>16</b>
<b>3</b>	<b>ELEMENTOS DE PROCESSAMENTO DIGITAL DE IMAGENS</b>	<b>20</b>
3.1	Definição . . . . .	20
3.2	Amostragem e quantização . . . . .	21
3.3	Digitalização de uma imagem monocromática . . . . .	22
<b>4</b>	<b>ELEMENTOS DE ÁLGEBRA LINEAR E OTIMIZAÇÃO</b>	<b>24</b>
4.1	Método do gradiente . . . . .	24
4.2	Tensores . . . . .	25
<b>5</b>	<b>ELEMENTOS DE INTELIGÊNCIA ARTIFICIAL</b>	<b>27</b>
5.1	Inteligência artificial . . . . .	27
5.2	Aprendizado de máquina . . . . .	29
5.3	Aprendizado profundo . . . . .	32
5.4	Redes neurais convolucionais . . . . .	35
<b>6</b>	<b>MÉTODO DE MONITORAMENTO DE CHAMAS ATRAVÉS DE IMA-</b>	

<b>GENS E REDES NEURAIAS ARTIFICIAIS</b>	<b>38</b>
6.1 Método . . . . .	38
6.2 Coleta de dados . . . . .	39
6.2.1 Aparato experimental . . . . .	39
6.2.2 Ensaaios . . . . .	40
6.3 Processamento dos dados . . . . .	41
6.4 Processamento das imagens . . . . .	42
6.4.1 Recorte . . . . .	43
6.4.2 Filtro de média móvel . . . . .	43
6.5 Método de inferência . . . . .	44
6.5.1 Arquitetura proposta . . . . .	46
6.5.2 Visualizações de camadas intermediárias . . . . .	47
<b>7 RESULTADOS</b>	<b>53</b>
7.1 Métricas . . . . .	54
7.2 Validações . . . . .	55
7.2.1 Cenário I . . . . .	55
7.2.2 Cenário II . . . . .	58
7.2.3 Cenário III . . . . .	59
<b>8 CONCLUSÃO</b>	<b>62</b>
8.1 Sugestões para trabalhos futuros . . . . .	63
<b>Referências</b>	<b>64</b>

# 1 INTRODUÇÃO

## 1.1 Motivação

No contexto geopolítico da década de 2010 e do início da década de 2020 há uma preocupação crescente por fontes de energia renováveis capazes de sustentar o desenvolvimento humano. Isso ocorre pois o mundo ainda depende da queima de combustíveis fósseis que, além de serem finitos, são extremamente danosos ao meio ambiente e a saúde humana. Não resta dúvida de que essa fonte de energia gerou um enorme progresso socioeconômico nos últimos séculos, como: a disseminação em massa do uso de energia elétrica; possibilitou as revoluções industriais; o avanço dos automóveis e da logística. Esse avanço acelerado, entretanto, trouxe consigo a poluição do ar para níveis alarmantes e agregou um custo alto à vida: a exposição à poluentes do ar relacionados à combustão vem causando problemas cardiorrespiratórios e neuro-cognitivos, como morte prematura (BROOK *et al.*, 2010); ansiedade; depressão e redução de QI (PAYNE-STURGES *et al.*, 2019). Como reporta a BBC (2016), estima-se que mais de 5,5 milhões de pessoas morrem anualmente de forma prematura devido à poluição do ar, segundo dados reunidos como parte do projeto *Global Burden of Disease* (“Peso Global das Doenças”, em tradução livre).

Ainda que os problemas mencionados sejam alertantes, o processo de migração de combustíveis fósseis para os renováveis é feito de forma lenta devido ao custo de adaptação na matriz energética, ao ainda elevado custo de geração em escala e à inadaptabilidade dessa em relação a variações da demanda. Há ainda, como fator exacerbante, o fato que países em desenvolvimento necessitam de uma fonte de energia de baixo custo para se manterem competitivos (AKIWUMI, 2021). Sob esta ótica, o gás natural, atualmente terceira fonte de energia mais utilizada no mundo, como cita BP (2021), representa uma ótima alternativa como fonte de transição, visto que apresenta baixo custo, é menos prejudicial à saúde e ao meio ambiente, já está inserido na matriz

energética atual e sua combustão pode ser regulada baseada na demanda do momento.

Observa-se então a importância da combustão de gás natural feita de forma eficiente, pois libera mínimas quantidades de gases nocivos à atmosfera como  $\text{NO}_x$ ,  $\text{SO}_2$ ,  $\text{CO}_2$ ,  $\text{CO}$  e praticamente nenhum particulado quando comparada à queima de carvão e derivados do petróleo (LIANG *et al.*, 2012). No entanto, alcançar eficiência elevada e reduzir a quantidade de combustível consumido necessita de um controle robusto e preciso que muitas vezes não se encontra presente devido à baixa capacidade de monitoramento das chamas, como notam Ballester e Hernández (2010). Isso ocorre pois o processo de transferência de calor por convecção é extremamente complexo e possui muitas fontes de variação, como mudanças na qualidade do combustível ou variações de concentração e temperatura do ar. Para representar esse tipo de processo, Ballester e Hernández (2010) sugerem que é mais comum o uso de modelos baseados nas emissões dos queimadores em estratégias de controle, no entanto, estes apresentam por natureza um atraso na medição de estado que causa desperdício de combustível, como observa Chen *et al.* (2012), além de não ser próprio para câmaras multi-chamas, isto é, que possuem diversos queimadores e apenas uma exaustão (BALLESTER; HERNÁNDEZ, 2010).

À frente de tal desafio, o monitoramento por imagens de chama se mostra uma alternativa conveniente: são não-intrusivos, conseqüentemente não influenciando no processo de combustão; pode ter um custo muito menor do que sensores para altas temperaturas e proveem características da chama individualmente e de forma instantânea. Este trabalho busca explorar mais o conceito, desenvolvendo um método de inferência através da utilização de imagens de chamas capturadas por uma câmera CCD e redes neurais artificiais para estimar as emissões de  $\text{O}_2$  e  $\text{CO}_2$  em uma fornalha de combustão de gás natural.



## 1.2 Objetivos

### 1.2.1 Objetivo geral

Desenvolver um método de inferência baseado em redes neurais capaz de estimar emissões de  $O_2$  e  $CO_2$  da combustão de gás natural em um forno industrial a partir de imagens de chamas geradas sob diferentes razões de equivalência, isto é, a proporção da mistura ar-combustível.

### 1.2.2 Objetivos específicos

Os objetivos específicos para desenvolvimento do método são:

- a) Pré-processar imagens a fim de reduzir custo computacional;
- b) Sincronizar imagens obtidas às amostras de emissões;
- c) Identificar a arquitetura de rede neural mais adequada ao problema;
- d) Treinar modelo utilizando um subconjunto das imagens obtidas e dados sincronizados;
- e) Validar método de inferência utilizando conjunto diferente de imagens.

## 1.3 Descrição do trabalho

O trabalho está formatado, de forma geral, em 4 seções: *i)* a revisão da literatura, onde será explorado o estado da arte e os últimos avanços no campo de pesquisa; *ii)* a introdução de tópicos especiais, onde serão revisados todos os conceitos necessários para entendimento do tema; *iii)* o método de monitoramento, em que serão apresentados mais a fundo os procedimentos de tratamento de dados e aprendizado de máquina aplicados no trabalho; e *iv)* os resultados e conclusões do trabalho, incluindo sugestões para trabalhos futuros.

## 2 REVISÃO DA LITERATURA

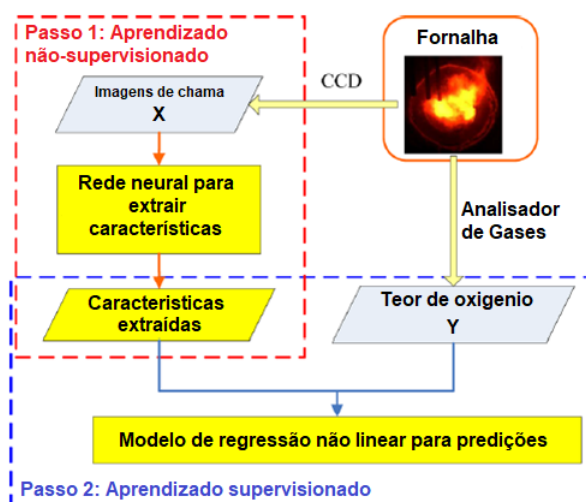
Para uma combustão de melhor desempenho geral, se faz necessário o monitoramento e controle contínuo da estabilidade da chama (SILVA NETO, 2021). Devido à relevância do campo, estudos já foram feitos correlacionando parâmetros geométricos e de luminosidade presentes em imagens da chama e parâmetros físico-químicos utilizados no controle da combustão, como temperatura de emissão por Toth, Garami e Csordás (2017), estabilidade da chama por Han *et al.* (2020) e teor de oxigênio nas emissões por Liu, Fan e Chen (2017). Alguns dos estudos mais relevantes para este trabalho foram listados abaixo como forma de indicar o quão atual é o tema citado.

Embora seja possível coletar amostras de parâmetros de combustão como os teores de  $O_2$ ,  $CO_2$  e  $NO_x$  na exaustão através de dispositivos analisadores de gás, as amostras obtidas dessa forma são bastante atrasadas (CHEN *et al.*, 2012), isto é, as emissões da chama demoram algum tempo após o processo de queima até chegarem nos dispositivos de análise. Esse atraso faz com que o controlador em malha fechada compense excessivamente o sistema, resultando em um controle pouco robusto e lento (LIU; FAN; CHEN, 2017). Para solucionar os problemas de atraso na amostragem, Chen, Chan e Cheng (2013) identificaram o modelo de uma fornalha a óleo que faria previsões do teor de oxigênio nas emissões utilizando regressão não-linear com Processos Gaussianos e validaram-no através de redes neurais artificiais. No artigo a rede neural apresentou resultados muito similares aos obtidos através de regressão, no entanto os autores não entram em mais detalhes sobre a arquitetura utilizada ou as entradas processadas.

Com o passar de alguns anos Liu, Fan e Chen (2017) revisitaram o problema, porém aplicando redes neurais profundas e imagens obtidas através de uma câmera CCD capaz de capturar 1 imagem a cada 5 segundos. O método proposto consiste em utilizar duas redes neurais artificiais conforme arquitetura representada na Figura 1, onde a primeira rede neural extrai características arbitrárias das imagens e a segunda

rede neural processa essas características para inferir o teor instantâneo de oxigênio na exaustão do processo.

Figura 1: Arquitetura proposta por Liu, Fan e Chen (2017).



Fonte: Adaptado de Liu, Fan e Chen (2017) pelo Autor.

Os resultados obtidos através desse modelo foram mais precisos do que os obtidos por Chen, Chan e Cheng (2013) anteriormente, demonstrando a eficácia do método, mas não menciona a predição de outras emissões além do  $O_2$ .

No mesmo ano, Toth, Garami e Csordás (2017) publicaram um estudo sobre o uso de redes neurais profundas utilizando parâmetros operacionais e imagens de chama em tempo real para prever a temperatura de saída de uma caldeira de biomassa. A arquitetura proposta encontrou melhores resultados do que as redes utilizadas na época, além de propôr previsões precisas de temperatura para até 28 minutos após a captura da imagem.

Abdurakipov *et al.* (2018) utilizaram redes neurais convolucionais desenvolvidas nos *frameworks* TensorFlow e Keras para classificar regimes de combustão baseados em imagens de chama em um queimador de propano e obtiveram grande sucesso, alcançando uma acurácia de 98% mesmo em um conjunto de imagens com grandes variações. Já em trabalhos mais recentes: i) Omiotek e Kotyra (2021) apresentaram 4 modelos classificatórios baseados na arquitetura de redes neurais VGG16<sup>1</sup>, capazes de analisar e classificar condições do processo de combustão com uma precisão média entre 82% e 98%; e ii) Han *et al.* (2021) propuseram um método utilizando autoco-

<sup>1</sup>Arquitetura de rede neural convolucional com 16 camadas, amplamente utilizada para tarefas de classificação de imagens.

dificadores com filtragem de ruídos (*DAE - Denoising Autoencoder*), Redes Neurais Adversariais Generativas (*GAN - Generative Adversarial Network*) e classificação por processos gaussianos (*GPC - Gaussian Process Classifier*) para alcançar resultados mais precisos. A proposta de Omiotek e Kotyra (2021), assim como a deste trabalho de conclusão de curso, possui a vantagem de não requerer etapas de pré-processamento de dados, design e extração de características de imagem. Já a de Han *et al.* (2021) introduz complexidade ao introduzir etapas ao processo, requerendo mais poder computacional para ganhos marginais de precisão. Apesar de importante, a identificação do regime em que o processo está ocorrendo –estacionário ou transiente– pode não ser suficiente para controlá-lo.

Han *et al.* (2020) propuseram uma rede neural baseada em autocodificadores esparsos empilhados (*SSAE - Stacked Sparse Autoencoder*) para monitorar a estabilidade do processo de combustão de etileno em um forno sob diferentes condições de operação. Uma das maiores vantagens da arquitetura escolhida é o aprendizado não supervisionado, uma característica que permite que o aprendizado ocorra sem a necessidade de informações extras além das imagens, como por exemplo uma tabela com dados referentes à cada amostra. A partir de uma máquina de vetores de suporte (*SVM - Support Vector Machine*) e utilizando métodos estatísticos e de redução de ruídos no conjunto de dados, alcançou-se uma pontuação  $f1$  de 0,99 – métrica entre 0 e 1 que avalia classificadores baseado no número de falso-positivos, falso-negativos e verdadeiro-positivos. Como citam Ballester e Hernández (2010), a estabilidade da chama é um indicador importante de performance do processo de combustão e, de acordo com Han *et al.* (2020), é essencial para controlá-lo de forma ótima. Os autores do texto afirmam que o método desenvolvido pode ser aplicado a novos experimentos sem ser retreinado, mas o artigo não menciona como integrá-lo ao controle da chama.

Através de um modelo autorregressivo de médias móveis com entradas exógenas, Silva Neto *et al.* (2021) desenvolveram um método capaz de prever emissões de  $O_2$  e  $CO_2$  em uma fornalha de gás natural a partir de parâmetros extraídos de imagens de chamas como perímetro, luminosidade, dispersão e coordenadas do centroide. Apesar dos resultados promissores, o método proposto requer etapas de pré-processamento *offline* de dados – como limiarização, segmentação e extração de características – e, por se basear em um modelo autorregressivo, depende do estado do sistema em instantes anteriores, impossibilitando que sejam inferidas as emissões apenas com o instante atual.

Por meio desta revisão pode-se concluir que é possível relacionar imagens de chama com características da combustão através de redes neurais, mas não foram encontrados estudos sobre a correlação entre imagens e emissões de  $O_2$  e  $CO_2$  do processo em tempo real, parâmetros de alta relevância que permitem verificar as condições do processo de combustão (SILVA NETO, 2021). Este é o problema ainda sem solução que procura-se solucionar: obter um mecanismo de inferência baseado em redes neurais capaz de estimar emissões de  $O_2$  e  $CO_2$  a partir de imagens de chama e que possa futuramente ser incluído em um sistema de controle em malha fechada.

Para facilitar a compreensão do leitor, a seguir serão introduzidos alguns assuntos essenciais para o entendimento da pesquisa, tais como processamento digital de imagens, tensores, aprendizado profundo e redes neurais convolucionais.

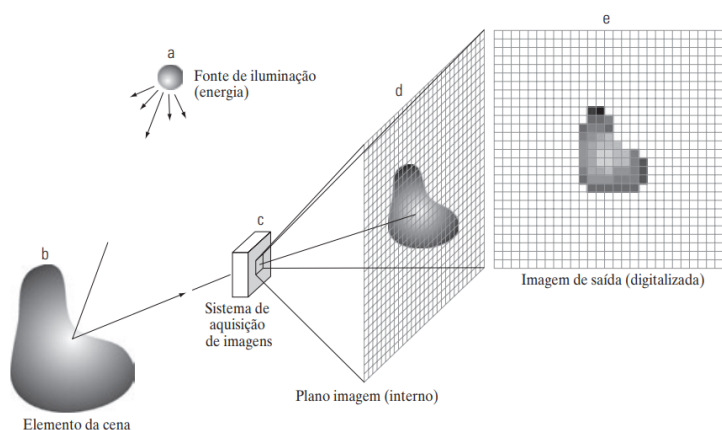
### 3 ELEMENTOS DE PROCESSAMENTO DIGITAL DE IMAGENS

O capítulo introduz a definição de imagens digitais e como são geradas, guiando o leitor pelo processo de discretização. Também é discutido uma das formas mais práticas de representá-las, que será utilizada na Seção 6.4

#### 3.1 Definição

De acordo com Gonzalez (2009), uma imagem pode ser definida como uma função bidimensional,  $f(x, y)$ , onde  $x$  e  $y$  são coordenadas no plano e a amplitude de  $f$  representa a intensidade – ou nível de cinza – do ponto representado por elas. Quando  $x$ ,  $y$  e a amplitude de  $f$  são quantidades discretas e finitas, a imagem pode ser chamada de *imagem digital*, e o par de coordenadas e intensidade de *pixel*. A digitalização de uma imagem então é feita através dos processos de amostragem e quantização, explicados na Seção 3.2, onde se digitalizam a coordenada e a amplitude de um ponto, respectivamente (Figura 2).

Figura 2: Processo de discretização de uma imagem.

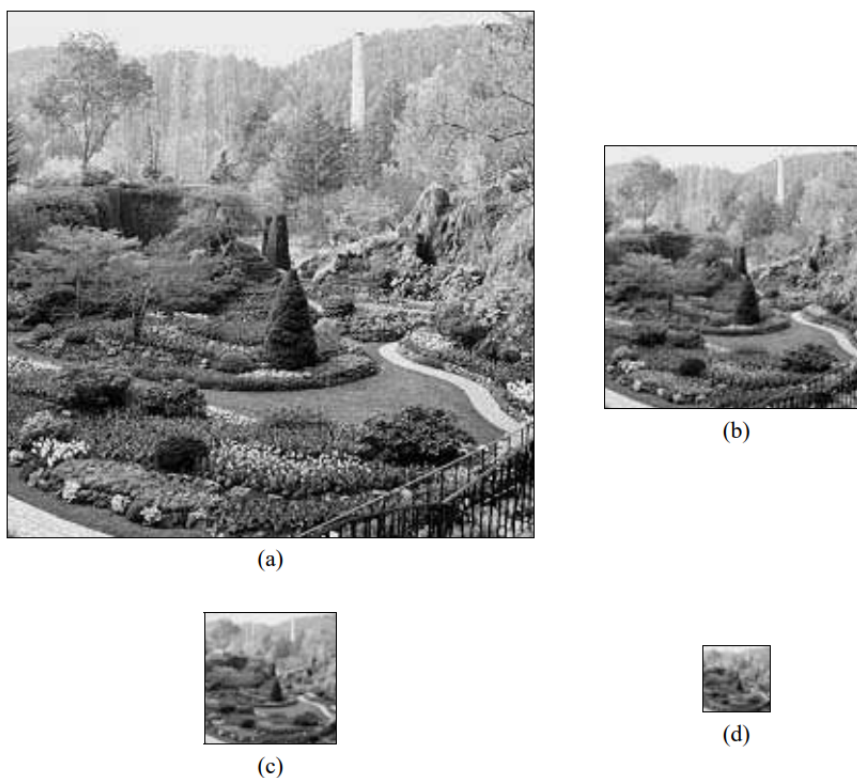


Fonte: Gonzalez (2009)

## 3.2 Amostragem e quantização

No processo de amostragem, obtém-se amostras de um sinal analógico em instantes de tempo discretos conforme uma frequência pré-determinada que deve ser maior que duas vezes a frequência do sinal amostrado (SHANNON, 1949). No caso analisado, o processo de amostragem realiza a discretização espacial da cena capturada, definindo assim o número de *pixels* na imagem digital resultante (FILHO; NETO, 1999). Para facilitar a visualização, a Figura 3 representa 4 imagens digitais com 256 tons de cinza e 256x256, 128x128, 64x64 e 32x32 *pixels*, respectivamente. Percebe-se como as imagens com menos *pixels* vão diminuindo em tamanho e representando cada vez menos detalhes.

Figura 3: Imagens digitais após o processo de amostragem.

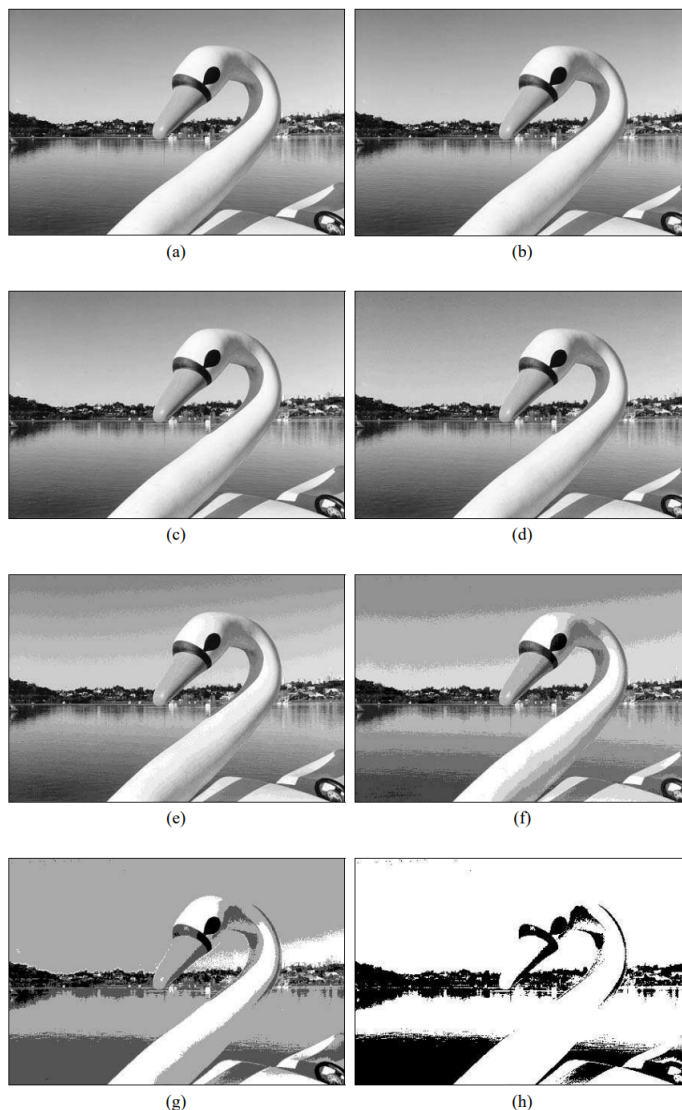


Fonte: Filho e Neto (1999)

Na etapa de quantização, as amostras discretizadas são mapeadas para uma faixa de valores pré-definidos e finitos. No contexto de digitalização de imagens, este processo realiza a discretização em amplitude (FILHO; NETO, 1999), definindo assim o número de cores possíveis na imagem digital resultante. A Figura 4 mostra cópias da mesma imagem digital com 256, 128, 64, 32, 16, 8, 4 e 2 tons de cinza, respectiva-

mente. Pode-se perceber através do exemplo que a redução da quantidade de cores possíveis também reduz a fidelidade da cena original.

Figura 4: Imagem digital após os processos de Amostragem e Quantização.



Fonte: Filho e Neto (1999)

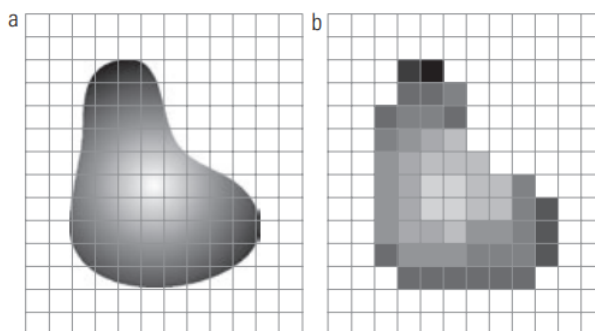
### 3.3 Digitalização de uma imagem monocromática

Naturalmente, o processo de discretização de uma imagem infere perda de dados, como pode ser visto na Figura 5 pela redução do nível de detalhes. Isso ocorre pois o conjunto  $\mathbb{N}^2$  em que estão contidos os elementos da imagem digital é um subconjunto do espaço  $\mathbb{R}^2$  onde está o plano imagem. Ou seja, mesmo utilizando todos os elementos do conjunto  $\mathbb{N}^2$  para representar uma imagem, ainda não se alcançaria uma



cópia integralmente fiel da imagem original. Com um número suficientemente grande de *pixels*, no entanto, a diferença entre as imagens real e digital se torna imperceptível ao olho humano e ganha-se a valiosíssima habilidade de armazenar essas imagens em recursos finitos como fotografias ou memórias eletrônicas.

Figura 5: Imagem digital após os processos de Amostragem e Quantização.



Fonte: Gonzalez (2009)

O armazenamento e a representação de uma imagem digital monocromática, então, podem ser feitos por uma matriz  $M_{i,j}$  como ilustrado na Equação 3.1, onde  $i$  e  $j$  indicam a posição no eixo horizontal e vertical, respectivamente, e  $f$  o valor de intensidade amostrada, em uma escala arbitrária predeterminedada, naquele ponto da imagem.

$$M_{i,j} = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,j-1) \\ f(1,0) & f(1,1) & \cdots & f(1,j-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(i-1,0) & f(i-1,1) & \cdots & f(i-1,j-1) \end{bmatrix} \quad (3.1)$$

Essa representação é especialmente útil pois permite manipular e transformar imagens de forma algébrica e com baixo custo computacional utilizando operações matriciais. Duas transformações que serão vistas no Subcapítulo 6.4 são o recorte, onde a parte da imagem sem valor agregado é descartada, e o filtro de média móvel (ou *imagem média*), que gera uma nova imagem a partir de varias através do calculo da média dos tons de cinza em cada *pixel* da imagem resultante da operação.

## 4 ELEMENTOS DE ÁLGEBRA LINEAR E OTIMIZAÇÃO

O campo de aprendizado profundo se baseia fortemente em alguns conceitos de álgebra linear e cálculo, como vetores, matrizes, tensores, calculo diferencial, continuidade de funções e tantos outros. Para entender como o algoritmo de aprendizado funciona, o leitor precisa deter conhecimento de alguns tópicos que as vezes não são visitados durante a graduação. Com intuito de tornar o trabalho autocontido, os temas necessários e menos comuns serão introduzidos neste capítulo.

### 4.1 Método do gradiente

O método do gradiente, como explicam Carneiro *et al.* (2005) e Korn e Korn (2000), é um modelo de otimização que utiliza o gradiente  $\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  de uma função objetivo  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  diferenciável em um ponto qualquer  $p = (x_1, x_2, \dots, x_n)$  para determinar a direção da busca pelo ponto ótimo. Após encontrar-se a direção da busca, se dá origem a uma nova busca que atenda melhor a função objetivo, até que se encontre um ponto de mínima ou máxima, dependendo da meta. Cada novo ponto é dado pela Equação 4.1.

$$p_{k+1} = p_k + \lambda_k \nabla f(p_k). \quad (4.1)$$

Onde  $\lambda_k$  é um escalar que determina o tamanho do passo na direção ótima. O gradiente  $\nabla f$  é calculado de acordo com a Equação 4.2.

$$\nabla f(p) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial f}{\partial x_n}(p) \end{bmatrix} \quad (4.2)$$

## 4.2 Tensores

Gaelzer (2021) define tensores como objetos geométricos que estabelecem relações lineares entre escalares, vetores e outros tensores. Um tensor pode ser representado por uma matriz multidimensional de valores numéricos, cuja dimensão se dá o nome de posto ou ordem (GAELZER, 2021).<sup>1</sup> O conceito, que a primeira vista pode parecer complexo, se resume a generalização de estruturas algébricas bem conhecidas como escalares (sem índices), vetores (um índice) e matrizes (dois índices) para um número arbitrário de índices. Isto é, um tensor de grau  $n$  em um espaço dimensional  $m$  é uma estrutura algébrica que possui  $n$  índices e  $m^n$  elementos. Para facilitar a visualização do conceito, segue-se alguns exemplos:

- Exemplo 1: Um tensor de grau 0 não possui índices, de forma similar a um escalar.

$$T_0 = \left[ S_0 \right] \quad (4.3)$$

- Exemplo 2: Um tensor de grau 1 possui um único índice, de forma similar um vetor.

$$T_i = \begin{bmatrix} S_0 \\ S_1 \\ \vdots \\ S_{i-1} \end{bmatrix} \quad (4.4)$$

- Exemplo 3: Um tensor de grau 2 possui dois índices, de forma similar a uma matriz.

$$T_i^j = \begin{bmatrix} S_{0,0} & S_{0,1} & \cdots & S_{0,j-1} \\ S_{1,0} & S_{1,1} & \cdots & S_{1,j-1} \\ \vdots & \vdots & \ddots & \vdots \\ S_{i-1,0} & S_{i-1,1} & \cdots & S_{i-1,j-1} \end{bmatrix} \quad (4.5)$$

---

<sup>1</sup>Definições mais formais, além das propriedades, axiomas e transformações que regem os espaços tensoriais podem ser encontradas em Marcus (1975).

- Exemplo 4: Um tensor de grau 3 possui três índices, de forma similar a um vetor de matrizes.

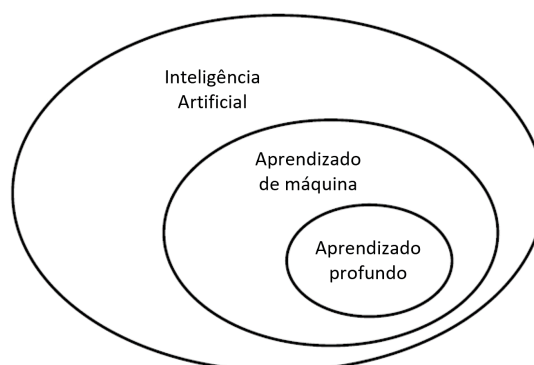
$$T_i^{jk} = \begin{bmatrix} S_{0,0} & S_{0,1} & \dots & S_{0,j-1} \\ S_{1,0} & S_{1,1} & \dots & S_{1,j-1} \\ \vdots & \vdots & \ddots & \vdots \\ S_{i-1,0} & S_{i-1,1} & \dots & S_{i-1,j-1} \end{bmatrix} \quad (4.6)$$

Tensores de grau  $N$  possuem  $N$  índices e podem armazenar  $Z$  escalares, onde  $Z$  é o produto do número de elementos contidos em cada um de seus índices. Nas ciências da computação comumente dá-se o nome de forma (*shape*) para as características de tamanho de um tensor (WHITE, 2021). Por exemplo, um tensor de forma (10, 340, 340, 3) armazena  $10 \times 340 \times 340 \times 3 = 3.468.000$  escalares – o equivalente a 10 imagens digitais de  $340 \times 340$  *pixels* e 3 canais de cores.

## 5 ELEMENTOS DE INTELIGÊNCIA ARTIFICIAL

Antes de introduzir aprendizado profundo, é necessário esclarecer a sua associação com outros dois tópicos que são muito confundidos atualmente: Inteligência artificial e aprendizado de máquina (*machine learning*). A Figura 6 ilustra a relação entre eles de forma simples, mas o leitor precisará de um conhecimento maior, introduzido ao longo deste capítulo, para entender melhor os conceitos propostos no trabalho.

Figura 6: Relação entre inteligência artificial, aprendizado de máquina e aprendizado profundo.



Fonte: Chollet (2021).

### 5.1 Inteligência artificial

Em 1950 o matemático inglês Alan Turing propôs o questionamento se computadores poderiam ser instruídos para replicar o comportamento humano de forma que uma pessoa não soubesse se estava conversando com uma máquina ou outra pessoa (TURING; HAUGELAND, 1950). Isso instigou o desenvolvimento de computadores que pudessem realizar tarefas comumente associadas à seres inteligentes como tomar decisões, aprender com experiências passadas, procurar sentido nas coisas ou até

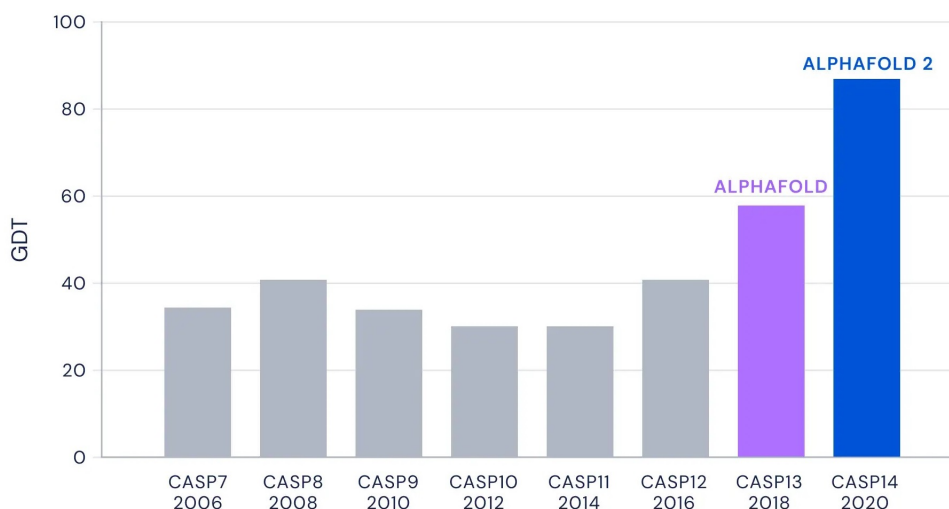
mesmo jogar xadrez (CHOLLET, 2021).

Alguns anos mais tarde, Rich (1983), definiu Inteligência artificial como:

*Inteligência Artificial é o estudo de fazer com que os computadores façam coisas que, no momento, os humanos fazem melhor<sup>1</sup>*

E desde então os progressos nesse campo de pesquisa permitiram que inteligências artificiais desempenhassem atividades de formas mais eficazes e eficientes até mesmo do que humanos, como pode se observar no exemplo recente de um grande avanço no campo da Bioinformática: o sistema *AlphaFold*, desenvolvido pela empresa *DeepMind* para predição de estruturas e enovelamento de proteínas, conquistou o primeiro lugar no CASP – Avaliação Crítica de Predição de Estruturas (*Critical Assessment of Structure Prediction*, em inglês) – atingindo uma média de 92,4% de precisão no teste de distância global, que compara estruturas preditas com estruturas determinadas experimentalmente. Os resultados obtidos pela IA são muito superiores ao estado da arte prévio, que alcançava apenas em torno de 40% de precisão (ALPHAFOLD TEAM, 2021). Divergindo do que ocorre na maioria das linhas de pesquisa, tal avanço não se deu por alguma descoberta no campo de estruturas proteicas, mas por avanços no campo de aprendizado de máquina que permitiram treinar tal inteligência artificial (JUMPER *et al.*, 2021).

Figura 7: Histórico CASP de primeiros colocados na precisão média de modelagem livre.



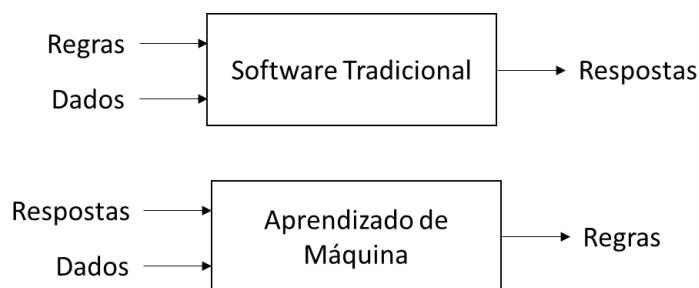
Fonte: AlphaFold Team (2021)

<sup>1</sup>Tradução livre de “Artificial Intelligence is the study of making computers do things which, at the moment, humans do better.”

## 5.2 Aprendizado de máquina

*Machine Learning* – ou aprendizado de máquina – pode ser explicado como um subsistema do campo de inteligência artificial que tenta gerar um modelo de predição de respostas através do aprendizado obtido em um conjunto de entradas e saídas esperadas. A forma de programar esse sistema passa então a ser o oposto da de *softwares* tradicionais: ao invés de se criar um conjunto de regras que gera uma resposta a partir de dados introduzidos, cria-se um algoritmo capaz de analisar entradas possíveis e respostas esperadas para gerar seu próprio conjunto de regras.

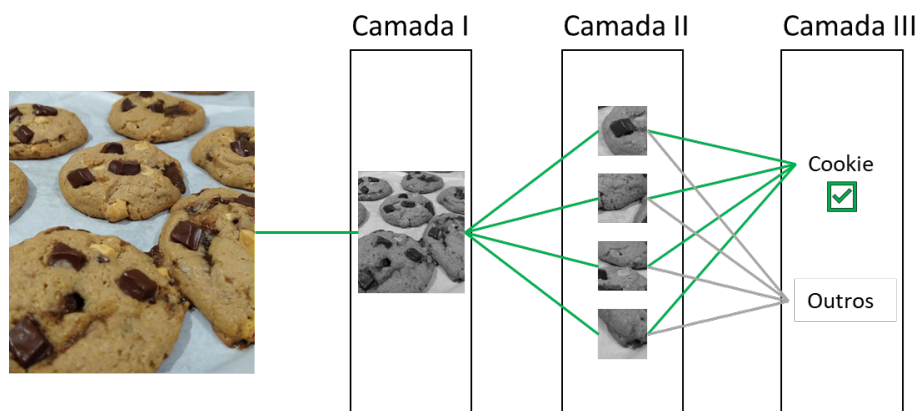
Figura 8: Aprendizado de máquina vs *softwares* tradicionais.



Fonte: Adaptado de Chollet (2021) pelo Autor.

O algoritmo principal para implementar este paradigma são as redes neurais artificiais, nomeada desta forma por imitar como neurônios biológicos se comunicam entre si. Em uma visão macro, estas redes são compostas por camadas (*layers*), estruturas responsáveis por transformar dados, que procuram formas melhores de relacionar os dados de entrada às saídas esperadas.

Figura 9: Representação visual de camadas ocultas.



Fonte: Autor.

Para o computador, uma camada é um objeto composto de um tensor de pesos ( $\vec{W}$ ), um tensor de vieses ( $\vec{b}$ ) e uma função de ativação diferenciável e não-linear ( $f$ ) que opera em uma entrada de formato pré-definido de acordo com a Equação 5.1 (CHOLLET, 2021).

$$\vec{Y} = f(\vec{X} \cdot \vec{W} + \vec{b}). \quad (5.1)$$

Vale notar que, assim como a soma,  $f$  deve ser aplicada elemento-por-elemento, e que quando  $\vec{W}$  e  $\vec{b}$  possuem graus ou formas diferentes a operação de *broadcast*<sup>1</sup> deve ser aplicada antes da soma/produto tensorial. Para fins de conhecimento, as funções de ativação mais comumente utilizadas são o degrau, a rampa, a sigmoide e a tangente hiperbólica devido ao menor custo computacional (SZANDAŁA, 2021).

Uma variante do método do gradiente utilizada em aprendizado de máquina é chamada de Gradiente Descendente Estocástico, onde o cálculo do gradiente é feito baseado em um subconjunto amostral escolhido aleatoriamente do conjunto de dados, ao invés do todo. Isso torna o algoritmo mais rápido devido a menor quantidade de dados para manipular a cada iteração.

Nesse contexto, “aprender” significa encontrar um conjunto de pesos e vieses para todas as camadas da rede que mapeie corretamente o conjunto de testes às saídas esperadas. O que faz dessa atividade algo complexo é o número de parâmetros possíveis – normalmente na casa dos milhões ou até bilhões (PANAGAKIS *et al.*, 2021) – e o fato de que alterar um parâmetro pode mudar o comportamento de todo o resto da rede neural. Dessa forma, a otimização de parâmetros comumente é feita de forma iterativa utilizando retropropagação e o método do gradiente com intuito de aumentar sua precisão e capacidade de generalizar até que seja capaz de gerar respostas promissoras em entradas nunca antes processadas.

Com esses conceitos, espera-se que o leitor esteja mais familiarizado com o necessários de como redes neurais artificiais aprendem e se desenvolvem com o passar de iterações.

No início do treinamento os tensores  $\vec{W}$  e  $\vec{b}$  são preenchidos de forma aleatória para que seja realizado um “passo à frente” (*forward pass*) de dados por todas as camadas da rede neural. Com os resultados dessa passada, uma função de perda  $C = f(W_1, b_1, W_2, b_2, \dots, W_n, b_n)$  calcula quão próximas as predições obtidas estão dos

<sup>1</sup>Para resolver o problema de operações aritméticas entre matrizes ou tensores de grau diferente, a operação de *broadcast* replica o tensor de menor grau ao longo da dimensão incompatível.

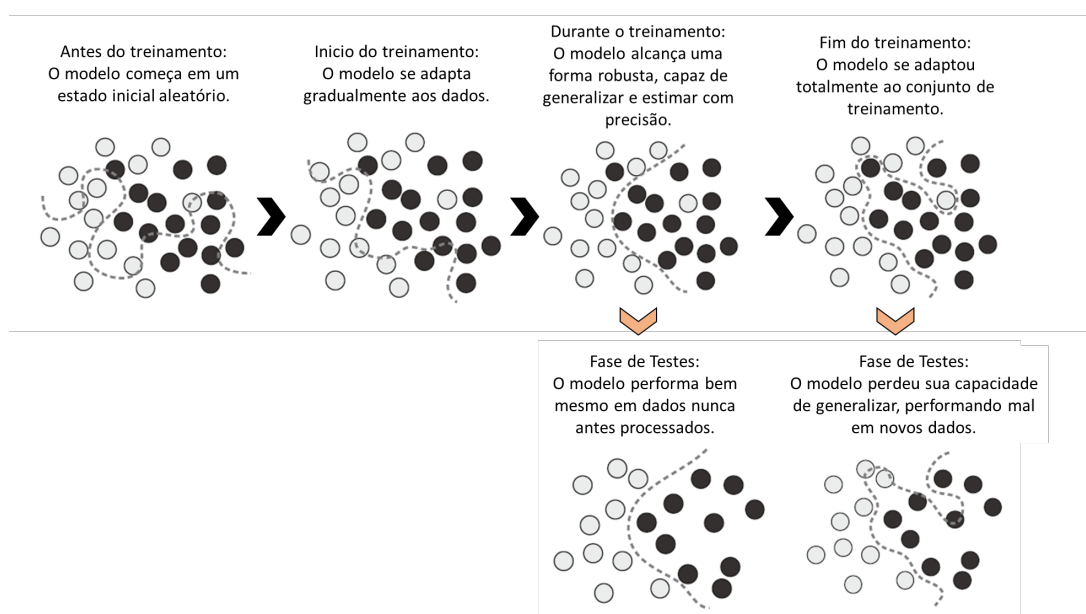


resultados esperados. Logo em seguida ocorre a etapa de retropropagação, onde o aprendizado efetivamente acontece: calcula-se o gradiente  $\vec{\nabla}C$  em respeito a todos os pesos do sistema para definir quais desses devem ser alterados e em qual magnitude. Matematicamente, a operação pode ser descrita como na Equação 5.2, onde  $g$  é a taxa de aprendizado: um escalar com o intuito de acelerar ou refrear o processo de aprendizado para evitar o travamento em mínimos locais da função de perda (CHOLLET, 2021).

$$W_{novo}^{\vec{}} = W_{antigo}^{\vec{}} - \vec{\nabla}C \cdot g \quad (5.2)$$

Conforme as iterações se acumulam e mais retro-propagações são realizadas, o algoritmo de aprendizado de máquina vai se aperfeiçoando e encontra formas melhores de correlacionar entradas e saídas no seu espaço de hipóteses, efetivamente aumentando sua precisão no conjunto de treinamento e reduzindo sua habilidade de generalizar para novos dados, isto é, se contamina com vieses e ruídos do conjunto e retorna resultados ruins para imagens nunca antes visitadas (Figura 10). A essa perda da capacidade de generalizar se dá o nome sobreajuste (*overfit*). Vale notar que, durante o processo iterativo de aprendizado, quando a rede termina de processar todo o conjunto de dados, recomeça-se da primeira amostra e diz-se que uma "época" passou.

Figura 10: O processo de aprendizado de máquina.



Fonte: Adaptado de Chollet (2021) pelo Autor.

## 5.3 Aprendizado profundo

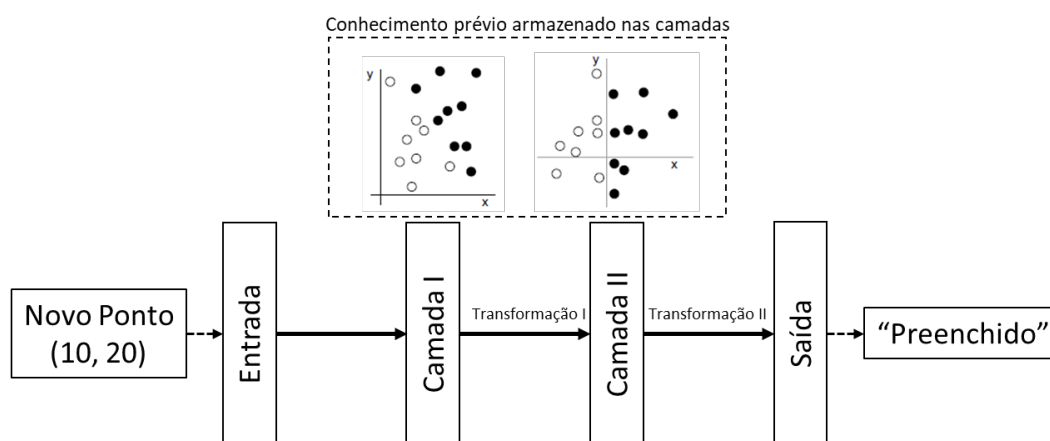
Como nota Chollet (2021), analisando os grandes avanços do campo de redes neurais historicamente, pode-se perceber que estes estão diretamente correlacionados aos avanços em 4 áreas:

- **Hardware:** Desde o início da década de 90, avanços em processadores e, eventualmente, em placas de vídeo aumentaram exponencialmente o número de cálculos por segundo que um computador pode realizar. Hoje em dia, uma única GPU – *Graphics Processing Unit* – topo de linha como a NVIDIA Titan V entrega 110 teraflops de performance, isto é, 110 trilhões de operações ponto-flutuante ou mais de 3400 vezes o número de operações que o Intel Touchstone Delta, o supercomputador mais rápido do mundo em 1990. Esse avanço possibilitou o desenvolvimento de algoritmos mais complexos e o aumento de escala de algoritmos já existentes à época, trazendo assim modelos de redes neurais artificiais com mais capacidade de aprendizado e precisão;
- **Dados:** O compartilhamento de informações tem se tornado cada vez mais fácil com os avanços em sistemas de coleta, transferência e armazenamento de dados. Visto que o sucesso de redes neurais artificiais está diretamente ligado a quantidade e qualidade dos dados disponíveis, é fácil notar como bases de dados cada vez maiores e mais confiáveis tem ajudado nos avanços da disciplina;
- **Algoritmos:** Melhores sistemas de otimização; funções de ativação; arquiteturas de redes e métodos do gradiente mais avançados trouxeram resultados extremamente promissores na última década. Tais inovações tiraram o campo do chamado “inverno IA” para uma era dourada de manchetes e investimentos bilionários, enquanto as gigantes do ramo de tecnologia procuram novos avanços;
- **Aprendizado Profundo:** Até o início da última década, as redes neurais existentes eram “rasas” devido ao seu baixo número de camadas e não possuíam poder representacional suficiente para obter resultados relevantes sem um pré-processamento manual de entradas. Nesse contexto, engenheiros e cientistas projetavam representações características (*features*) que acreditavam concentrar os dados mais relevantes para “facilitar” o trabalho desenvolvido pela rede em si, já que a mesma só poderia transformar os dados um número limitadamente

pequeno de vezes antes da saída. Hoje, com os avanços na área de aprendizado profundo, esse trabalho já não é mais necessário uma vez que há poder computacional suficiente para criar redes neurais artificiais com uma quantidade arbitrária de camadas.

Cada camada adicionada é acompanhada por um espaço predefinido de possibilidades, onde o mecanismo de aprendizado buscará a melhor representação dos dados através da retroalimentação da função de perda. Devido a isso pode-se dizer que a topologia de uma rede neural artificial determina o seu “espaço de hipóteses”, isto é, o espaço de todas as hipóteses que poderão ser utilizadas pela rede para mapear entradas e saídas (SAMMUT; WEBB, 2011)(Figura 11).

Figura 11: Exemplo de melhoria nas representações de uma rede neural.



Fonte: Autor.

Ao conceito de criar redes com cada vez mais camadas, capazes de aprender representações sucessivamente mais significativas dos dados processados, se deu o nome Aprendizado Profundo (*Deep Learning*). Essas redes, que podem atingir centenas de camadas, são capazes de mapear qualquer entrada à qualquer saída desde que haja uma correlação real por trás dos dados, executar tarefas de classificação e até mesmo gerar textos e imagens a partir do que já aprenderam. Atualmente, redes neurais profundas tem assistido à humanidade em campos como a ciência, medicina, manufatura, logística, desenvolvimento de *softwares* e até mesmo arte (CHOLLET, 2021).

Na subsecção 6.5.1 serão utilizados alguns elementos de redes neurais profundas, então é importante definir as equações utilizadas, como vistas em Chollet (2021), que podem ser encontradas a seguir:

- Passo à frente:

$$z_i^k(W, b) = \sum_{j=1}^{n^{(k-1)}} (w_{i,j}^k \cdot a_j^{(k-1)}) + b_i^k, \quad (5.3)$$

$$a_i^k(z) = g(z_i^k). \quad (5.4)$$

Onde:

$z_i^k(W, b)$  = função linear de um nó  $i$  na camada  $k$ ;

$n^{(k-1)}$  = quantidade de nós na camada anterior;

$w_{i,j}^k$  = peso da conexão entre o nó  $j$  da camada  $(k - 1)$  e o nó  $i$  da camada  $k$ ;

$a_j^{(k-1)}$  = valor de um nó  $j$  na camada  $(k - 1)$ ;

$b_i^k$  = viés do nó  $i$  da camada  $k$ ;

$a_i^k$  = valor de um nó  $i$  na camada  $k$ ;

$g(x)$  = função de ativação.

- Função de Ativação (Relu):

$$g(x) = \max(0, x); x \in \mathbb{R}. \quad (5.5)$$

- Função de Perda (MSE - Mean Squared Error):

$$E(\vec{y}, \hat{\vec{y}}) = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2. \quad (5.6)$$

Onde:

$\hat{\vec{y}}$  = vetor de valores inferidos ( $\hat{y} \in \mathbb{R}$ ).

$\vec{y}$  = vetor de valores observados ( $y \in \mathbb{R}$ ).

- Passo atrás (retropropagação):

$$w_{i,j}^{k+1} = w_{i,j}^k - \alpha \frac{\partial E}{\partial w_{i,j}^k}, \quad (5.7)$$

$$b_i^{k+1} = b_i^k - \alpha \frac{\partial E}{\partial b_i^k}. \quad (5.8)$$

Onde:

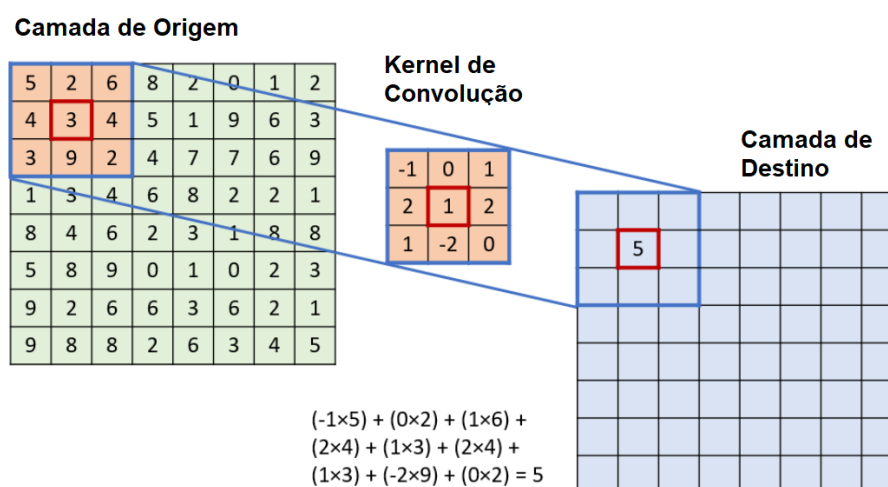
$\alpha$  = taxa de aprendizado ( $\alpha \in \mathbb{R}$ ).

## 5.4 Redes neurais convolucionais

Redes neurais convolucionais são arquiteturas de redes muito aplicadas a imagens devido aos resultados promissores que apresentam nos problemas de classificação e regressão (HAN *et al.*, 2020). Esses modelos normalmente são baseados em 4 tipos de camadas:

- **Camadas Convolutivas:** executam a operação matemática de convolução na entrada recebida, aplicando filtros (*kernels*) e procurando padrões em pequenas janelas de  $m \times n$  pixels, onde  $m$  e  $n$  são parâmetros configuráveis da camada. É possível entender o processo de forma mais didática através da Figura 12: partindo da origem da imagem, separa-se uma amostra de  $m \times n$  pixels e aplica-se o filtro de mesmo tamanho através da soma dos produtos entre os elementos, efetivamente atribuindo à intensidade do pixel central a intensidade média de seus  $m \times n$  vizinhos (PODAREANU *et al.*, 2019).

Figura 12: Exemplo da operação de convolução.

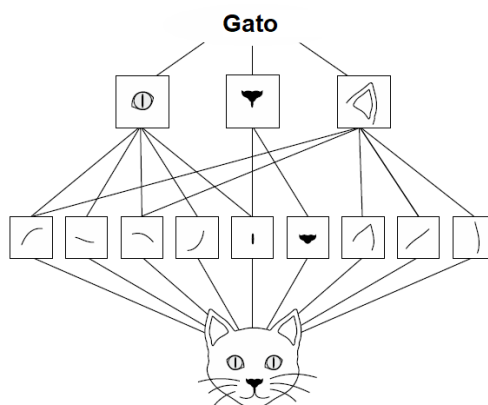


Fonte: Adaptado de Podareanu *et al.* (2019) pelo Autor.

Como nota Chollet (2021), esse comportamento traz ao modelo duas características importantes: o reconhecimento de padrões independente de translação, um número ou letra em uma folha escaneada identificado no fim de uma página será identificado com a mesma facilidade no início de uma página, por exemplo, e uma hierarquia de padrões, onde um padrão pode ser subdividido em diversos outros padrões. Isso permite que redes convolucionais encontrem relações en-

tre padrões cada vez mais complexos e representações cada vez mais abstratas, como demonstra a Figura 13.

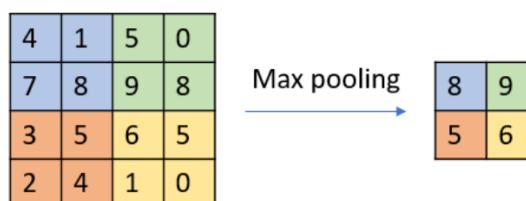
Figura 13: Exemplo de hierarquia de padrões.



Fonte: Adaptado de Chollet (2021) pelo Autor.

- Camadas de Agrupamento (*Pooling*):** consolidam *pixels* através de uma operação predefinida – normalmente obtenção de valores de intensidade máximos ou médios – da entrada recebida através de filtros (*kernels*) – de  $m \times n$  *pixels*, onde  $m$  e  $n$  são parâmetros configuráveis da camada. É possível entender o processo de forma mais didática através da Figura 14: partindo da origem da imagem, separa-se uma amostra de  $m \times n$  *pixels* e extrai-se o valor máximo/médio, efetivamente reduzindo a amostragem da imagem e mantendo apenas os *pixels* mais importantes. Essa operação é crucial em redes convolucionais pois, além de reduzir o número de parâmetros a serem treinados, formam padrões hierárquicos mais complexos – as janelas visualizam cada vez mais informações em relação ao tamanho original da imagem – e abstratos – *pixels* mais importantes influenciam mais no resultado final (PODAREANU *et al.*, 2019).

Figura 14: Exemplo da operação de *pooling*.



Fonte: Podareanu *et al.* (2019).

- **Camadas de Achatamento (*Flatten*):** transformam uma entrada tensorial de qualquer grau em um tensor de grau 1. Por exemplo, para uma entrada tensorial de grau 2 com forma (100, 3) a saída seria o tensor de grau 1 com forma (300). Essa camada desempenha a função de reduzir a dimensão dos dados processados em redes convolucionais para que possam ser conectados à uma camada densa (CHOLLET, 2021).
- **Camadas de Abandono (*Dropout*):** eliminam algumas ativações ou pesos da rede neural em cada iteração do processo de treinamento com base em uma probabilidade, visando a redução do sobreajuste. De forma resumida, a camada *Dropout* força a rede neural a se tornar mais robusta, pois quebra a dependência entre camadas (SRIVASTAVA *et al.*, 2014).
- **Camadas Densas (*Dense*):** geralmente são usadas na saída, como uma tradução entre a melhor representação obtida e a saída esperada. Essa camada é chamada assim pois conecta todos os nós da camada anterior à camada atual, utilizando uma função de ativação linear para problemas de regressão ou uma função de ativação limitada para problemas de classificação (CHOLLET, 2021).

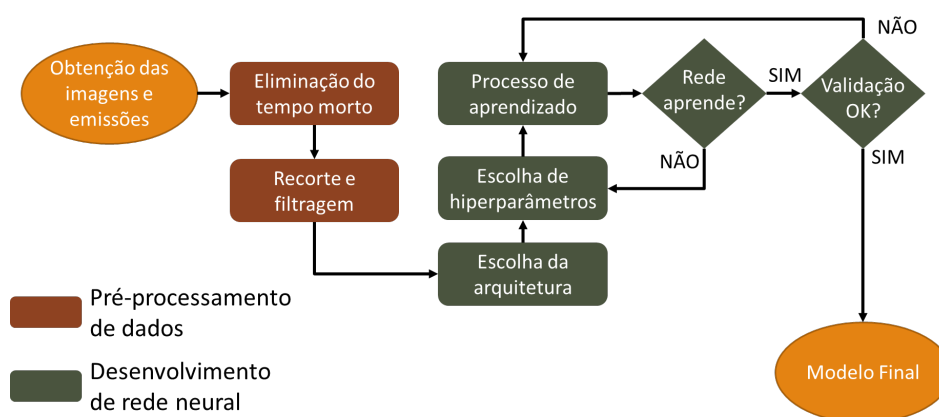
## 6 MÉTODO DE MONITORAMENTO DE CHAMAS ATRAVÉS DE IMAGENS E REDES NEURAS ARTIFICIAIS

Neste capítulo será explicado como as imagens de chamadas utilizadas foram coletadas, incluindo o aparato experimental; como as imagens foram processadas e o método de criação do sistema de inferência.

### 6.1 Método

A Figura 16 representa o método proposto dividido em etapas: i) Obtém-se as imagens e amostras de emissões em que será treinada a rede neural (Seção 6.2); ii) Elimina-se o tempo morto relacionado ao analisador de gases (Seção 6.3); iii) Equalizam-se as taxas de amostragem das imagens às emissões (Seção 6.4); e iv) Escolhem-se a arquitetura e parâmetros da rede neural para então treiná-la através de um processo iterativo (Seção 6.5). Ao fim desse processo, caso a rede neural seja validada, chega-se ao modelo final.

Figura 15: Diagrama do método proposto.



Fonte: Autor

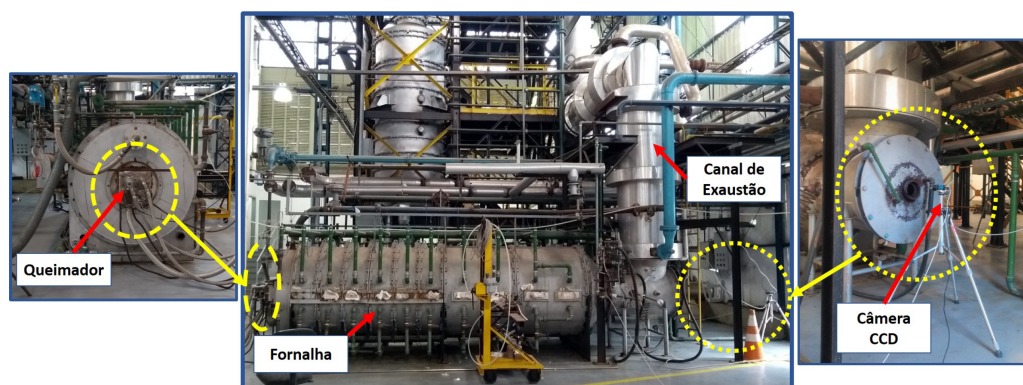


## 6.2 Coleta de dados

### 6.2.1 Aparato experimental

Os experimentos foram realizados em uma fornalha horizontal cilíndrica (Figura 15) situada no Laboratório de Engenharia Térmica do Instituto de Pesquisas Tecnológicas do Estado de São Paulo (IPT) por Silva Neto *et al.* (2021).

Figura 16: Fornalha horizontal utilizada nos experimentos.



Fonte: Silva Neto *et al.* (2021)

Em uma das extremidades da fornalha encontra-se um queimador que injeta gás natural com oxigênio puro para o processo de combustão e, na outra extremidade, há um duto para saída das emissões e a pequena abertura onde uma câmera foi posicionada a fim de capturar as imagens utilizadas por Silva Neto *et al.* (2021) e por este trabalho de conclusão de curso.

A fornalha está equipada com um painel de comando eletrônico, que controla as vazões de gás natural e oxigênio, ligado a um sistema supervisor que coleta essas informações; a temperatura em vários pontos ao longo da fornalha e a concentração de  $O_2$  e  $CO_2$  nos gases de emissão, capturados através de um analisador de gases *Horiba* modelo *PG-350E* a cada intervalo de tempo de 1 segundo.

As imagens foram capturadas através de uma câmera CCD monocromática *Imaging Source* modelo *DMK 23G618* com lente objetiva (12 mm, f1,2) à uma taxa de 30 quadros por segundo, no formato 640x480 e resolução 0,3 *MegaPixels*.

## 6.2.2 Ensaio

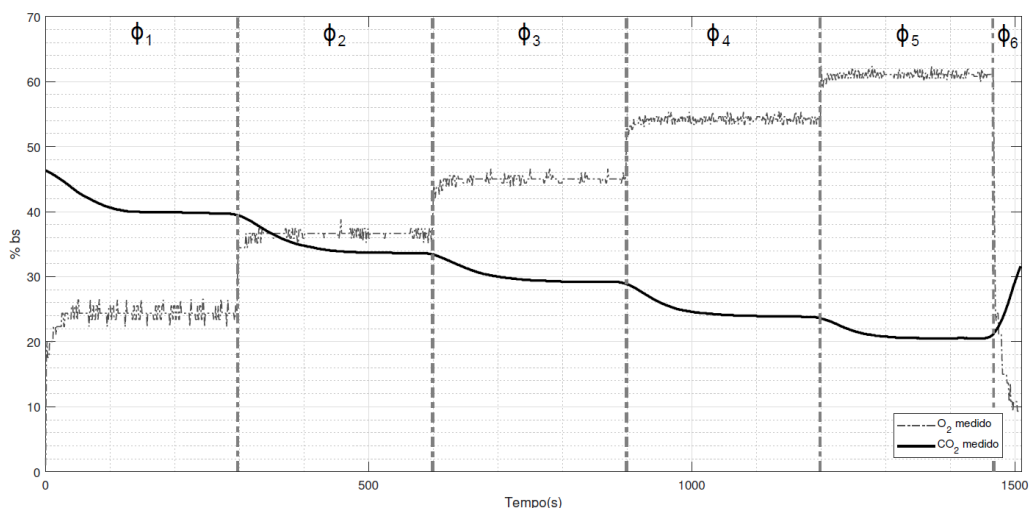
Os experimentos realizados por Silva Neto *et al.* (2021) foram conduzidos em 6 diferentes razões de equivalência da mistura ar-combustível (Tabela 1) ao longo de um período de aproximadamente 25 minutos, tempo este cedido pelo IPT. Os efeitos dessa variação foram captados pelo analisador de gases e os resultados obtidos foram plotados na Figura 17. Correlacionando a Tabela 1 à Figura 17 pode-se perceber que conforme a mistura se torna mais pobre, isto é, passa a ter mais ar e menos combustível, as emissões de O<sub>2</sub> aumentam e as de CO<sub>2</sub> diminuem até estabilizarem e então a razão é alterada novamente para observar como o sistema se comporta.

Tabela 1: Razões de equivalência da mistura ar-combustível utilizadas e duração dos ensaios.

Condição	$\Phi$	Duração (s)
1	0,85	296
2	0,70	302
3	0,60	300
4	0,50	301
5	0,42	267
6	1,05	44

Fonte: Silva Neto *et al.* (2021)

Figura 17: Emissões de O<sub>2</sub> e CO<sub>2</sub> obtidas durante o ensaio sob diferentes razões de equivalência.



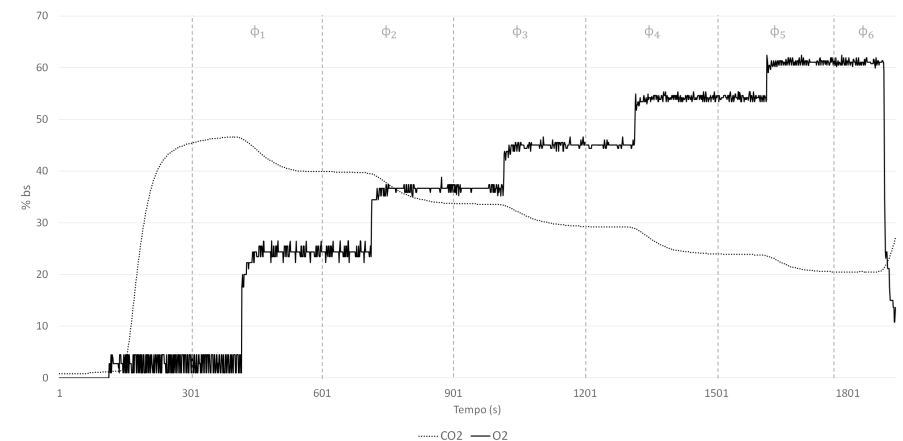
Fonte: Silva Neto *et al.* (2021)

A câmera foi posicionada externamente em uma abertura existente no final da câmara de combustão, obtendo assim imagens de topo da chama. O vídeo gravado na taxa de 30 quadros por segundo foi extraído após o experimento e convertido em imagens, porém com uma placa de captura de imagens (*frame grabber*) seria possível processar essas imagens em tempo real.

### 6.3 Processamento dos dados

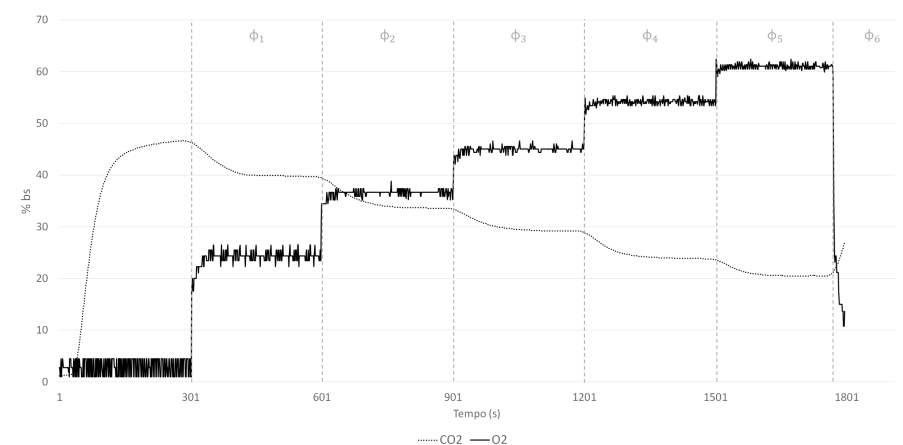
A fim de sincronizar a imagem da chama à emissão gerada naquele momento, deve-se considerar o “tempo morto” – o tempo entre a ocorrência de um fenômeno e o momento em que se observa o mesmo – entre a queima do gás natural e o instante que as emissões geradas chegam ao analisador de gases. Para tal, analisou-se o atraso médio observado entre variações da razão  $O_2$ /gás natural na entrada e variações de  $O_2$  e  $CO_2$  na saída. A análise indica que, na fornalha do experimento, o produto da combustão leva 114 segundos para alcançar o analisador de gases no duto de exaustão do forno. Este atraso, que inevitavelmente afeta os sensores tradicionais posicionados no duto de exaustão, é o que faz com que técnicas de controle compensem excessivamente o sistema (LIU; FAN; CHEN, 2017), resultando em um controle pouco robusto e lento como mencionado no início do Capítulo 2. Uma das grandes vantagens do monitoramento por imagens é a medição em tempo real dos parâmetros desejados. A partir desse contexto, os dados coletados referentes ao teor de  $O_2$  e  $CO_2$  nas emissões foram atrasados em 114 segundos para ficarem sincronizados com as imagens de chama e variações nas razões de equivalência, mesmo procedimento utilizado em Silva Neto *et al.* (2021). O comportamento das emissões antes e após o atraso podem ser encontrados respectivamente nas Figuras 18 e 19. Percebe-se pelas imagens citadas que antes do atraso, ao alterar-se a razão de equivalência (linhas verticais segmentando o gráfico) as emissões continuavam estáveis por um período. Esse comportamento é explicado pelo tempo morto do analisador de gases, que só consegue observar a variação das emissões quando os gases gerados na combustão chegam ao duto de exaustão.

Figura 18: Emissões de O<sub>2</sub> e CO<sub>2</sub> ao longo do experimento, com tempo morto.



Fonte: Autor

Figura 19: Emissões de O<sub>2</sub> e CO<sub>2</sub> ao longo do experimento, atrasados em 114s.



Fonte: Autor

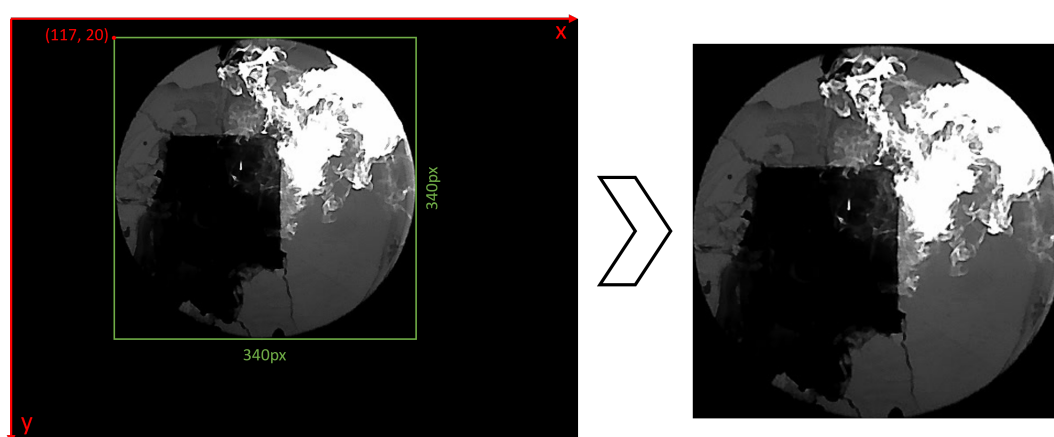
## 6.4 Processamento das imagens

As imagens obtidas, coletadas na taxa de 30 quadros por segundo devido ao modelo da câmera, não estão na mesma frequência de amostragem dos dados de emissões, coletados na taxa de 1 amostra por segundo, e também contêm informações sem valor agregado como o entorno causado pela vigia da fornalha que pode ser visualizada na Figura 20. Desta forma, foram realizadas 2 manipulações de pré-processamento que podem ser realizadas em tempo real no caso de uma implementação *online*, o recorte e o filtro de média móvel.

### 6.4.1 Recorte

As imagens originais foram recortadas utilizando uma máscara quadrada de 340px na posição (117,20), como ilustra a Figura 20. A transformação das imagens de 640x480px em imagens 340x340px reduziu o espaço necessário para armazená-las em 62,6%, além de reduzir o custo computacional nas etapas seguintes sem perder nenhum dado relevante. Vale salientar que as coordenadas utilizadas e o tamanho da imagem final dependem do tipo do forno, da câmera e da posição relativa entre os dois, devendo ser ajustado de acordo.

Figura 20: Ilustração do processo de recorte das imagens.



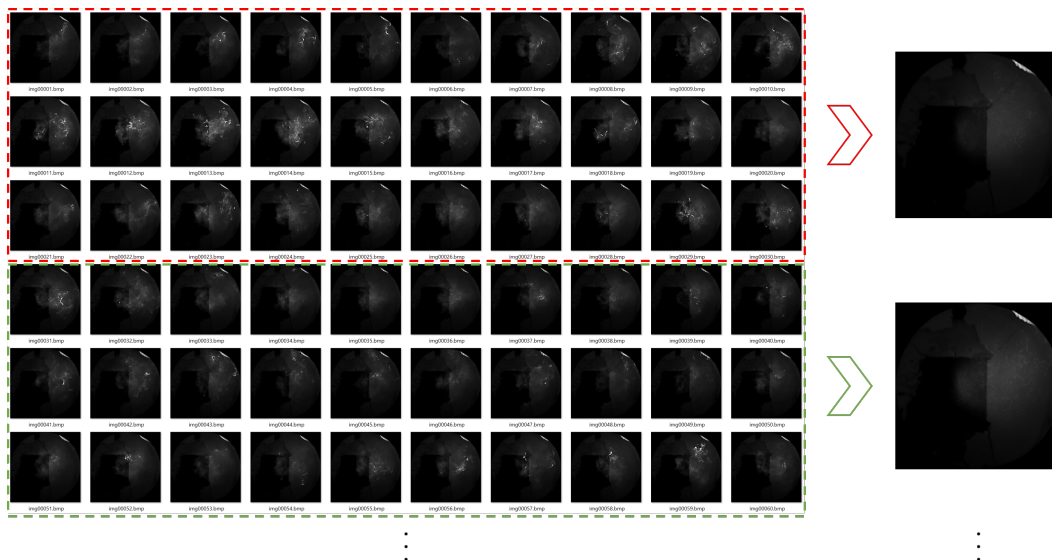
Fonte: Autor

### 6.4.2 Filtro de média móvel

Para sincronizar as imagens obtidas, coletadas a 30 quadros por segundo, com os dados de emissões, coletados a 1 amostra por segundo, optou-se por aplicar um filtro de média móvel nas imagens a fim de equalizar as taxas de amostragem conforme visto em Silva Neto (2021). A aplicação do filtro apresenta também outros benefícios, como eliminar ruídos nas imagens e reduzir em 98,5% o espaço necessário para armazenar as imagens (antes 38.5GB e 168MB após), além do tempo para processá-las nos próximos passos em um fator de 30 vezes. Ao gerar uma imagem média dos 30 quadros capturados, o algoritmo está efetivamente realçando características mais importantes da combustão – *pixels* que ficaram luminosos por mais tempo apresentam alta luminosidade na imagem final – e desvanecendo ruídos e interferências – *pixels* que ficaram luminosos por menos tempo apresentam baixa luminosidade na imagem final. Um exemplo da aplicação do filtro de média móvel pode ser visto na Figura 21, onde 60

imagens se tornam 2.

Figura 21: Ilustração do filtro de média móvel com 30 amostras.



Fonte: Autor

## 6.5 Método de inferência

Para o desenvolvimento do método de inferência capaz de relacionar imagens de chama às emissões de CO<sub>2</sub> em tempo real, foram escolhidos os *frameworks* de aprendizado de máquina mais utilizados no momento da publicação: *TensorFlow* e *Keras*, utilizando a linguagem de programação *Python 3.7.10*. Para atingir resultados eficientes, diversas arquiteturas de redes neurais artificiais tradicionais e convolutivas foram implementadas e testadas, das quais a arquitetura proposta na Subseção 6.5.1 apresentou melhor equilíbrio entre precisão, complexidade e custo computacional. A arquitetura proposta pode ser treinada e avaliada de forma relativamente rápida – 4 minutos no computador pessoal do Autor<sup>1</sup> – de forma que viabilize a sua execução em hardware embarcado e microcontroladores, isto é, computadores dedicados que são anexados ao sistema que controlam, e ainda mantenha uma baixa margem de erro nas condições testadas.

Após a etapa de pré-processamento, o conjunto de 1492 imagens obtidas foi embaralhado aleatoriamente através de uma distribuição uniforme, de forma que a rede

<sup>1</sup>Windows 10 com processador Intel Core i5 9600K, duas memórias RAM DDR4 8GB e sem placa de vídeo.

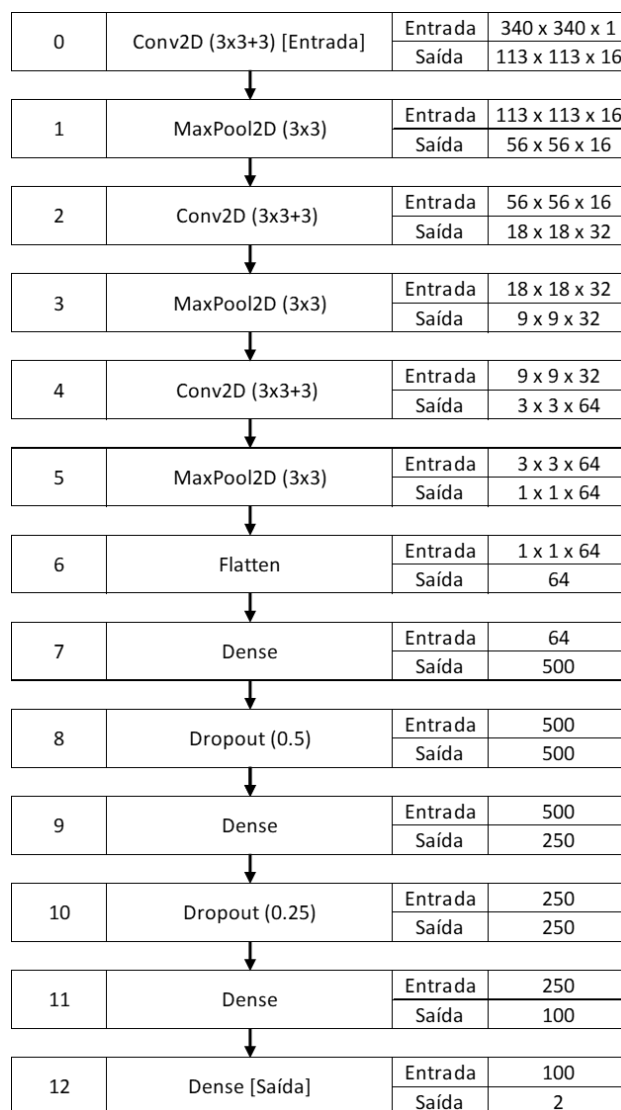
neural pudesse observar diferentes razões de equivalência a cada passada, e dividido em 3 categorias conforme é comum na literatura:

- **Subconjunto de aprendizado:** Equivalente a 55% do total de imagens, esse subconjunto contou com 821 imagens embaralhadas aleatoriamente e serviu para que a rede pudesse “aprender” a relacionar imagens de chama à medição correspondente de CO<sub>2</sub> encontrada durante os experimentos;
- **Subconjunto de validação:** Equivalente a 15% do total de imagens, esse subconjunto contou com 223 imagens embaralhadas aleatoriamente e serviu para que a rede pudesse validar as relações e conclusões encontradas durante o processo de aprendizado. É importante ressaltar que a rede não aprende com esses dados, os mesmos são usados exclusivamente para validar se a mesma está progredindo no caminho correto ou se afastando dele. Ainda assim, o processo de “tentar novamente” quando a validação é ruim cria um viés equivalente ao “vazamento” indireto de informações do conjunto de validação. Na prática, isso pode fazer com que a rede aprenda como se desempenhar bem no conjunto de validação, invalidando o processo. Para resolver tal dilema, cria-se um terceiro subconjunto de testes onde a rede será avaliada ao final do processo de aprendizagem em dados nunca antes vistos;
- **Subconjunto de testes:** Equivalente a 30% do total de imagens, esse subconjunto contou com 448 imagens e serviu para que a rede pudesse ser avaliada de forma definitiva em dados nunca antes processados. Após esta etapa, a rede não deve mais ser alterada ou iterada, mas sim descartada caso não alcance os indicadores desejados, seja quais forem, para evitar o mesmo vazamento apresentado no subconjunto de validação.

## 6.5.1 Arquitetura proposta

A arquitetura proposta é uma versão simplificada e adaptada do modelo de rede neural convolucional proposto por Bojarski *et al.* (2016). Na versão do Autor, o modelo consiste em uma rede com 12 camadas e 206.246 parâmetros treináveis, dispostos de acordo com a Figura 22. No modelo, algumas escolhas arbitrárias foram tomadas baseadas em estudo da literatura (CHOLLET, 2021), como as camadas convolutivas que utilizam filtros 3x3 com passo 3 ou as probabilidades das camadas de *Dropout*. Visando a completude do estudo, foram testados também outros tamanhos comuns de filtros e passos, cujos resultados serão discutidos posteriormente na Tabela 2 do Capítulo 7.

Figura 22: Diagrama do macroprocesso da arquitetura proposta.



Fonte: Autor



## 6.5.2 Visualizações de camadas intermediárias

O processo de desenvolvimento de uma rede neural é iterativo e cheio de experimentação, mas visualizações podem acelerá-lo e ajudar a solucionar problemas que surgem durante os experimentos (JHA, 2022). Recentemente, o interesse em entender como redes neurais interpretam dados e aprendem vem crescendo e, para tal, novas formas de observar o modelo caixa-preta<sup>1</sup> da rede neural são desenvolvidos. Como cita Krause, Perer e Bertini (2016), as principais razões para se analisar redes neurais de forma gráfica são:

**Análise de dados** O aprendizado de máquina é tipicamente utilizado como ferramenta para fazer previsões, mas também pode ser utilizado como uma ferramenta para compreender e observar dados complexos.

**Modelos de confiança** Quando os modelos tomam decisões importantes, é necessário ter uma melhor compreensão do que uma rede faz e por quê. Assim, o aumento da transparência desempenha um papel fundamental no seu nível de confiabilidade.

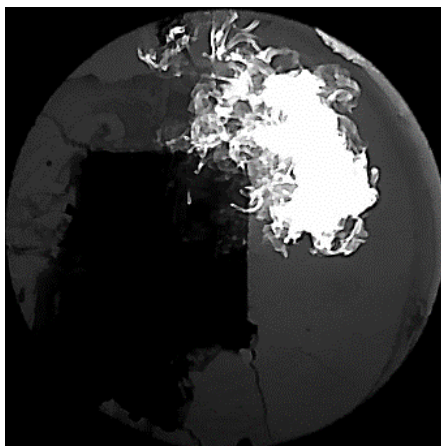
**Comparação de modelos e diagnósticos** Os desenvolvedores precisam muitas vezes diagnosticar onde um modelo falha para gerar versões mais eficientes.

Nesse contexto, escolheu-se extrair e plotar a saída das camadas intermediárias da rede sequencialmente quando uma imagem escolhida de forma aleatória passa pelo modelo. Esse procedimento torna o modelo caixa-preta um pouco mais transparente, permitindo a análise do processo de aprendizado e servindo também como ferramenta de diagnóstico durante os experimentos de otimização de parâmetros. A título de exemplo, serão apresentadas nas Figuras 24, 25 e 26 uma análise da arquitetura proposta na Subseção 6.5.1 quando a Figura 23 é utilizada como entrada.

---

<sup>1</sup>Um modelo caixa-preta é identificado através da análise dos dados de entrada e saída, ao invés de observar a física do sistema (AGUIRRE, 2004).

Figura 23: Entrada escolhida para gerar exemplos de representações intermediárias da rede, imagem de chama no instante  $t = 419, 8s$ .



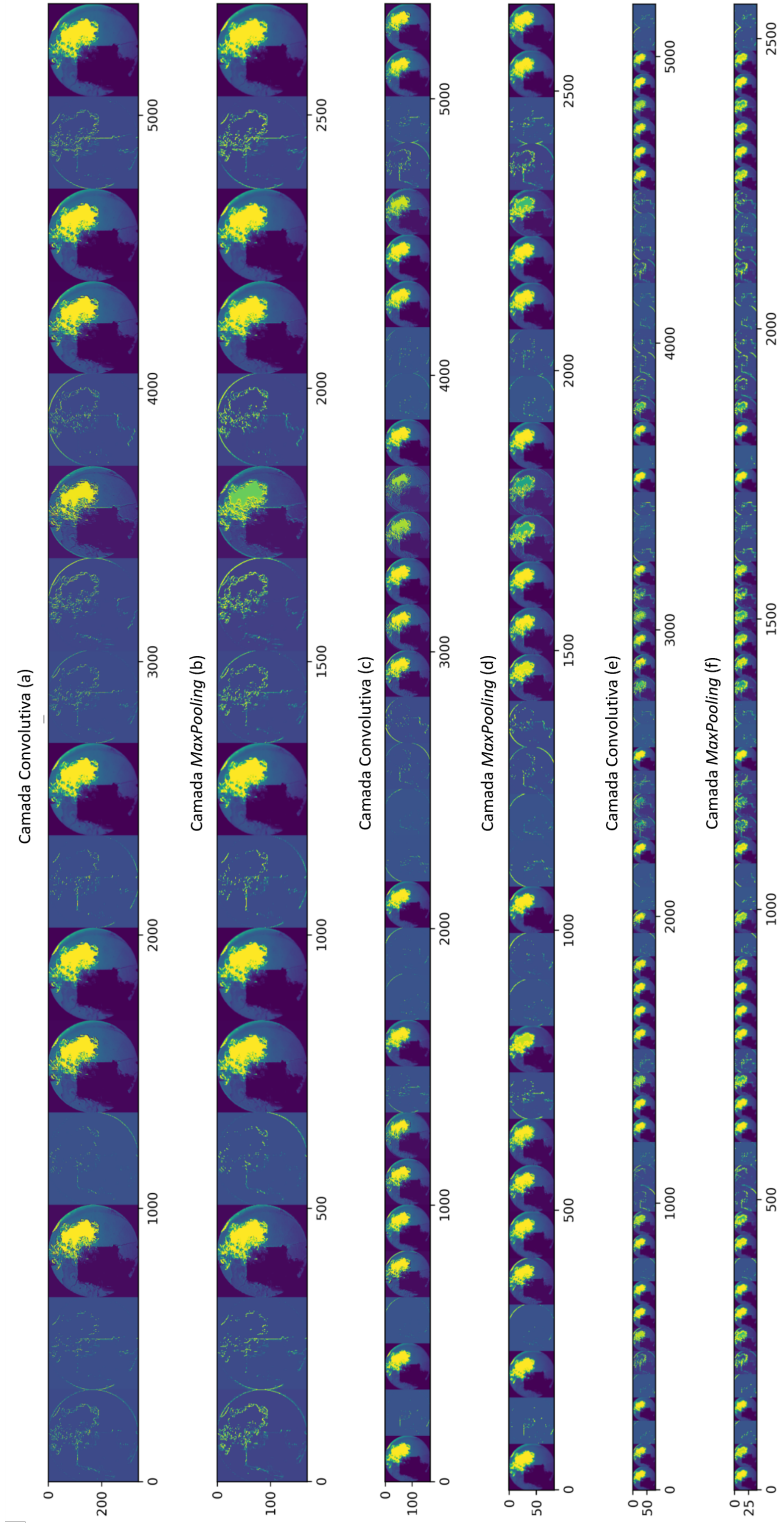
Fonte: Autor

Na Figura 24, que consolida as saídas de cada camada convolutiva e de agrupamento da rede antes do processo de aprendizado, pode-se observar:

- Como os filtros são inicializados propositalmente de forma randômica para evitar tendências no processo de aprendizado, as saídas nas camadas convolutivas (a), (c) e (e) são gerados por padrões aleatórios e serão alterados durante o processo de treinamento através do método do gradiente;
- As camadas de agrupamento por valor máximo (*MaxPooling*) (b), (d) e (f) reduzem em uma escala de 2:1 as imagens obtidas enquanto mantém boa fidelidade à imagem original da chama, como mencionado na Seção 5.4.

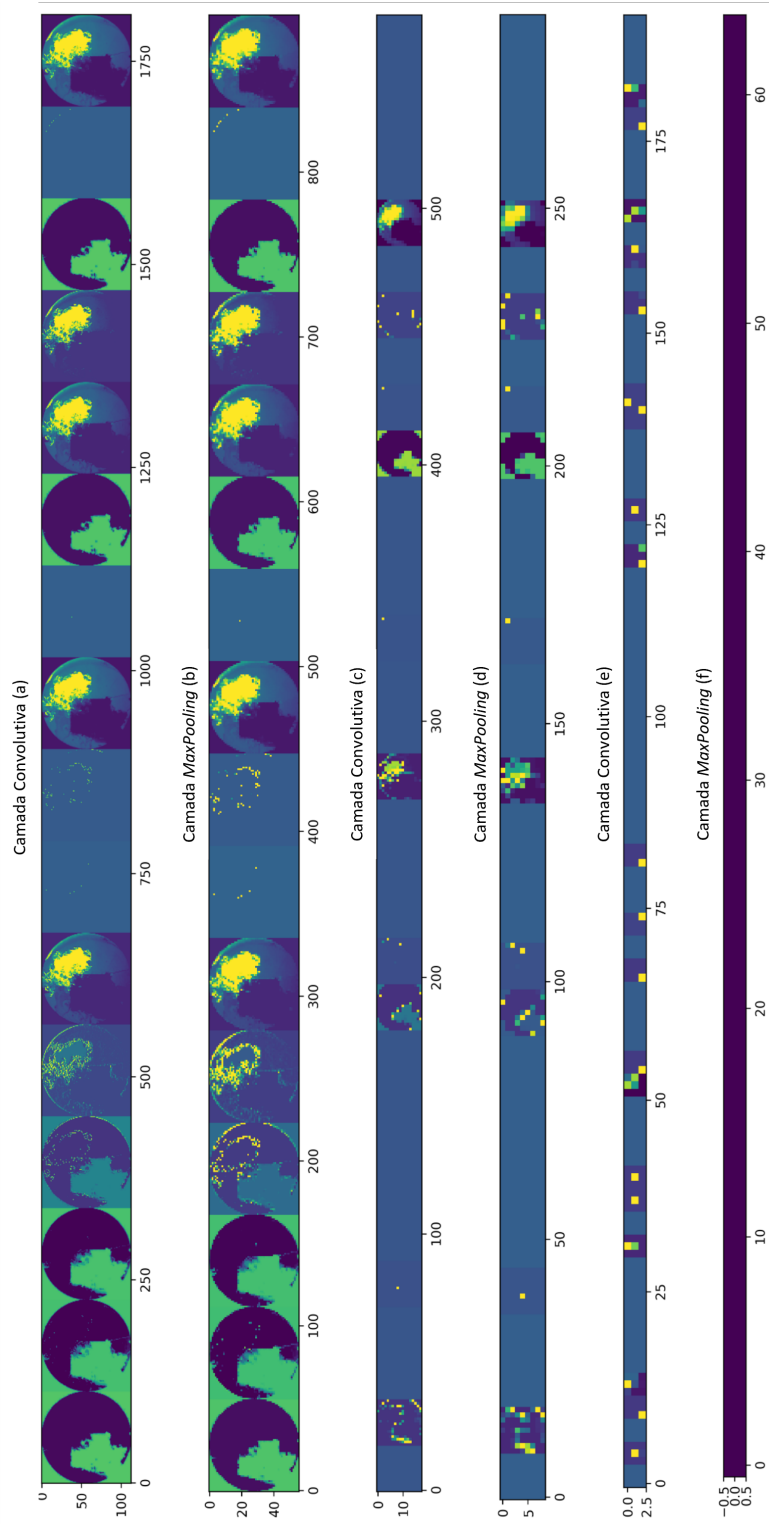
As Figuras 25 e 26 representam o mesmo processo, porém após o treinamento da rede. Nota-se como os filtros mudam quando comparados aos da Figura 24 e que as imagens se tornam cada vez menores e mais abstratas em camadas mais profundas, onde são retidas apenas as características mais importantes das imagens para o processo de classificação ou regressão. Durante o início do processo de desenvolvimento, essas visualizações foram utilizadas para diagnosticar o baixo desempenho da rede e a otimização do número de representações extraídas em cada camada.

Figura 24: Representações intermediárias das camadas convolutivas pré-treinamento, coloridas artificialmente para facilitar visualização.



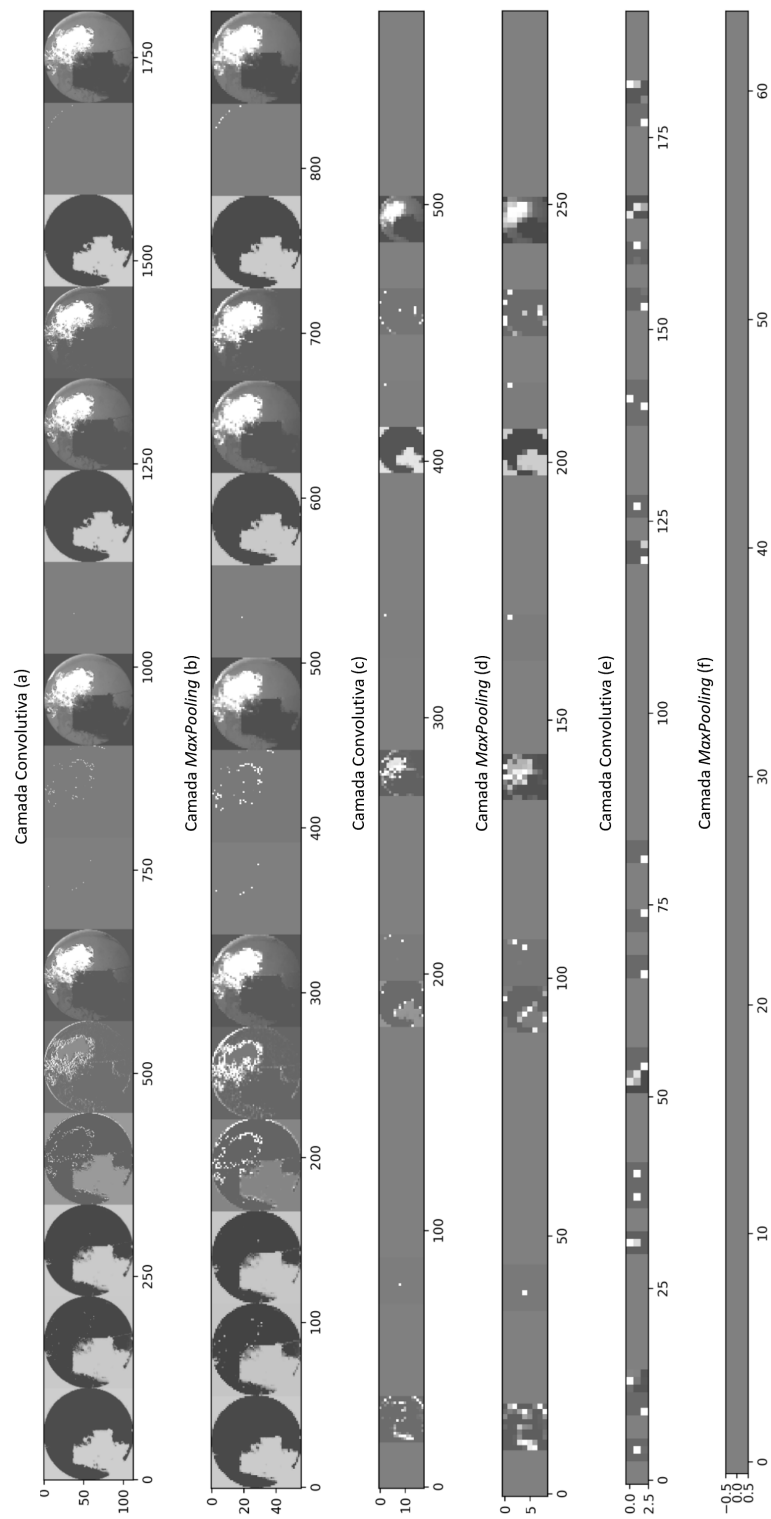
Fonte: Autor

Figura 25: Representações intermediárias das camadas convolutivas pós-treinamento, coloridas artificialmente para facilitar visualização.



Fonte: Autor

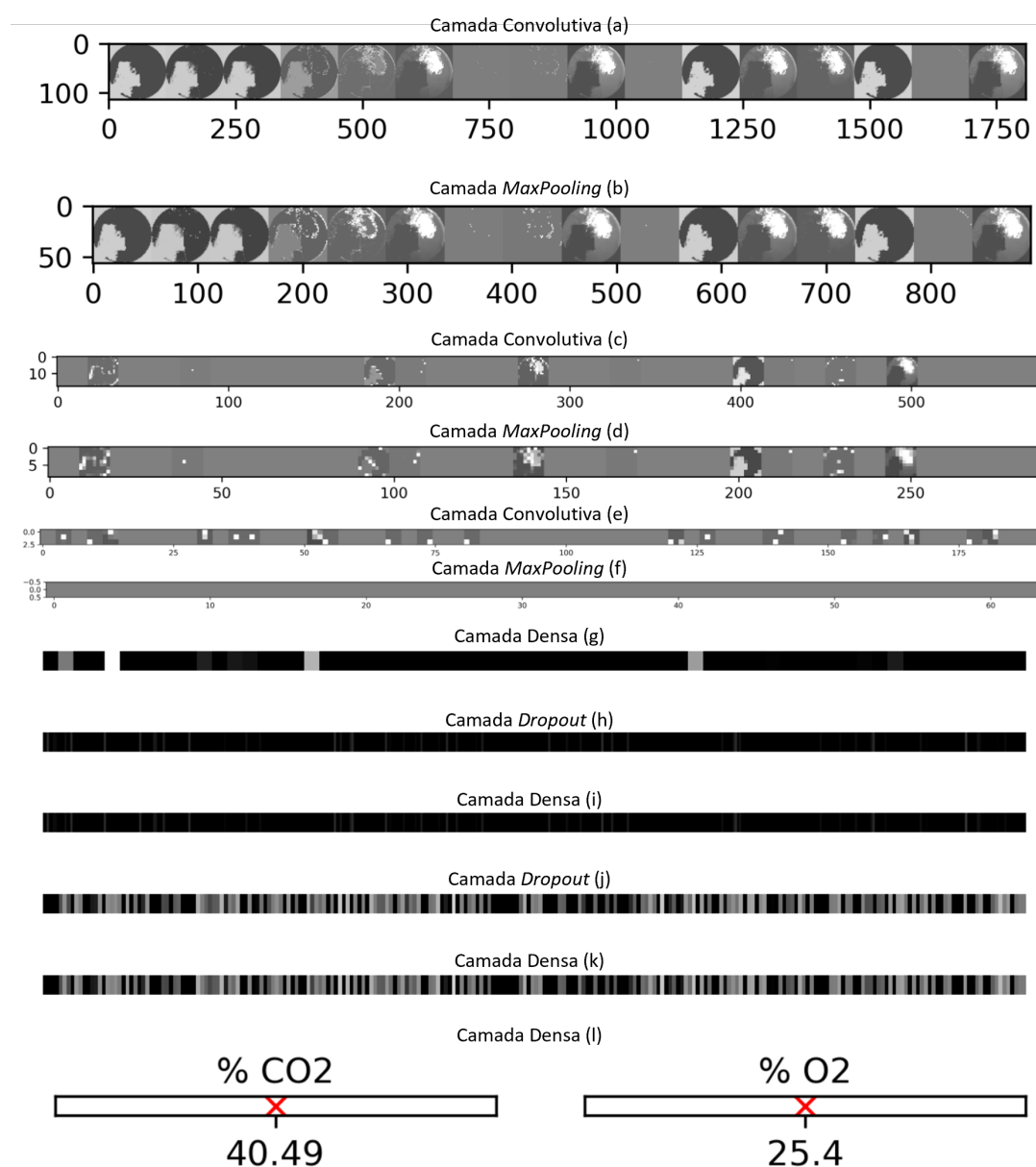
Figura 26: Representações intermediárias das camadas convolutivas pós-treinamento, não manipuladas.



Fonte: Autor

A Figura 27 reúne as saídas de todas as camadas da arquitetura proposta, possibilitando algumas observações. Nas camadas (a), (b), (c), (d), (e), (f) a rede neural está procurando similaridades entre a Figura 23 e as imagens em que foi treinada, atribuindo um peso maior aos *pixels* em comum (visualizado pela tonalidade, onde um peso maior indica uma cor mais clara). Conforme se observa camadas mais profundas, as características procuradas vão ficando menores, mais numerosas e abstratas afim de generalizar para novos dados. Já nas camadas (g), (h), (i), (j), (k) e (l) a rede neural utiliza o modelo caixa preta identificado durante o treinamento para mapear as características abstratas encontradas no passo anterior às saídas esperadas.

Figura 27: Representações intermediárias da rede completa.



Fonte: Autor

## 7 RESULTADOS

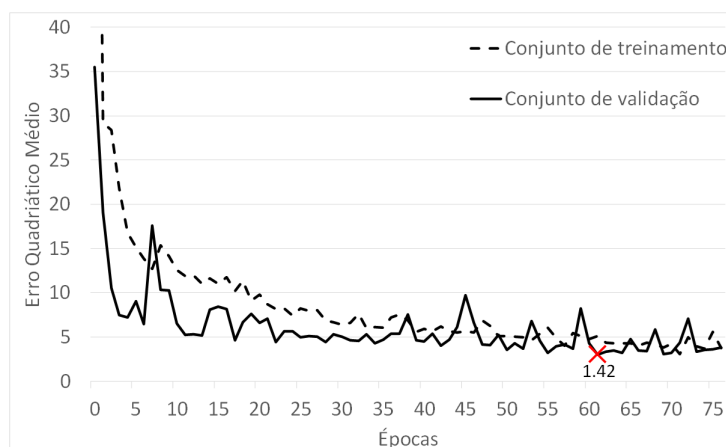
O treinamento ótimo da rede foi atingido após 62 épocas com as seguintes configurações:

- **Método de otimização:** Método do gradiente descendente estocástico (SGD) com uma taxa de aprendizado (*learning rate*) de 0,001. Um dos métodos mais comuns no desenvolvimento de redes neurais devido ao seu menor custo computacional (CHOLLET, 2021);
- **Função de custo:** Erro quadrático médio (MSE). Entre as funções de custo mais utilizadas em problemas de regressão, isto é MSE/RMSE e MAE, o MSE foi escolhido por ser o único que leva em consideração o sinal do erro e por penalizar mais os erros elevados, o que pode elevar a performance do sistema (CHAI; DRAXLER, 2014);
- **Tamanho de Lote:** Os pesos da rede foram atualizados a cada 4 imagens, resultando em 224 passadas pela rede a cada época;
- **Regularização:** O sobreajuste (*overfit*) aos dados foi combatido através de 2 camadas de abandono (*Dropout*), com taxas de 0,5 e 0,25 respectivamente;
- **Inicialização:** Os pesos da rede foram inicializados de forma uniformemente randômica, como é comum no campo, para evitar a criação de tendências no modelo;
- **Normalização:** Os valores dos *pixels* na entrada do modelo foram normalizados na faixa de 0 a 1, pois, sabe-se empiricamente que o passo ajuda na propagação de gradientes acelerando o processo de aprendizado (CHOLLET, 2021) e (IOFFE; SZEGEDY, 2015).

## 7.1 Métricas

Após a época de número 62, a rede neural começou a apresentar evolução no erro quadrático médio do conjunto de aprendizado e involução no conjunto de validação, indicando sobreajuste (Figura 28, com MSE mínimo realçado em vermelho).

Figura 28: Evolução do erro quadrático médio ao longo do período de treinamento.



Fonte: Autor

Os parâmetros não treináveis das camadas convolutivas – tamanho do filtro (*kernel*) e espaçamento (*stride*) – foram escolhidos através de experimentos sistemáticos como sugere Brownlee (2018), partindo dos valores mais comuns na literatura de redes neurais convolucionais e alterar um parâmetro por vez até encontrar valores ótimos. Os valores testados e o resultado dos experimentos no conjunto de dados trabalhado se encontra compilado na Tabela 2.

Tabela 2: Métricas de desempenho da arquitetura proposta sob diferentes hiperparâmetros.

Filtro	Passo	RMSE	Tempo de Treinamento
2x2	1	3,93	0:20:34
2x2	2	1,41	0:04:17
2x2	3	2,27	0:02:35
3x3	1	3,96	0:22:09
3x3	2	1,70	0:06:04
<b>3x3</b>	<b>3</b>	<b>1,19</b>	<b>0:02:19</b>
4x4	1	4,44	0:20:55
4x4	2	1,41	0:05:25
4x4	3	1,55	0:02:59

Fonte: Autor.



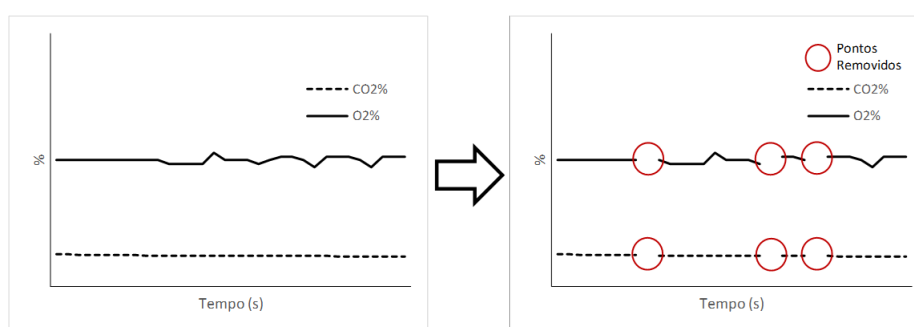
A iteração de melhor ajuste aos dados foi a com filtro 3x3 e passo 3, que apresentou raiz do erro quadrático médio (RMSE) absoluto igual a 1,19 e erro máximo absoluto igual a 6,21 no conjunto de testes – imagens nunca antes processadas pelo modelo – o que indica um bom equilíbrio entre precisão e generalização. <sup>1</sup>

## 7.2 Validações

### 7.2.1 Cenário I

No primeiro cenário, imagens de chamas foram removidas de forma aleatória do conjunto de aprendizado pré-treinamento para compor o conjunto de testes, como exemplifica a Figura 29.

Figura 29: Conjunto de aprendizado antes (esquerda) e depois (direita) da remoção de dados para compor o conjunto de testes.

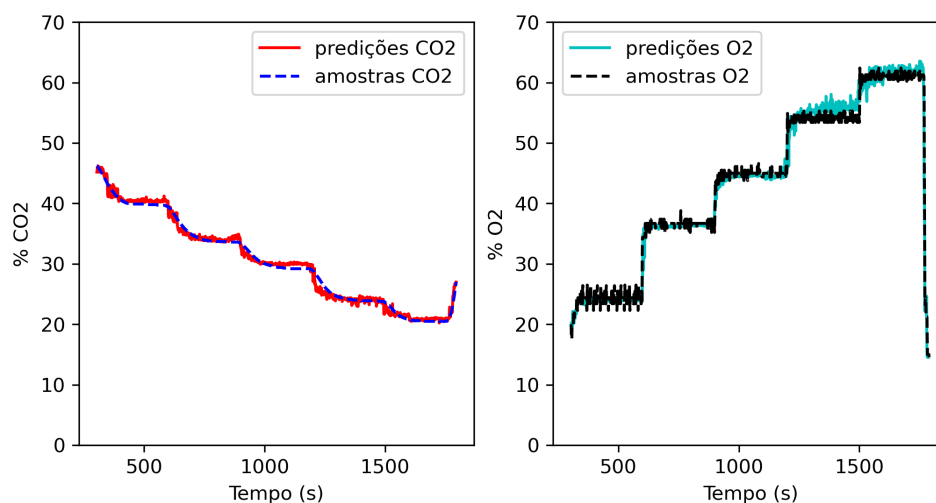


Fonte: Autor

Desta forma, o mecanismo de inferência foi exposto à todas as razões de equivalência e aprendeu características das imagens de chamas para cada uma delas. Isso levou o mecanismo a desempenhar bem dentro da faixa observada, como pode ser visto na Figura 30 através da semelhança entre as linhas de predições e amostras coletadas.

<sup>1</sup>Durante reproduções os números podem variar devido à natureza estocástica do processo de aprendizado profundo, advinda da inicialização randômica dos tensores de peso e do embaralhamento das imagens pré-divisão em conjuntos de aprendizado, validação e teste.

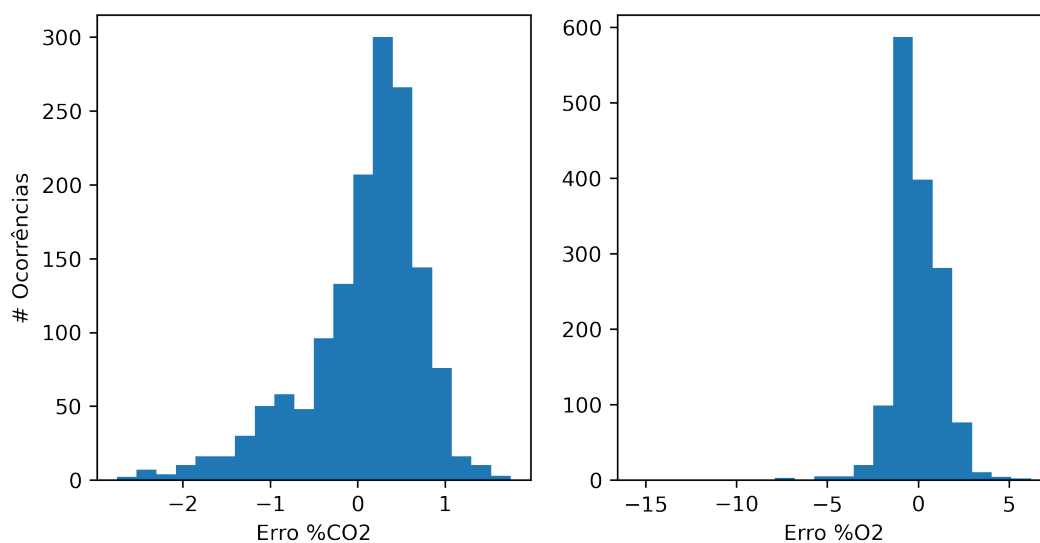
Figura 30: Teor de O<sub>2</sub> e CO<sub>2</sub> amostrado vs estimado ao longo do experimento.



Fonte: Autor

Devido à não continuidade dos dados embaralhados, não é possível plotar as séries temporais de amostras vs predições. Desta forma, optou-se por analisar a distribuição da raiz do erro quadrático médio. O desvio padrão igual à 1,046 e a média igual à 0,048 indicam uma distribuição bem centralizada e com pouca variância, validando o modelo (Figura 31).

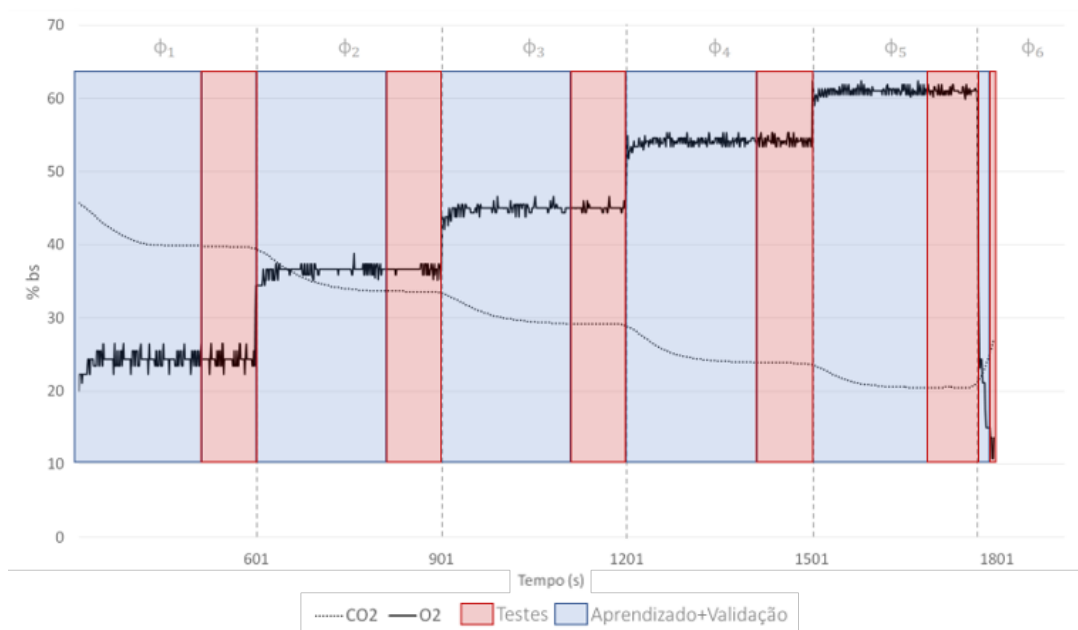
Figura 31: Histograma da raiz do erro quadrático médio amostrado vs estimado no conjunto de testes.



Fonte: Autor

Devido à todas as amostras terem a mesma chance de ser selecionadas para compôr o conjunto de testes, levantou-se a hipótese de que uma razão de equivalência poderia estar mais presente que as outras, gerando uma tendência nas métricas. Para garantir que os dados do conjunto de testes estavam sendo uniformemente distribuídos, o procedimento foi repetido da seguinte forma: i) As razões de equivalência foram separadas e tiveram seus dados embaralhados; ii) Cada razão de equivalência foi fracionada na proporção 55:15:30 em subconjuntos de aprendizado, validação e testes, respectivamente; iii) Os subconjuntos foram agrupados em suas respectivas categorias e embaralhados novamente; e iv) A rede neural proposta foi retreinada e validada com os mesmos hiperparâmetros nos novos conjuntos. O método descrito é ilustrado na Figura 32.

Figura 32: Divisão do conjunto de dados entre aprendizado e testes após nova segmentação.



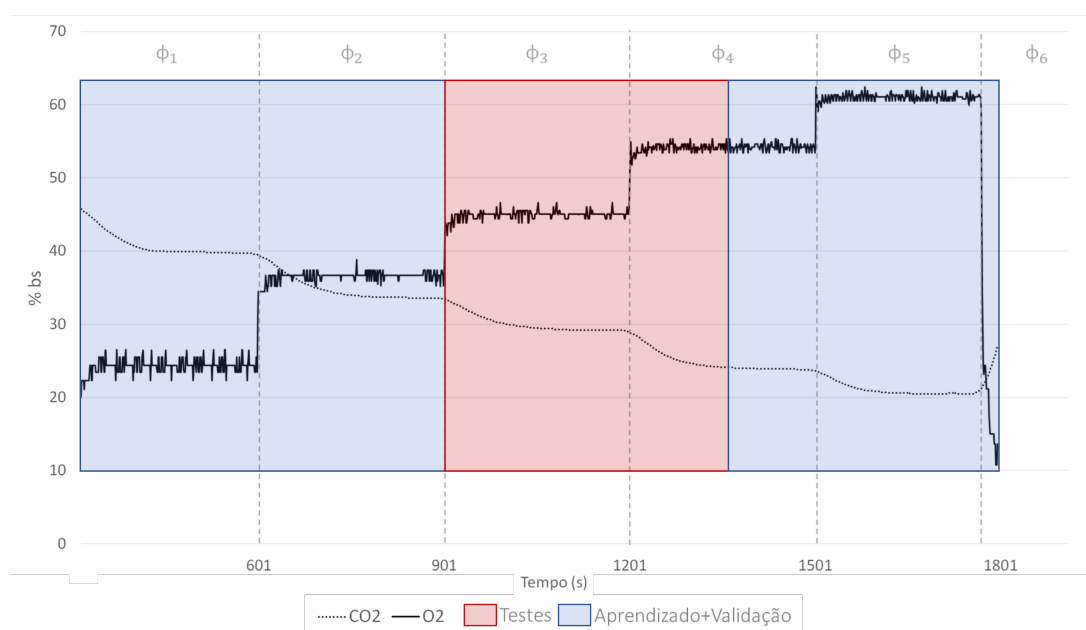
Fonte: Autor

Os resultados obtidos foram erro quadrático médio igual à 1,41 (vs 1,19 anteriormente), desvio padrão igual à 1,416 (vs 1,046 anteriormente) e a média igual à 0,079 (vs 0,048 anteriormente). A similaridade das métricas encontradas em ambos os casos indica que não houve viés relevante na primeira forma de escolha do conjunto de testes.

## 7.2.2 Cenário II

No segundo cenário, o mecanismo de inferência foi treinado até o instante  $t = 901s$  e após o instante  $t = 1353s$ , como ilustra a Figura 33. O intuito deste experimento é medir a capacidade da rede proposta de estimar resultados para novas razões de equivalência dentro da faixa treinada, isto é, treinando a rede neural em proporções ar-combustível altas e baixas deseja-se verificar o desempenho em proporções ar-combustível intermediárias.

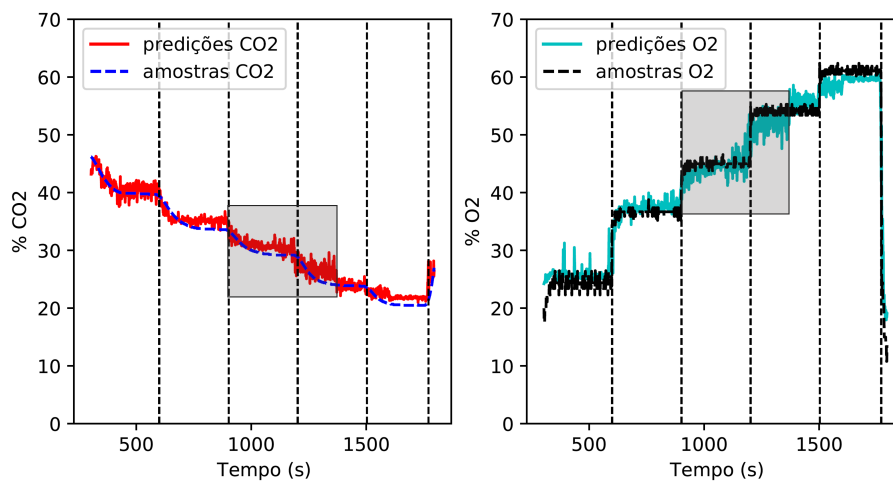
Figura 33: Divisão do conjunto de dados entre aprendizado e testes para o segundo cenário.



Fonte: Autor

O comportamento da rede neural pode ser visualizados na Figura 34, onde a área sombreada corresponde ao conjunto de testes. Percebe-se que as estimativas foram razoavelmente precisas, mesmo em uma razão de equivalência completamente nova para a rede, através do erro quadrático médio igual à 2,01 e da similaridade entre as linhas de previsões e a de amostras. Esse resultado é promissor pois indica que o método proposto não precisa ser treinado em todas as razões de equivalência do processo em que será utilizado, possivelmente simplificando a sua implementação.

Figura 34: Teor de  $O_2$  e  $CO_2$  amostrado vs estimado no cenário II.

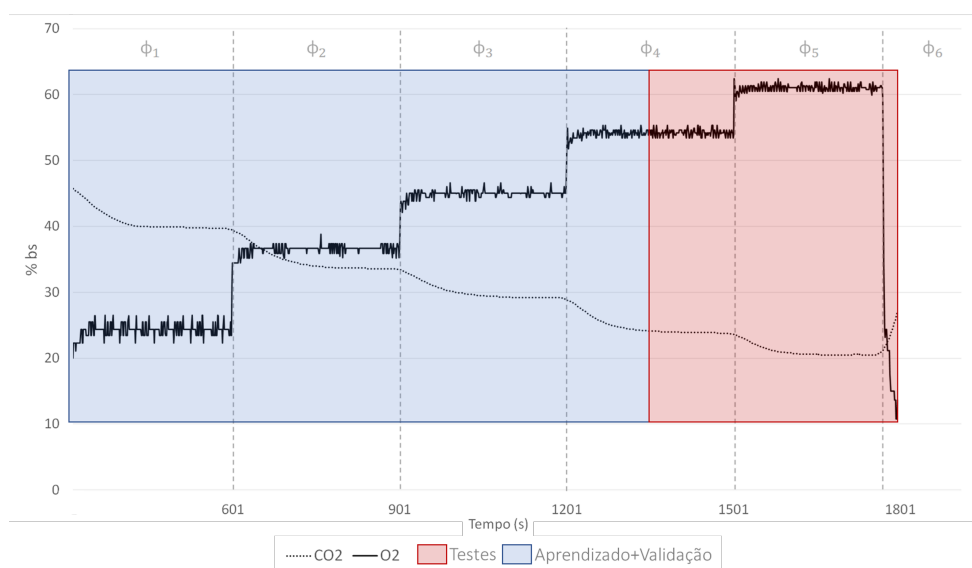


Fonte: Autor

### 7.2.3 Cenário III

No terceiro cenário, o mecanismo de inferência foi treinado até o instante  $t = 1348s$ , que representa os primeiros 70% dos dados (55% conjunto de aprendizado e 15% conjunto de validação) e os 30% restantes foram utilizados para validação (conjunto de testes), como ilustrado na Figura 35. O intuito desta vez é medir a capacidade da rede proposta de estimar resultados para novas razões de equivalência fora da faixa treinada, simulando uma condição de processo fora das faixas de trabalho esperadas.

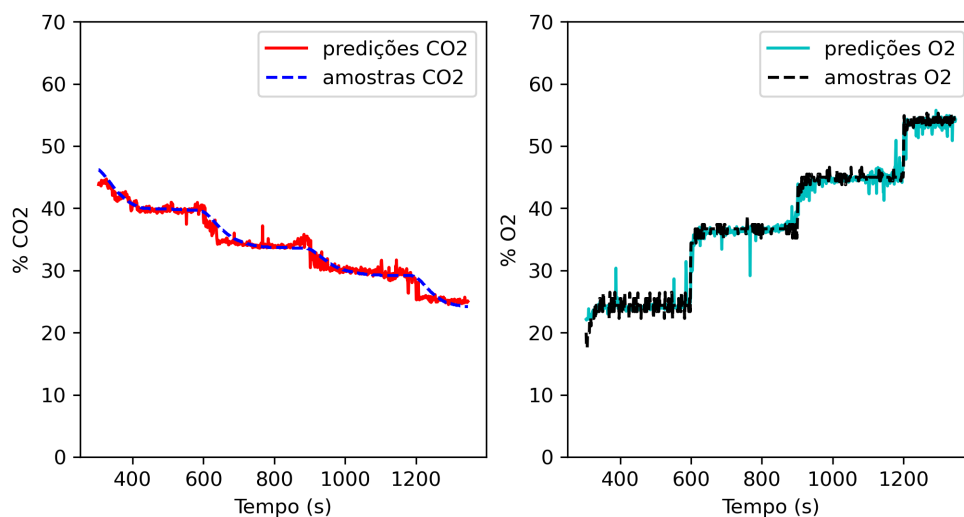
Figura 35: Divisão do conjunto de dados entre aprendizado e validação para o terceiro cenário.



Fonte: Autor

A Figura 36 representa o desempenho da rede no conjunto de aprendizado apenas para fins de comparação à Validação I, não devendo ser considerada uma métrica para a eficiência do modelo, pois os conjuntos de testes e aprendizado não devem se sobrepor. Percebe-se que no conjunto de aprendizado a rede realiza estimações bem próximas aos valores medidos, como é o esperado.

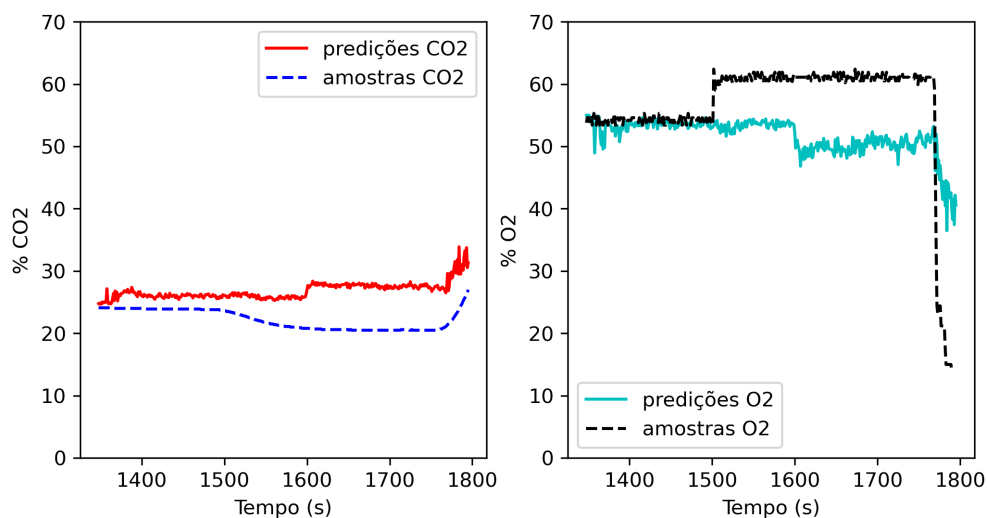
Figura 36: Teor de  $O_2$  e  $CO_2$  amostrado vs estimado ao longo do experimento.



Fonte: Autor

Na Figura 37, por outro lado, pode-se perceber que a rede perde precisão quando tenta prever teores de  $O_2$  e  $CO_2$  em razões de equivalência fora da faixa em que foi treinada, envolvendo de um erro quadrático médio igual à 2,01 para 9,67.

Figura 37: Teor de  $O_2$  e  $CO_2$  amostrado vs estimado ao longo do experimento.



Fonte: Autor

Apesar do resultado parecer negativo, é interessante notar que o método foi robusto o suficiente para estimar teores próximos aos valores amostrados, ao invés de sugerir resultados mirabolantes. Esse resultado leva a crer que mesmo em um processo ocorrendo fora das faixas de trabalho esperadas após alguma anomalia, o método ainda conseguiria continuar funcionando, apenas com menor precisão.

Conclui-se, pela boa precisão ao estimar novos valores na faixa de razões de equivalência treinadas e pela baixa capacidade da rede em gerar previsões nas razões de equivalência não treinadas, que o processo de aprendizado encontrou uma relação entre as imagens de chama e o teor de  $O_2$  e  $CO_2$  nas emissões sem modelar satisfatoriamente o processo físico de combustão. Essa conclusão deve ser considerada empírica, como notam Ballester e Hernández (2010), pois não se pode justificar fisicamente a relação entre a imagem da radiação emitida por uma chama e parâmetros do seu processo de combustão. O resultado reforça a teoria de Hestness *et al.* (2017), sugerindo que a rede performaria melhor e teria mais capacidade de inferência quando treinada em um conjunto de dados maior e mais diversificado.

## 8 CONCLUSÃO

Este trabalho investigou a viabilidade de um mecanismo de predição de emissões de  $O_2$  e  $CO_2$  baseado em redes neurais convolucionais profundas utilizando imagens de chamas da combustão de gás natural sob diferentes razões de equivalência. O modelo proposto foi treinado utilizando imagens de chama coletadas de forma offline, porém junto a uma placa de captura de imagens seria possível replicar a técnica para monitoramento em tempo real da fornalha.

O sistema apresentando foi capaz de estimar com alta precisão os teores de  $O_2$  e  $CO_2$  emitidos pelo processo de combustão nas razões de equivalência testadas utilizando apenas imagens da chama, o que indica que a rede proposta encontrou um modelo eficiente para a tarefa, além de generalizar bem para dados novos. Para as razões de equivalência treinadas, o método proposto apresentou raiz do erro quadrático médio (RMSE) absoluto de 1,19 e erro máximo absoluto de 6,21, além de poder ser treinado em poucos minutos e não requerer pré-processamento *offline* de dados, como gerar *features* manualmente. Para novas imagens em razões de equivalência fora da faixa treinada, o método proposto apresentou resultados, por um lado, menos promissores, indicando que o modelo encontrado pela rede é eficaz em reconhecer nas novas imagens características similares às apresentadas durante o seu treinamento, ao invés de modelar o processo físico-químico da combustão, porém, por outro lado, positivos, dando indícios de que mesmo em um processo fora das faixas de trabalho esperadas, o método continuaria funcionando, apenas com menor precisão. Baseado no trabalho de Hestness *et al.* (2017), que demonstra como a capacidade de redes neurais cresce de forma proporcional ao conjunto de dados, este autor estima que com um conjunto de dados maior os resultados poderiam ser ainda mais promissores. Essa afirmação é reiterada pelo trabalho de Toth, Garami e Csordás (2017), que obtiveram resultados extremamente precisos ao estimar a temperatura de saída em um caldeira utilizando



280 mil imagens e 2200 amostras.

A rede foi capaz de aprender relações intuitivas entre a imagem média gerada através de 1 segundo de vídeo do processo de combustão e o teor de CO<sub>2</sub> emitido por este, demonstrando que o monitoramento de chamas por imagens aliado a redes neurais profundas podem aprimorar o tempo de resposta e precisão de um controle de malha fechada na combustão de gás natural, tornando o controle mais simples e o processo mais eficiente. O resultado é promissor, pois, além de se demonstrar viável, o método pode ser implementado de forma mais simples que outros métodos tradicionais de monitoramento em indústrias e refinarias.

## **8.1 Sugestões para trabalhos futuros**

Com o objetivo de dar continuidade à pesquisa, abordando aspectos não estudados no presente trabalho, faz-se a seguir algumas sugestões e considerações para trabalhos futuros:

- Treinar rede em conjunto de dados maior e que englobe mais razões de equivalência.
- Investigar resultados obtidos através de redes neurais recorrentes (RNN) e sua capacidade de predição.
- Aplicar o método a dados coletados em tempo real.
- Testar controle em malha fechada utilizando o método proposto como sensor virtual.

## REFERÊNCIAS

- ABDURAKIPOV, S. *et al.* Combustion regime monitoring by flame imaging and machine learning. **Optoelectronics, Instrumentation and Data Processing**, Springer, v. 54, n. 5, p. 513–519, 2018.
- AGUIRRE, L. A. **Introdução à identificação de sistemas–Técnicas lineares e não-lineares aplicadas a sistemas reais**. [S.l.]: Editora UFMG, 2004.
- AKIWUMI, P. Least developed countries cannot afford to strand their assets, given their development challenges. 2021. Acesso em: 25 de junho de 2022. Disponível em: <<https://tinyurl.com/3m62rc54>>.
- ALPHAFOLD TEAM. Alphafold: A solution to a 50-year-old grand challenge in biology. 2021.
- BALLESTER, J.; HERNÁNDEZ, R. Image-based techniques for the monitoring of flames. **Industrial Combustion Testing**, CRC Press, p. 361–378, 2010.
- BBC. **Poluição do ar causa 5,5 milhões de mortes por ano**. 2016. Acesso em: 28 de maio de 2022. Disponível em: <[http://www.bbc.com/portuguese/noticias/2016/02/160213\\_poluicao\\_mortes\\_fn](http://www.bbc.com/portuguese/noticias/2016/02/160213_poluicao_mortes_fn)>.
- BOJARSKI, M. *et al.* End to end learning for self-driving cars. **arXiv preprint arXiv:1604.07316**, 2016.
- BP. Statistical review of world energy. 2021.
- BROOK, R. D. *et al.* Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the american heart association. **Circulation**, American Heart Association, v. 121, n. 21, p. 2331–2378, 2010.
- BROWNLEE, J. How to configure the number of layers and nodes in a neural network. **Machine Learning Mastery: Vermont, Australia**, 2018.
- CARNEIRO, R. L. *et al.* Métodos de gradiente para otimização simultânea: Estudo de casos de sistemas alimentares. Campo Mourao, 2005.
- CHAI, T.; DRAXLER, R. R. Root mean square error (rmse) or mean absolute error (mae). **Geoscientific Model Development Discussions**, v. 7, n. 1, p. 1525–1534, 2014.
- CHEN, J.; CHAN, L. L. T.; CHENG, Y.-C. Gaussian process regression based optimal design of combustion systems using flame images. **Applied energy**, Elsevier, v. 111, p. 153–160, 2013.
- CHEN, J. *et al.* Design of image-based control loops for industrial combustion processes. **Applied energy**, Elsevier, v. 94, 2012.

- CHOLLET, F. **Deep learning with Python**. [S.l.]: Simon and Schuster, 2021.
- FILHO, O. M.; NETO, H. V. **Processamento digital de imagens**. [S.l.]: Brasport, 1999.
- GAELZER, R. **Apostila de Física-Matemática**. UFRGS, 2021. Acesso em: 10 de junho de 2022. Disponível em: <<https://professor.ufrgs.br/rgaelzer/pages/fismat-apost>>.
- GONZALEZ, R. C. **Digital image processing**. [S.l.]: Pearson education india, 2009.
- HAN, Z. *et al.* Combustion stability monitoring through flame imaging and stacked sparse autoencoder based deep neural network. **Applied Energy**, Elsevier, v. 259, 2020.
- HAN, Z. *et al.* Prediction of combustion state through a semi-supervised learning model and flame imaging. **Fuel**, Elsevier, v. 289, 2021.
- HESTNESS, J. *et al.* Deep learning scaling is predictable, empirically. **arXiv preprint arXiv:1712.00409**, 2017.
- IOFFE, S.; SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: PMLR. **International conference on machine learning**. [S.l.], 2015. p. 448–456.
- JHA, A. Visualizing machine learning models: Guide and tools. 2022. Acesso em: 30 de maio de 2022. Disponível em: <<https://neptune.ai/blog/visualizing-machine-learning-models>>.
- JUMPER, J. *et al.* Highly accurate protein structure prediction with alphafold. **Nature**, Nature Publishing Group, v. 596, n. 7873, p. 583–589, 2021.
- KORN, G. A.; KORN, T. M. **Mathematical handbook for scientists and engineers: definitions, theorems, and formulas for reference and review**. [S.l.]: Courier Corporation, 2000.
- KRAUSE, J.; PERER, A.; BERTINI, E. Using visual analytics to interpret predictive machine learning models. **arXiv preprint arXiv:1606.05685**, 2016.
- LIANG, F.-Y. *et al.* The role of natural gas as a primary fuel in the near future, including comparisons of acquisition, transmission and waste handling costs of as with competitive alternatives. **Chemistry Central Journal**, BioMed Central, v. 6, n. 1, p. 1–24, 2012.
- LIU, Y.; FAN, Y.; CHEN, J. Flame images for oxygen content prediction of combustion systems using dbn. **Energy & Fuels**, ACS Publications, v. 31, n. 8, p. 8776–8783, 2017.
- MARCUS, M. **Finite Dimensional Multilinear Algebra, Part II**. 1. ed. [S.l.]: Dekker, 1975. (Monographs and Textbooks in Pure and Applied Mathematics, Volume 23).

OMIOTEK, Z.; KOTYRA, A. Flame image processing and classification using a pre-trained vgg16 model in combustion diagnosis. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 21, n. 2, p. 500, 2021.

PANAGAKIS, Y. *et al.* Tensor methods in computer vision and deep learning. **Proceedings of the IEEE**, IEEE, v. 109, n. 5, p. 863–890, 2021.

PAYNE-STURGES, D. C. *et al.* Healthy air, healthy brains: advancing air pollution policy to protect children's health. **American journal of public health**, American Public Health Association, v. 109, n. 4, p. 550–554, 2019.

PODAREANU, D. *et al.* **Best Practice Guide-Deep Learning**. [S.l.]: Feb, 2019.

RICH, E. **Artificial Intelligence**. USA: McGraw-Hill, Inc., 1983. ISBN 0070522618.

SAMMUT, C.; WEBB, G. I. **Encyclopedia of machine learning**. [S.l.]: Springer Science & Business Media, 2011.

SHANNON, C. E. Communication in the presence of noise. **Proceedings of the IRE**, IEEE, v. 37, n. 1, p. 10–21, 1949.

SILVA NETO, G. C. *et al.* Identification of  $CO_2$  and  $O_2$  emissions dynamics in a natural gas furnace through flame images, armax models, and kalman filtering. **Journal of the Brazilian Society of Mechanical Sciences and Engineering**, Springer, v. 43, n. 5, p. 1–13, 2021.

SILVA NETO, G. C. d. **Projeto de um sensor virtual baseado em modelos ARMAX e filtro de Kalman para predição das emissões de  $O_2/CO_2$  aa partir de imagens de chamas em um forno industrial**. Tese (Doutorado) — Universidade de São Paulo, 2021.

SRIVASTAVA, N. *et al.* Dropout: a simple way to prevent neural networks from overfitting. **The journal of machine learning research**, JMLR. org, v. 15, n. 1, p. 1929–1958, 2014.

SZANDAŁA, T. Review and comparison of commonly used activation functions for deep neural networks. In: **Bio-inspired neurocomputing**. [S.l.]: Springer, 2021. p. 203–224.

TOTH, P.; GARAMI, A.; CSORDÁS, B. Image-based deep neural network prediction of the heat output of a step-grate biomass boiler. **Applied energy**, Elsevier, v. 200, p. 155–169, 2017.

TURING, A. M.; HAUGELAND, J. Computing machinery and intelligence. **The Turing Test: Verbal Behavior as the Hallmark of Intelligence**, p. 29–56, 1950.

WHITE, A. D. **Deep Learning for Molecules and Materials**. [s.n.], 2021. Acesso em: 28 de maio de 2022. Disponível em: <<https://dmol.pub>>.