



**UFAM**

**UNIVERSIDADE FEDERAL DO AMAZONAS**

**FACULDADE DE TECNOLOGIA - FT**

**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA -  
PPGEE**

Roberlânio de Oliveira Melo

**Aplicação de redes neurais convolucionais para detectar  
vazamentos de gás natural em imagens de cabeça de  
poço petrolíferos *onshore***

Manaus – AM

2021

Roberlânio de Oliveira Melo

**Aplicação de redes neurais convolucionais para detectar vazamentos de gás natural em imagens de cabeça de poço petrolíferos *onshore***

Tese apresentada no Programa de Pós-Graduação em Engenharia Elétrica da Faculdade de Tecnologia como requisito parcial para obtenção do título de Doutor em Engenharia Elétrica.

Universidade Federal do Amazonas

Orientador: Dr. Cícero Ferreira Fernandes Costa Filho  
Coorientadora: Dra. Marly Guimarães Fernandes Costa

Manaus- AM

2021

## Ficha Catalográfica

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

M528a Melo, Roberlânio de Oliveira  
Aplicação de redes neurais convolucionais para detectar vazamentos de gás natural em imagens de cabeça de poço petrolíferos onshore / Roberlânio de Oliveira Melo . 2021  
137 f.: il. color; 31 cm.

Orientador: Cícero Ferreira Fernandes Costa Filho  
Coorientadora: Marly Guimarães Fernandes Costa  
Tese (Doutorado em Engenharia Elétrica) - Universidade Federal do Amazonas.

1. Vazamento de gás. 2. Detecção de gás. 3. Rede neural convolucional. 4. Imagem de cabeça de poço onshore. 5. Indústria de gás e petróleo. I. Costa Filho, Cícero Ferreira Fernandes. II. Universidade Federal do Amazonas III. Título

ROBERLÂNIO DE OLIVEIRA MELO

**APLICAÇÃO DE REDES NEURAS CONVOLUCIONAIS PARA DETECTAR  
VAZAMENTOS DE GÁS NATURAL EM IMAGENS DE CABEÇA DE POÇO  
PETROLÍFEROS ONSHORE**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Doutor em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Aprovado em 03 de dezembro de 2021.

**BANCA EXAMINADORA**

*Cícero F. F. Costa Filho*

Prof. Dr. Cícero Ferreira Fernandes Costa Filho, Presidente

Universidade Federal do Amazonas

*José Kaimundo Gomes Pereira*

Prof. Dr. José Kaimundo Gomes Pereira, Membro

Universidade Federal do Amazonas

*Clahildek Matos Xavier*

Prof. Dr. Clahildek Matos Xavier, Membro

Sistema de Proteção da Amazônia- SIPAM

*Fabiola Guerra Nakamura*

Profª. Dra. Fabiola Guerra Nakamura

Universidade Federal do Amazonas

*Jozias Parente de Oliveira*

Prof. Dr. Jozias Parente de Oliveira

Universidade Estadual do Amazonas



*Para minhas queridas perolas: Maria Eduarda e Giane Zupellari.  
As maiores relíquias da minha vida.*



# Agradecimentos

Esta tese é o resultado de anos de estudos. Esse resultado só foi possível graças ao apoio e incentivo de várias pessoas e instituições, a quem não posso deixar de agradecer.

Um agradecimento especial ao meu Orientador. Professor Dr. Cícero Ferreira Fernandes Costa Filho, obrigado pela orientação, ética, paciência e magnífica capacidade de escuta, mesmo naqueles momentos que eu relutava. O seu conhecimento, experiência profissional e a capacidade de encorajar o orientando a vencer novos desafios, são características marcantes, as quais me inspirou a seguir forte e com foco, principalmente durante as madrugadas de estudos.

A minha Coorientadora, Professora Dra. Marly Guimarães Fernandes Costa, uma professora com uma habilidade de raciocínio formidável e ágil, agradeço por todo o empenho e disponibilidade em sempre me ajudar e fornecer as principais orientações.

Nada disso seria possível sem o apoio e compreensão da minha esposa, Giane Zupellari, e da minha filha, Maria Eduarda, as quais agradeço do fundo do meu coração por estarem sempre ao meu lado, diuturnamente, dando-me forças para conquistar mais um degrau na jornada da minha vida.

Então, gostaria de agradecer aos funcionários, aos meus amigos e amigas do PPGEE da UFAM por todo o apoio recebido.

Esta jornada começou na Universidade Federal de Santa Catarina (UFSC), no ano de 2018. Portanto, agradeço aos professores e amigos daquela universidade que me encorajaram a nunca desistir do doutoramento.

Agradeço imensamente a Universidade Federal do Amazonas (UFAM), especialmente ao Centro de P&D em Tecnologia Eletrônica da Informação (CETELI), que disponibilizou seu espaço físico e seus computadores para o desenvolvimento desta tese.

Obrigado, DEUS!



*Tudo tem o seu tempo determinado, e  
há tempo para todo o propósito  
debaixo do céu.  
Eclesiastes 3:1*



## Resumo

Grandes vazamentos de gás natural levam a acidentes que muitas vezes são letais para os humanos e altamente destrutivos para a propriedade, além de liberar metano na atmosfera, impactando negativamente nosso meio ambiente. Logo, a detecção de vazamentos de gás natural é uma ação importante da indústria do petróleo para a prevenção de acidentes. A literatura fornece diferentes técnicas para detecção de vazamentos de gás natural. Porém, com exceção de estudos anteriores desenvolvidos pelo autor desta tese sobre o tema, ainda existe uma lacuna na literatura sobre detecção de nuvens de vazamentos de gás natural por meio de imagens digitais, sem a necessidade de sensores ou câmeras especiais calibradas para o espectro de moléculas de metano. Nos estudos anteriores foram utilizadas técnicas de processamento de imagem associadas a um classificador de filtro de novidade para investigar a presença ou ausência de nuvem visível de vapores de hidrocarbonetos, ou seja, uma pluma de gás natural em quadros de circuito fechado de televisão instalados em cabeças de poços petrolíferos terrestres. Nesta tese, é apresentado um novo método de detecção de nuvens visível de vapores de hidrocarbonetos que é constituído de redes neurais convolucionais, as quais são aplicadas para classificar imagens (*frames* do circuito fechado de televisão) como pertencentes às classes com ou sem vazamento de gás natural em cabeças de poço *onshore*, aprimorando os resultados obtidos anteriormente. Com o objetivo de realizar a tarefa de detecção de vazamentos, o estudo propõe e apresenta o resultado de três arquiteturas de redes neurais convolucionais, além de investigar a técnica de transferência de conhecimento com as arquiteturas pré-treinadas DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18, e também a técnica de *Feature Fusion* aplicada a estes modelos pré-treinados. O estudo fornece ainda uma explicação visual provendo a região da imagem de entrada que foi relevante para o classificador tomar a decisão de que ocorreu um vazamento, através da técnica *Gradient-weighted Class Activation Mapping*. Os resultados experimentais mostraram que o modelo de melhor desempenho apresentou acurácia de 100,00% e taxa de falsos negativos de 0,00%, superando o desempenho de métodos anteriores de detecção de vazamento de gás natural por imagens de circuito fechado de televisão.

**Palavras-chave:** vazamento de gás; detecção de gás; rede neural convolucional; imagem de cabeça de poço *onshore*; indústria de gás e petróleo.



## Abstract

Large natural gas leaks lead to accidents that often are lethal to humans and highly destructive to property, as well as release methane into the atmosphere, negatively impacting our environment. Therefore, the detection of natural gas leaks is an important action of the oil industry to prevent accidents. The literature provides different techniques for detecting natural gas leaks. However, except for previous studies proposed by the author of this thesis on the subject, there is still a gap in the literature on the detection of clouds of natural gas leaks through digital images, without the need for sensors or specific cameras calibrated for the spectrum of methane molecules. In previous studies, image processing techniques associated with a novelty filter classifier were used to investigate the presence or absence of a visible cloud of hydrocarbon vapors, that is, a plume of natural gas in closed circuit television frames installed on terrestrial oil wellheads. In this thesis a new method for detecting visible clouds of hydrocarbon vapors is presented, which consists of convolutional neural networks, which are applied to classify images (closed circuit television frames) as belonging to classes with or without leakage of natural gas in onshore wellheads, improving the results obtained previously. To carry out the leak detection task, the study proposes and presents the results of the three convolutional neural networks architectures, in addition to investigating the learning transfer technique with the pre-trained architectures DenseNet-201, GoogLeNet, MobileNetV2 and ResNet-18, and also the feature fusion technique between these pre-trained models. The study also provides a visual explanation providing the region of the input image that was relevant for the classifier to decide that a leak occurred, through the Gradient-weighted Class Activation Mapping Algorithm. The experimental results showed that the best performing model presented an accuracy of 100.00% and a false negative rate of 0.00%, surpassing the performance of previous methods of detecting leakage of natural gas by closed circuit television images.

**Keywords:** gas leak; gas detection; convolutional neural network; onshore wellhead image; gas and oil industry.



## Lista de ilustrações

Figura 1.1 - Operações de petróleo e gás natural: oportunidades de mitigação de metano. Fonte: Adaptada e traduzida de relatório da Environmental Protection Agency (2011-2019 GREENHOUSE GAS REPORTING PROGRAM INDUSTRIAL PROFILE: PETROLEUM AND NATURAL GAS SYSTEMS, 2018). .....	32
Figura 2.1 - Consumo mundial de energia por tipo de combustível no ano de 2020. Fonte: Adaptada de Statistical Review of World Energy (MURVAY; SILEA, 2012; ZAMAN et al., 2020). .....	37
Figura 2.2 - Geologia esquemática dos recursos de gás natural. Fonte: Adaptada e traduzida de Mokhatab, Poe e Mak (2018). .....	40
Figura 2.3 - Estratégias de gerenciamento de vazamentos hidráulicos. Fonte: Adaptada e traduzida de Murvay e Silea (2012) e Zaman et al. (2020). .....	42
Figura 2.4 - Estratégias de gerenciamento de vazamento hidráulicos. Fonte: Adaptada e traduzida de Murvay e Silea (2012), Adegboye, Fung e Karnik (2019) e Zaman et al. (ZAMAN et al., 2020). .....	44
Figura 2.5 - Sensor baseado em combustão catalítica: (a) elementos catalisadores ativo (1) e de referência (2) e (b) Ponte Wheatstone com elemento catalítico. Fonte: Adaptada e traduzida de Wang et al.(2011).....	47
Figura 2.6 - Estratégias de gerenciamento de vazamento hidráulicos. Fonte: Adaptada e traduzida de Murvay e Silea (2012), Adegboye, Fung e Karnik (2019) e Zaman et al. (ZAMAN et al., 2020). .....	48
Figura 2.7 - Proposta de Wang et al. (2020) para predição de vazamento de GN através de imagens térmicas: (a) instalação do estudo, um separador multifásico; (b) imagem do vazamento de gás natural obtida pela câmera infravermelha e (c) subtração de fundo da figura (b). Fonte: Adaptada de Wang et al. (2020).....	49
Figura 2.8 - Proposta de Melo et al. (2013) para predição de vazamento de GN através de imagens de CFTV: (a) imagem original com vazamento de gás próximo ao nível do solo e (b) imagem de novidade correspondente da Fig. 3 (a). Fonte: Costa Filho, Melo e Costa (2013). .....	50
Figura 2.9 - Ilustração do conceito de filtro de novidade no espaço $R^3$ . Conceito usando o método de ortogonalização de Gram-Schmidt. Fonte: Costa Filho, Melo e Costa(2012). .....	50

Figura 3.1 - Rede neural convolucional, aplicada às imagens de dígitos manuscritos, proposta por LeCun em 1989, onde o autor já menciona que os coeficientes que compõem o kernel $5 \times 5$ eram determinados automaticamente pela rede. Fonte: LeCun et al. (1989). .....	54
Figura 3.2 - LeNet-5: Rede neural convolucional para reconhecimento de dígitos de 1998. Fonte: LeCun et al. (1989). .....	54
Figura 3.3 - Processo de convolução realizado entre a entrada $I$ de dimensão $4 \times 4$ e filtro de dimensão $3 \times 3$ , passo de $1 \times 1$ e sem acréscimos de zeros nas extremidades do tensor. Esta configuração resulta na saída da função convolucional $C$ de dimensão $2 \times 2$ , a obtenção de cada valor de $C$ é expressa nas figuras a, b, c e d, respectivamente. Fonte: Compilação do autor. ....	57
Figura 3.4 - Processo de convolução entre tensor $I$ e kernel $K$ com acréscimo de zeros de dimensão 1. Fonte: Compilação do autor. ....	58
Figura 3.5 - Comportamento da função de ativação ReLU. Fonte: Adaptada de Aggarwal (2018). .....	60
Figura 3.6 - Resultado da aplicação da função de ativação ReLU ao mapa de característica da Figura 3.4 (d). Fonte: Compilação do autor. ....	60
Figura 3.7 - Representação visual de duas camadas de amostragem (max e avg): (a) mapa de característica da camada anterior que é entrada para camada de amostragem; (b) saída da camada de amostragem max e (c) saída da camada de amostragem avg. Fonte: Adaptada de Ozkaya, Ozturk e Barstugan (2020). ....	61
Figura 3.8 - Dropout em Redes Neurais: (a) modelo de rede neural; (b) dropout aplicado ao modelo com probabilidade de 0,5 de abandono de neurônios (x indica os neurônios abandonados). Fonte: Adaptada de Srivastava et al. (2014). ....	62
Figura 3.9 - Processo de transferência de aprendizagem aplicado as imagens do wellhead, incluindo: (a) carregamento de um modelo pré-treinado e remoção das quatro últimas camadas fully connected (FC) FC6 e F7; softmax e classification); (b) substituição das camadas removidas por camadas novas com valores de pesos aleatórios; (c) ajustar as camadas anteriores e treinar as novas camadas com o conjunto de dados do wellhead. Fonte: Adaptada e traduzida de Huang et al. (2019). .....	64
Figura 3.10 - Mapeamento de ativação de classe: a pontuação de classe prevista $c$ é mapeada na última camada convolucional $A$ , gerando os mapas de ativação de classe (CAMs)	

que são redimensionados, destacando as regiões discriminativas específicas da classe <i>c</i> . Fonte: Adaptada de Zhou et al. (2016).....	66
Figura 3.11 - Uma visão geral do método Grad-CAM. Fonte: Adaptado e traduzido de Selvaraju et al.(2017) e Chattopadhyay et al. (2018). .....	67
Figura 3.12 - Tarefa de classificação binária: esferas azuis ou esferas vermelhas. Fonte: Compilação do autor.....	69
Figura 3.13 – Exemplo de CNN com overfitting: (a) desempenho da CNN após treinamento e validação e (b) desempenho da mesma CNN ao ser testada com novo conjunto de dados, resultando em falhas de falso positivo (FP) e falso negativo (FN). Fonte: Compilação do autor.....	69
Figura 4.1 - Câmera do CFTV: CCD FC15U da Yokogawa®. Fonte: Fabricante Yokogawa (2009).....	71
Figura 4.2 - Amostra de imagens do conjunto de dados de imagens db1: (a) Imagem diurna sem vazamento de gás (sem a nuvem de hidrocarbonetos), (b) Imagem diurna com vazamento de gás (com a nuvem de hidrocarbonetos), (c) Imagem noturna sem vazamento de gás (sem a nuvem de hidrocarbonetos ), (d) Imagem noturna com vazamento de gás (com a nuvem de hidrocarbonetos). Fonte: Melo, Costa e Costa Filho (2014). .....	72
Figura 4.3 - Imagens de amostra do segundo conjunto de dados (db2) das novas cabeças de poço, em destaque na cor amarela: (a) Imagem panorâmica da cabeça de poço “X”, (b) Imagem de perfil da cabeça de poço “Y”, (c) Imagem panorâmica da cabeça de poço “W” e (d) imagem de fundo da cabeça do poço “Z”. Fonte: Compilação do autor.....	73
Figura 4.4 - Visão geral da proposta. Fonte: Compilação do autor.....	75
Figura 4.5 - Visualização de exemplos de transformações aleatórias aplicadas a quatro imagens db1: (a) rotação de + 20°; (b) rotação de -20 °; (c) translação de ± 3 pixels; (d) dimensionamento da cena com fator 0,8. Fonte: Compilação do autor.....	76
Figura 4.6 - Arquiteturas de rede propostas: a) CNL, com 18 camadas de convolução; b) CNM, com 6 camadas de convolução; c) CNS, com 2 camadas de convolução Fonte: Adaptada de Miyagawa et al.(2019).....	78
Figura 4.7 - Exemplo de um DenseNet com três blocos densos. As camadas entre dois blocos adjacentes são chamadas de camadas de transição e mudam os tamanhos do mapa de características, por meio de convolução e amostragem. Fonte: Adaptada e traduzida de (HUANG et al., 2017).....	81

Figura 4.8 - Exemplo da arquitetura de um módulo Inception da CNN GoogLeNet. Fonte: Adaptada e traduzida de Szegedy et al. (2015).	81
Figura 4.9 - Ideia original dos blocos residuais das CNN do tipo ResNet. Fonte: Adaptada e traduzida de He et al. (2016).	82
Figura 4.10 - Estrutura do bloco de construção do MobileNetV2. Fonte: Adaptada e traduzida de Sandler et al. (2018).	83
Figura 4.11 - Arquitetura da CNN FF proposta, constituída pela fusão de característica das seguintes redes CNN pré-treinadas: DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18. Fonte: Compilação do autor.	85
Figura 4.12 - Método proposto de fusão de característica das CNN DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18. Fonte: Compilação do autor.	86
Figura 4.13 - Diagrama de blocos para treinamento, avaliação e teste: (a) Treinamento e validação dos 27 modelos de CNN para obtenção do modelo MMD, treinamento das 4 CNNs pré-treinadas, treinamento da rede FF; (b) Teste do MMD, das 4 CNNs pré-treinadas e da rede FF com o conjunto de dados db1_test e obtenção do Melhor Modelo Geral (MMG); (c) Teste do MMG com o conjunto de dados db2. Fonte: Compilação do autor.	87
Figura 4.14 - Camadas substituídas novas layers, conforme método de transfer learning: (a) DenseNet-201; (b) GoogLeNet, (c) ResNet-18 e (d) MovileNetV2. Fonte: Compilação do autor.	94
Figura 5.1 - Esquemático das etapas percorridas na seção de resultados e discussão. Fonte: Compilação do autor.	98
Figura 5.2 - Treinamento e validação do MMD: (a) acurácia durante o treinamento e validação; (b) perda durante o treinamento e validação. Fonte: Compilação do autor.	100
Figura 5.3 - Imagem com vazamento real classificada pelas redes DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18, como sem vazamento. Fonte: Compilação do autor.	104
Figura 5.4 - (a), (b), (c) e (d) acurácia durante o treinamento e validação da DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18, respectivamente. Fonte: Compilação do autor.	104
Figura 5.5 - Gráfico de comportamento do modelo FF durante o treinamento e validação: (a) acurácia e (b) perda. Fonte: Compilação do autor.	105
Figura 5.6 – As imagens (a) e (b) são exemplos dos elementos que constituem o conjunto de dados db2. Fonte: Compilação do autor.	109

Figura 5.7 - (a) exemplo de uma imagem original com vazamento; (b) mapa de calor com magnitudes pré-processadas; (c) identificação dos pixels que contribuíram para a previsão de vazamento. Os pixels com cor vermelha contribuíram fortemente para a pontuação da classe com vazamento. Fonte: Compilação do autor..... 110



## Lista de quadros

Quadro 4.1 - Descrição das arquiteturas DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18. ....	83
--	----



## Lista de tabelas

Tabela 2.1 - Análise elementar dos constituintes do petróleo. ....	39
Tabela 2.2 - Composição típica do gás natural associado e não associado. ....	41
Tabela 2.3 - Algumas tarefas que utilizaram as CNNs que como máquinas de reconhecimento de padrões. ....	43
Tabela 2.4 - Conjuntos de dados de imagens de cabeça de poço mostrando o número de imagens com vazamentos e sem vazamentos. ....	51
Tabela 4.1 - Conjuntos de dados de imagens de cabeça de poço mostrando o número de imagens com vazamentos e sem vazamentos. ....	74
Tabela 4.2 - 27 modelos de CNN propostos. ....	80
Tabela 4.3 - Parâmetros e hiperparâmetros adotados no desenvolvimento, treinamento, validação e teste das CNNs. ....	88
Tabela 4.4 – Modelo da matriz de confusão. ....	88
Tabela 4.5 - Tabela de contingência usada para calcular $\chi^2$ . ....	90
Tabela 5.1 - Métricas de desempenho dos 27 modelos CNN obtidas através do conjunto de dados <i>db1_valid</i> : acurácia (ACC), Taxa de Falsos Negativos (TFN), Sensibilidade, F1-Score, <i>Chi-quadrado</i> $\chi^2$ e tempo do treinamento e validação. ....	99
Tabela 5.2 - Desempenho do Teste ( a ), referente ao modelo MMD, quando testado com o <i>db1_test</i> . ....	101
Tabela 5.3 - Métricas de desempenho das CNN DenseNet-201, GoogLeNet, MobileNetV2, ResNet-18 e FF, obtidas com o conjunto de dados <i>db1_valid</i> : acurácia (ACC), Taxa de Falsos Negativos (TFN), Sensibilidade, F1-Score, <i>Chi-square</i> $\chi^2$ e tempo do treinamento e validação. ....	102
Tabela 5.4 - Desempenho das CNN DenseNet-201, GoogLeNet, MobileNetV2, ResNet-18 e FF, obtidos com o conjunto de dados <i>db1_test</i> mostrando métricas: acurácia (ACC), Taxa de Falsos Negativos (TFN), Sensibilidade, F1-Score, <i>Chi-square</i> $\chi^2$ e tempo do treinamento e validação Desempenho dos modelos. ....	103
Tabela 5.5 - Resumo das métricas finais dos modelos MMD, DenseNet-201, GoogLeNet, MobileNetV2, ResNet-18 e FF obtidas com o conjunto <i>db1_test</i> . ....	106
Tabela 5.6 - Matriz de confusão da CNN MMD, obtida com o conjunto <i>db1_test</i> . ....	106
Tabela 5.7 - Matriz de confusão da CNN DenseNet-201, obtida com o conjunto <i>db1_test</i> . .	107
Tabela 5.8 - Matriz de confusão da CNN GoogLeNet, obtida com o conjunto <i>db1_test</i> . ....	107

Tabela 5.9 - Matriz de confusão da CNN MobileNetV2, obtida com o conjunto <i>db1_test</i> ...	107
Tabela 5.10 - Matriz de confusão da CNN ResNet-18, obtida com o conjunto <i>db1_test</i> .....	107
Tabela 5.11 - Matriz de confusão da CNN FF, obtida com o conjunto <i>db1_test</i> . ....	108
Tabela 5.12 - Comparação entre os resultados obtidos com a CNN FF (MMG) e com o Filtro de Novidade. ....	108
Tabela 5.13 - Matriz de confusão da CNN FF (MMG), obtida com o conjunto <i>db2</i> . Fonte: Compilação do autor. ....	109
Tabela 5.14 - Desempenho da CNN FF obtido com o conjunto de dados <i>db2</i> .....	109

## Lista de abreviaturas e siglas

ACC	Acurácia
ADAM	<i>Adaptive Moment Estimation</i>
AI	<i>Artificial Intelligence</i>
AM	Estado do Amazonas
AUC	<i>Area Under Curve</i>
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CETELI	<i>Center for R&amp;D in Electronics and Information</i>
CFTV	Circuito Fechado de Televisão
CH <sub>4</sub>	Carbono-Hidrogênio (ligação química, metano)
CIFA	<i>Canadian Institute for Advanced Research</i>
CNN	<i>Convolutional Neural Network</i>
CNNL	Convolutional Neural Network Longa
CNNM	Convolutional Neural Network Média
CNNS	Convolutional Neural Network Sintética
CO <sub>2</sub>	Carbono-Oxigênio (ligação química, dióxido de carbono)
COVID	<i>Coronavirus Disease</i>
EPA	<i>Environmental Protection Agency</i>
FC	<i>Fully Connected</i>
FF	<i>Feature Fusion</i>
FN	Falsos Negativos
FP	Falsos Positivos
FT	Faculdade de Tecnologia
GHG	<i>Greenhouse Gas</i>
GHGRP	<i>Greenhouse Gas Reporting Program</i>
GN	Gás Natural
GNL	Gás Natural Liquefeito
GPU	<i>Graphics Processing Unit</i>
Grad-CAM	<i>Gradient-weighted Class Activation Mapping</i>
IEEE	<i>Institute of Electrical and Electronic Engineers</i>
ILSVRC	<i>ImageNet Large Scale Visual Recognition Challenge</i>
IR	<i>Infrared</i>
IVOCT	<i>Intravascular Optical Coherence Tomography</i>

MMD	Modelo com Melhor Desempenho
MMG	Melhor Modelo Geral
MMT CO <sub>2</sub> e	Milhões de Toneladas Métricas de Dióxido de Carbono Equivalente
MWIR	<i>Mid Wave Infrared</i>
PPGEE	Programa de Pós-Graduação em Engenharia Elétrica
ReLU	<i>Rectified Linear Unit</i>
RGB	<i>Red Green Blue</i> (modelo de cores baseado em cores primárias aditivas)
RMSProp	<i>Root Mean Square Propagation</i>
RNA	Rede Neural Artificial
RNC	Rede Neural Convolutacional
ROC	<i>Receiver Operating Characteristic</i>
SGDM	<i>Stochastic Gradient Descent with Momentum</i>
TFN	Taxa de Falsos Negativos
TL	<i>Transfer Learning</i>
UEA	Universidade do Estados do Amazonas
UFAM	Universidade Federal do Amazonas
VN	Verdadeiros Negativos
VP	Verdadeiros Positivos

## Sumário

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>31</b>
<b>1.1</b>	<b>Objetivo geral .....</b>	<b>34</b>
<b>1.2</b>	<b>Objetivos específicos.....</b>	<b>34</b>
<b>1.3</b>	<b>Organização da tese.....</b>	<b>35</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA.....</b>	<b>37</b>
<b>2.1</b>	<b>A importância econômica do petróleo e gás natural .....</b>	<b>37</b>
<b>2.2</b>	<b>O petróleo.....</b>	<b>38</b>
<b>2.3</b>	<b>Gás natural.....</b>	<b>39</b>
<b>2.4</b>	<b>Confiabilidade e manutenibilidade de sistemas para prevenção de vazamentos de gás natural .....</b>	<b>40</b>
<b>2.5</b>	<b>Técnicas contemporâneas de detecção de vazamento de GN .....</b>	<b>44</b>
2.5.1	Cães treinados.....	45
2.5.2	Humanos experientes .....	45
2.5.3	Veículo Autônomo/Drone/Robô .....	45
2.5.4	Monitoramento acústico .....	45
2.5.5	Método baseado em combustão catalítica .....	46
2.5.6	Método baseado em absorção do espectro infravermelho.....	47
2.5.7	Balanço de massa / volume .....	48
2.5.8	Pressão negativa .....	48
<b>2.6</b>	<b>Técnicas de detecção de vazamento de GN por imagem .....</b>	<b>48</b>
<b>2.7</b>	<b>Conclusão .....</b>	<b>51</b>

<b>3.</b>	<b>FUNDAMENTAÇÃO TEÓRICA.....</b>	<b>53</b>
<b>3.1</b>	<b>Redes neurais artificiais .....</b>	<b>53</b>
<b>3.2</b>	<b>Redes neurais convolucionais .....</b>	<b>54</b>
3.2.1	Camada convolucional.....	55
3.2.2	Camada de normalização em lote .....	58
3.2.3	Função de ativação retificadora linear ( <i>Rectifier Linear Unit - ReLU</i> ).....	59
3.2.4	Camada de amostragem ( <i>pooling</i> ) .....	61
3.2.5	Camada <i>dropout</i> .....	62
3.2.6	Camadas totalmente conectadas ( <i>fully-connected layers</i> ) .....	63
<b>3.3</b>	<b><i>Transfer Learning (TL)</i>.....</b>	<b>63</b>
<b>3.4</b>	<b>Grad-CAM .....</b>	<b>65</b>
<b>3.5</b>	<b>Aumento de dados.....</b>	<b>67</b>
<b>3.6</b>	<b>Conclusão.....</b>	<b>70</b>
<b>4.</b>	<b>MATERIAIS E MÉTODOS .....</b>	<b>71</b>
<b>4.1</b>	<b>Materiais .....</b>	<b>71</b>
4.1.1	Ambiente de desenvolvimento.....	71
4.1.2	Conjuntos de dados .....	72
4.1.2.1	Primeiro conjunto de dados: db1 .....	72
4.1.2.2	Segundo conjunto de dados: db2 .....	73
<b>4.2</b>	<b>Métodos.....</b>	<b>74</b>
4.2.1	Pré-processamento de dados .....	76
4.2.2	Aumento de dados .....	76
4.2.3	Arquitetura das CNN e parâmetros de treinamento.....	76
4.2.3.1	Propostas de arquiteturas de CNN .....	77
4.2.3.2	Quatro CNNs pré-treinadas utilizando a técnica TL .....	80
4.2.3.3	CNN FF baseada em fusão de características .....	84
4.2.4	Treinamento, validação e teste e hiperparâmetros de treinamento .....	86

4.2.5	Avaliação das CNN: métricas de desempenho.....	88
4.2.6	Transferência de conhecimento.....	91
4.2.7	Realce da área com vazamento .....	95
<b>4.3</b>	<b>Conclusão .....</b>	<b>95</b>
<b>5.</b>	<b>RESULTADOS E DISCUSSÃO .....</b>	<b>97</b>
<b>5.1</b>	<b>Resultados do treinamento e validação das CNN.....</b>	<b>98</b>
5.1.1	Resultados dos 27 modelos de CNN propostos.....	98
5.1.2	Resultados das CNN pré-treinados e CNN FF.....	101
<b>5.2</b>	<b>Determinação do Melhor Modelo Geral (MMG) .....</b>	<b>105</b>
<b>5.3</b>	<b>Comparação dos resultados da CNN MMG e filtro de novidade .....</b>	<b>108</b>
<b>5.4</b>	<b>Avaliação da capacidade de generalização do MMG (FF).....</b>	<b>109</b>
<b>5.5</b>	<b>Explicação visual .....</b>	<b>110</b>
<b>5.6</b>	<b>Conclusão .....</b>	<b>110</b>
<b>6.</b>	<b>CONCLUSÕES .....</b>	<b>113</b>
<b>6.1</b>	<b>Propor e avaliar o desempenho de três diferentes arquiteturas CNN e selecionar aquela que oferece o melhor desempenho na detecção de vazamentos de gás natural..</b>	<b>113</b>
<b>6.2</b>	<b>Avaliar o desempenho de redes pré-treinadas DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18, quanto a detecção de vazamentos de gás natural .....</b>	<b>113</b>
<b>6.3</b>	<b>Propor e avaliar o desempenho de um modelo CNN baseado em fusão de características (CNN FF), obtido por meio da extração simultânea de características das redes DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18, quanto a detecção de vazamentos de gás natural.....</b>	<b>114</b>
<b>6.4</b>	<b>Comparar o desempenho das CNN MMD, DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18 e FF, com respeito, principalmente, a acurácia e a taxa de falso negativos fornecidas por essas CNNs, e assim definir o Melhor Modelo Geral (CNN MMG), quanto a detecção de vazamentos de gás natural .....</b>	<b>114</b>
<b>6.5</b>	<b>Comparar o desempenho do modelo baseado em CNN com o do modelo baseado em filtro de novidade através do teste Chi-quadrado .....</b>	<b>114</b>

<b>6.6</b>	<b>Investigar a capacidade do melhor modelo geral em distinguir nuvens do céu de plumas de GN através da avaliação de imagens de diferentes cabeças de poço, obtidas em diversas condições de ambiente, desconhecidas do modelo .....</b>	<b>115</b>
<b>6.7</b>	<b>Limitações e desafios .....</b>	<b>115</b>
<b>6.8</b>	<b>Trabalhos futuros .....</b>	<b>115</b>
	<b>REFERÊNCIAS .....</b>	<b>117</b>
	<b>APÊNDICE 1.....</b>	<b>127</b>
	<b>APÊNDICE 2.....</b>	<b>137</b>

## 1 INTRODUÇÃO

Com a revolução industrial é dado início a uma nova era no setor de energia com o ganho em popularidade dos combustíveis fósseis. Desde então, os combustíveis fósseis desempenham um papel de destaque no mercado mundial de energia, principalmente o Gás Natural (GN) e os derivados do Petróleo, como óleo combustível, gasolina e óleo diesel.

O gás natural, extraído de poços petrolíferos *onshore*, também conhecidos como *wellhead* ou árvore de natal, é composto por metano ( $\text{CH}_4$ ), etano, propano, butano e pentano e outros hidrocarbonetos, além de compostos não hidrocarbonetos (água, sólidos, dióxido de carbono etc.). No entanto, 70% do GN é composto de metano (MCCAIN, 2017).

As emissões de  $\text{CH}_4$  para a atmosfera podem ser convertidas em emissões equivalentes de dióxido de carbono ( $\text{CO}_2$ ) (CALABRÒ, 2009), as quais contribuem para o aumento do efeito estufa (FORSTER et al., 2007; ICF INTERNATIONAL, 2015; MAGEE, 2015). Em um horizonte de tempo relativamente curto, considerando que o potencial de aquecimento global do  $\text{CH}_4$  é 25 vezes maior do que o do  $\text{CO}_2$ , em um período de 100 anos, a liberação do  $\text{CH}_4$  para o meio ambiente resultará no aumento relevante do efeito estufa global (FORSTER et al., 2007; ICF INTERNATIONAL, 2015; MYHRE et al., 2013).

Em 2020, a Agência de Proteção Ambiental dos Estados Unidos da América (EPA, do inglês, *Environmental Protection Agency*) publicou os dados de gases de efeito estufa (GHG, do inglês, *Greenhouse Gas*) de 2019 para Sistemas de Petróleo e GN coletados no Programa de Relatórios de Gases de Efeito Estufa (GHGRP, do inglês, *Greenhouse Gas Reporting Program*). Esses dados demonstram que emissões de GHG em 2019, em 2.350 instalações que realizam atividades de Sistemas de Petróleo e GN, como produção, coleta e dinamização, processamento, transmissão e distribuição, totalizaram 341 MMT  $\text{CO}_2e$  (milhões de toneladas métricas de dióxido de carbono equivalente) (2011-2019 GREENHOUSE GAS REPORTING PROGRAM INDUSTRIAL PROFILE: PETROLEUM AND NATURAL GAS SYSTEMS, 2018). O maior segmento da indústria de Petróleo e GN em termos de emissões de GHG relatadas pelo GHGRP é oriunda da produção *onshore*, com 117 MMT  $\text{CO}_2$ . Considerando as emissões das atividades de Sistemas de Petróleo e GN, cerca de um terço das emissões de  $\text{CO}_2$  vêm das extrações de gás natural.

Uma parte do  $\text{CH}_4$  presente na atmosfera é antropogênico, ou seja, são derivados de atividades humanas (SAPART et al., 2012). A Figura 1.1 ilustra os processos industriais

envolvendo GN que englobam a produção, processamento, transmissão e distribuição deste combustível, processos estes que favorecem a emissão de  $\text{CH}_4$ .

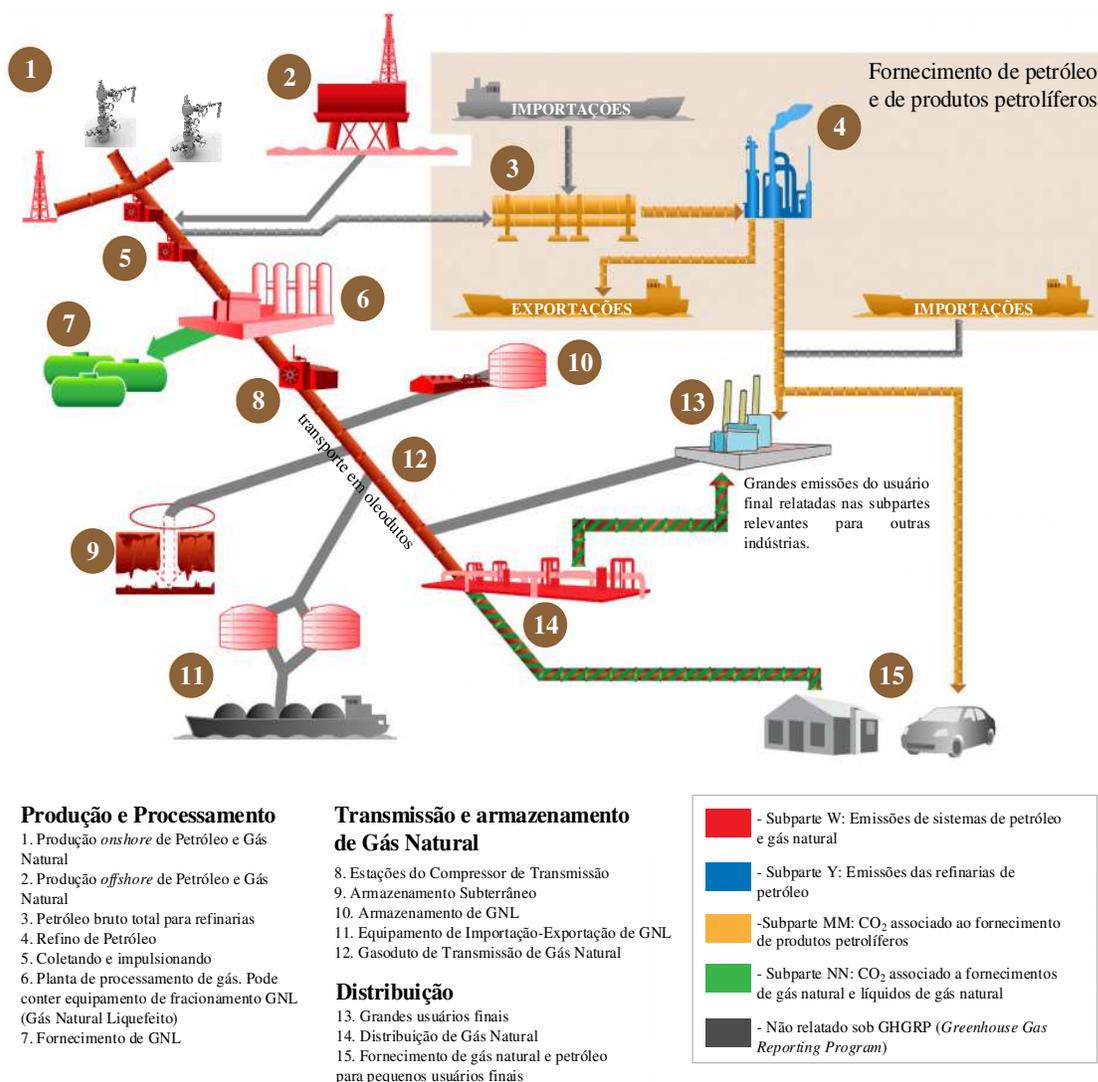


Figura 1.1 - Operações de petróleo e gás natural: oportunidades de mitigação de metano. Fonte: Adaptada e traduzida de relatório da *Environmental Protection Agency* (2011-2019 GREENHOUSE GAS REPORTING PROGRAM INDUSTRIAL PROFILE: PETROLEUM AND NATURAL GAS SYSTEMS, 2018).

As fontes de emissão de metano para a atmosfera são relacionadas a seguir:

- Procedimentos tais como completações de poços, purgas, dispositivos pneumáticos movidos a gás, *gas venting* (descarga de GN não queimados na atmosfera) para manutenção ou reparo de oleoduto e compressor (FULTON et al., 2011; HAUGLAND, 2019);

- ii) Vazamentos em estações de medição e regulação, dutos, estação de compressão, planta de processamento e em cabeças de poços petrolíferos (FULTON et al., 2011; HAUGLAND, 2019).

Portanto, para mitigar os efeitos do vazamento de GN sobre o efeito estufa, é necessário o desenvolvimento de estratégias de prevenção eficientes, como as ferramentas de detecção de vazamento de GN.

Além dos impactos ambientais, devido suas contribuições para as mudanças climáticas globais, os efeitos dos acidentes decorrentes de vazamento de gás natural têm consequências catastróficas, principalmente quando são seguidos de uma explosão, causando perdas financeiras e de vidas, capazes de destruir instalações de produção e processamento de óleo e gás (KHALIL, 2017).

No desastre da plataforma Piper Alpha em 1988, ocorrido depois que um condensado de GN vazou para a atmosfera e explodiu, 167 trabalhadores perderam suas vidas e ocorreu a perda total da plataforma (SOVACOOOL, 2008).

Para mitigar eventos que envolvem vazamentos e explosões de metano, existem técnicas que vão desde a classificação de área até o monitoramento e detecção de vazamentos desse gás em plantas de processamento industrial.

De rochas subterrâneas, as cabeças dos poços extraem o óleo, uma mistura gasosa de hidrocarbonetos e outros não hidrocarbonetos sob alta pressão. O vazamento de gás sob alta pressão produzirá uma nuvem de vapores visíveis que emanam do local do vazamento na cabeça do poço. Esta nuvem é o resultado de hidrocarbonetos e vapores de água que condensam a umidade do ar circundante (ATKINSON; HALL; MCGILLIVRAY, 2017; BOARD; INVESTIGATION, 2006).

Os primeiros trabalhos sobre sistemas inteligentes e autônomos de detecção de vazamento de GN, por meio da detecção de nuvens visíveis de vapores de hidrocarbonetos, com base na análise de imagens obtidas por meio de um sistema de vigilância, Circuito Fechado de Televisão (CFTV), e sem a necessidade de instalação de sensores específicos ou câmeras especiais, foram publicados entre os anos de 2012 e 2014 por Melo, Costa Filho e Costa (2012, 2013; 2014). O método proposto foi baseado no uso do componente R do espaço de cores RGB e no conceito de filtro de novidade (KOHONEN; OJA, 1976).

Esta tese investigou um novo método de detecção de vazamento de GN usando imagens de CFTV. Esse estudo é motivado pelo excelente desempenho das Redes Neurais Convolucionais (RNC), em inglês CNN de *Convolutional Neural Networks*, em tarefas de classificação de imagens (AGGARWAL, 2018; GOODFELLOW; BENGIO; COURVILLE,

2016). Por outro lado, registra-se a inexistência de estudos prévios que apliquem a CNN à predição de vazamento de GN utilizando imagens de CFTV.

Nesse contexto, este trabalho propõe um novo modelo para detecção de vazamento de gás natural baseado em redes neurais convolucionais (AGGARWAL, 2018) aplicado ao banco de dados utilizado por Melo, Costa Filho e Costa (2014), bem como a um novo banco de dados constituído de 145 imagens de 113 diferentes cabeça do poço com a presença de nuvens, afim de avaliar o poder de generalização do modelo proposto, bem como verificar se o modelo é capaz de distinguir as nuvens de vazamento de gás natural das nuvens do céu (fundo).

### 1.1 Objetivo geral

Propor um novo método de detecção de vazamento de gás natural em cabeça de poço petrolíferos *onshore* utilizando redes neurais convolucionais em imagens de Circuito Fechado de Televisão (CFTV). Propondo assim um novo método de predição de vazamento com desempenho superior as técnicas convencionais desenvolvidas por Melo, Costa Filho e Costa (2014).

### 1.2 Objetivos específicos

A seguir são enumerados os objetivos específicos desta tese:

- i) Propor e avaliar o desempenho de três diferentes arquiteturas CNN e selecionar aquela que oferece o melhor desempenho na detecção de vazamentos de gás natural;
- ii) Avaliar o desempenho das redes pré-treinadas DenseNet-201 (HUANG et al., 2017), GoogLeNet (SZEGEDY et al., 2015), MobileNetV2 (SANDLER et al., 2018) e ResNet-18 (HE et al., 2016), quanto a detecção de vazamentos de gás natural;
- iii) Propor e avaliar o desempenho de um modelo CNN baseado em fusão de características (CNN FF), obtido por meio da extração simultânea de características das redes DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18, quanto a detecção de vazamentos de gás natural;
- iv) Comparar o desempenho das CNN MMD, DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18 e FF, com respeito, principalmente, a acurácia e a taxa de falso negativos fornecidas por essas CNNs, e assim definir o Melhor Modelo Geral (CNN MMG), quanto a detecção de vazamentos de gás natural;
- v) Comparar o desempenho do modelo baseado em CNN com o do modelo baseado em filtro de novidade (MELO; COSTA FILHO; COSTA, 2014; MELO; COSTA; COSTA FILHO, 2014), através do teste Chi-quadrado (HOWELL, 2010);

- vi) Investigar a capacidade do melhor modelo geral em distinguir nuvens do céu de plumas de GN através da avaliação de imagens de diferentes cabeças de poço, obtidas em diversas condições ambientais, desconhecidas do modelo.

### 1.3 Organização da tese

Esta seção traz um esboço relativo à estruturação desta tese, cuja finalidade é passar ao leitor uma breve descrição do que compreende cada capítulo da tese.

No Capítulo 1 (Introdução) apresenta-se uma visão geral da pesquisa, incluindo a origem e a importância do gás natural para a sociedade e os impactos decorrentes de um vazamento de gás natural. Neste capítulo, descreve-se ainda a relevância e os objetivos do trabalho.

No Capítulo 2 são discutidos o estado da arte sobre as principais técnicas de detecção de vazamentos de GN.

No Capítulo 3 é apresentada a fundamentação teórica que respalda a pesquisa, com especial destaque para os conceitos referentes às CNN.

No Capítulo 4 descreve-se os materiais e métodos utilizados no desenvolvimento da máquina de reconhecimento de padrões baseada em CNN, cuja tarefa primordial é a detecção de vazamento de GN em uma imagem.

No Capítulo 5 descreve-se os resultados obtidos para máquina de reconhecimento de padrões.

Finalmente, no Capítulo 6 apresenta-se uma discussão sobre os resultados mais importantes obtidos, faz-se uma comparação com outros trabalhos previamente publicados na literatura, destaca-se as principais contribuições do trabalho, indica-se as limitações do método de reconhecimento de padrões proposto para a detecção de vazamento de GN e sugere-se recomendações para trabalhos futuros.

O Apêndice 1 inclui o primeiro artigo científico, decorrente dos resultados desta tese, publicado no periódico IEEE Access, ISSN 2169-3536, conforme dados abaixo.

MELO, Roberlanio De Oliveira; COSTA, Marly Guimarães Fernandes; COSTA FILHO, Cícero Ferreira Fernandes. Applying Convolutional Neural Networks to Detect Natural Gas Leaks in Wellhead Images. **IEEE Access**, [s. l.], v. 8, p. 191775–191784, 2020. doi: 10.1109/ACCESS.2020.3031683.

O Apêndice 2 inclui o segundo artigo científico, decorrente dos resultados desta tese, submetido para publicação no periódico Computers & Electrical Engineering, ISSN: 0045-7906. <https://www.sciencedirect.com/journal/computers-and-electrical-engineering/special-issues>.



## 2 REVISÃO BIBLIOGRÁFICA

Neste capítulo procura-se contextualizar o problema a ser resolvido nessa tese, qual seja a detecção de vazamentos de GN. Com esse intuito, são apresentadas informações a respeito da natureza e extração do petróleo e do gás natural, e as técnicas mais relevantes já desenvolvidas na detecção de vazamento de gás natural. A apresentação de tais técnicas tem por objetivo permitir uma comparação, no Capítulo 6, relativa à discussão e conclusões, dos resultados obtidos com o método proposto nessa tese para detecção de vazamentos de GN, com resultados de trabalhos previamente publicados na literatura.

### 2.1 A importância econômica do petróleo e gás natural

Desde o século 19, quando Edwin Laurentine Drake, conhecido como Coronel Drake, deu início ao processo de exploração comercial do petróleo, ao perfurar o primeiro poço petrolífero comercial e, posteriormente, o gás natural ser comercializados em larga escala, novas fontes de energia surgiram e são pesquisadas até os dias atuais (THOMAS, 2001). No entanto, em 2020, conforme dados mais recentes encontrados na publicação *Statistical Review of World Energy* (BP, 2021), e apresentados na Figura 2.1, juntos, o petróleo e o gás natural respondem por 55,93% do consumo mundial de energia. Portanto, esses dois combustíveis fósseis são a chave da economia mundial, propiciando a contínua expansão da sociedade moderna, pois são eles as principais fontes de energia desta sociedade.

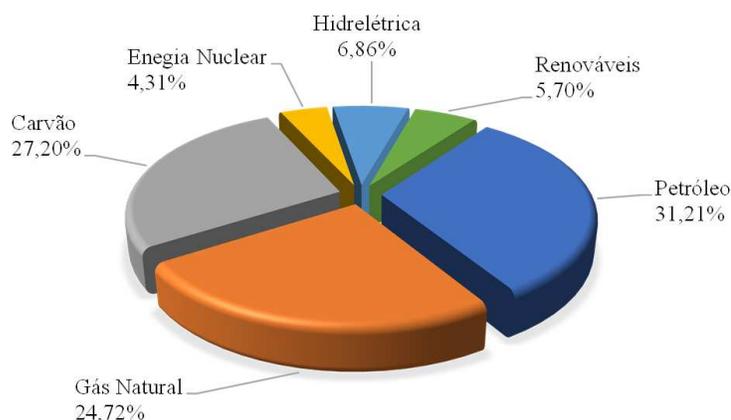


Figura 2.1 - Consumo mundial de energia por tipo de combustível no ano de 2020. Fonte: Adaptada de *Statistical Review of World Energy* (MURVAY; SILEA, 2012; ZAMAN et al., 2020).

## 2.2 O petróleo

Encontrado na natureza, especificamente no subsolo do planeta sob sedimentos, o petróleo encontra-se aprisionado há milhões de anos, em rochas permeáveis e porosas, rodeadas por materiais impermeáveis. A explicação para sua origem pode ser creditada à teoria abiogênica ou à teoria biogênica (FARAMAWY; ZAKI; SAKR, 2016).

O professor Thomas Gold, utilizando como base a teoria russo-ucraniana, desenvolveu uma das vertentes da teoria abiogênica (GOLD, 1985). Como astrofísico, Gold possuía conhecimento que o carbono é o quarto elemento mais abundante no universo, apresentando-se principalmente na forma de hidrocarbonetos, e em especial na molécula de metano, a qual compreende grande parte da atmosfera dos planetas, tais como Jupiter, Saturno, Urano e Netuno. Logo, para Gold, o petróleo e o gás natural, constituídos de metano, estão presentes no planeta Terra desde a sua formação, não sendo necessário artifícios especiais da natureza para a constituição do metano em nosso planeta (GLASBY, 2006).

A teoria orgânica também sugere que, no planeta Terra, os hidrocarbonetos originam-se de forma natural, através da sinterização de diversos tipos de matéria orgânica (vegetais, algas, plâncton, animais, etc.) armazenada sob sedimentos a milhões de anos atrás, sendo submetida a uma série de processos termoquímicos (combinação de ações bacteriológicas e mudanças de temperatura) conhecidos como estágios de diagênese, catagênese e metagênese (SPEIGHT, 2006).

O processo exploratório contínuo do homem em busca de tais reservas fornece evidência que ainda há grandes acúmulos de petróleo a serem descobertos no planeta terra, como ocorreu recentemente no Brasil, onde grandes reservatórios de petróleo foram localizados na costa brasileira em águas ultra profundas sob uma camada de sal no leito do oceano, denominado como região do pré-sal (ARAUJO RODRIGUES; LUÍS SAUER, 2015).

Do ponto de vista químico, o petróleo é composto principalmente por hidrocarbonetos, ou seja, de compostos orgânicos formados por ligações químicas entre átomos de hidrogênio (H) e carbono (C). Inúmeros outros elementos químicos também compõem esse combustível fóssil, mas os principais átomos e sua presença no petróleo possuem análise elementar dentro dos limites indicados na Tabela 2.1.

Tabela 2.1 - Análise elementar dos constituintes do petróleo.

Elemento químico	Composição (% em peso)
Carbono	84 – 87
Hidrogênio	11 – 14
Enxofre	0,06 – 2,00
Nitrogênio	0,1 – 2,0
Oxigênio	0,6 – 2,0

Fonte: Adaptada de McCain (MCCAIN, 2017)

### 2.3 Gás natural

Entre os combustíveis fósseis, o gás natural detém a maior eficiência energética e é considerado mais limpo e ecológico quando comparado com outros combustíveis fósseis (carvão e petróleo bruto, por exemplo), pois ele tem uma queima relativamente limpa (FARAMAWY; ZAKI; SAKR, 2016; MOKHATAB; POE; MAK, 2018). A combustão do GN resulta em emissão insignificante de dióxido de enxofre e as emissões reduzidas de óxido nítrico e o dióxido de carbono, em relação à queima dos demais combustíveis fósseis (FARAMAWY; ZAKI; SAKR, 2016). Porém, o hidrocarboneto metano presente no GN, ao ser enviado para atmosfera sem queimar é danoso ao meio ambiente, conforme descrito no Capítulo 1 (INTRODUÇÃO).

O termo gás natural refere-se a gás combustível rico composto de hidrocarbonetos, sendo considerado inerente aos campos petrolíferos, no entanto também pode ser encontrado em minas de carvão e xisto (SPEIGHT, 2018).

Nos reservatórios, o petróleo pode estar na forma sólida, líquida ou gasosa. O GN é a parcela do petróleo encontrado no estado gasoso (não associado) ou em solução no petróleo (associado). O GN associado está presente na rocha do reservatório do campo petrolífero, na forma de gás dissolvido no petróleo ou como gás situado acima do reservatório de óleo (gás de revestimento). O GN não associado está contido em uma formação geológica que normalmente contém pouco ou nenhum petróleo no estado líquido, e por isso muitas vezes a literatura o descreve como gás seco (o gás associado também descrito como gás úmido). Em ambos os tipos, caso ocorra um vazamento no *wellhead*, o mesmo irá para atmosfera em forma de nuvem, devido a mudança brusca de pressão e temperatura e, principalmente, devido aos vapores de água que condensam a umidade do ar circundante (ATKINSON; HALL; MCGILLIVRAY, 2017; BOARD; INVESTIGATION, 2006).

Os recursos de gás natural diferem pelas características geológicas de sua rocha reservatório, conforme ilustrado na Figura 2.2.

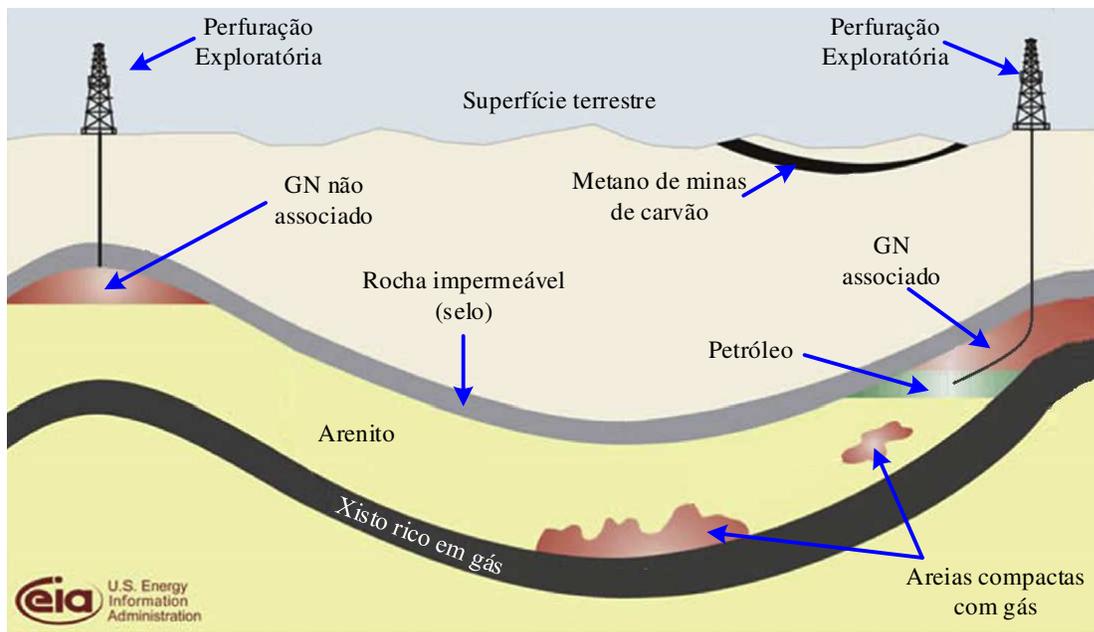


Figura 2.2 - Geologia esquemática dos recursos de gás natural. Fonte: Adaptada e traduzida de Mokhatab, Poe e Mak (2018).

O GN é formado por moléculas de hidrocarbonetos alifáticos saturados e, para fins de compreensão da sua composição, os dados da Tabela 2.2 mostram os tipos de hidrocarbonetos e suas proporções típicas em GN associado e não associado. Observa-se que o metano é o hidrocarboneto mais presente no GN (MCCAIN, 2017).

#### 2.4 Confiabilidade e manutenibilidade de sistemas para prevenção de vazamentos de gás natural

A cadeia de produção, transporte e distribuição de GN mostrada na Figura 1.1 é altamente eficiente e possui alto valor social e comercial. Logo, é mandatório que essa cadeia possua uma gestão de sua integridade através de prevenção e detecção de vazamento de GN, pois o vazamento é um dos piores eventos que pode ocorrer em instalações petrolíferas.

No entanto, por mais eficiente que seja a gestão, não significa que o processo de produção do GN seja imune a falhas. Falhas em sistemas envolvendo vazamento de GN são inevitáveis, sendo inúmeras as causas que podem propiciá-las, seja devido a corrosão, problemas no processo fabril, elevação de pressão e temperatura, expansão e afrouxamento de juntas, desastres naturais, entre outros (PSYRRAS; SEXTOS, 2018). Durante a operação destes

sistemas, uma das práticas para mitigar possíveis desastres é a detecção de vazamento de GN. Para este fim, foram desenvolvidas diversas técnicas.

Tabela 2.2 - Composição típica do gás natural associado e não associado.

Tipo de Componente	Componente	Fórmula molecular	Gás não associado	Gás associado
			(% em peso)	(% em peso)
Hidrocarbonetos	Metano	CH <sub>4</sub>	70 – 98	45 – 92
	Etano	C <sub>2</sub> H <sub>6</sub>	1 – 10	4 – 21
	Propano	C <sub>3</sub> H <sub>8</sub>	traços – 5	1 – 15
	Butano	C <sub>4</sub> H <sub>10</sub>	traços – 1	0,5 – 7
	Pentano	C <sub>5</sub> H <sub>12</sub>	traços – 0,5	traços – 3
	Hexano	C <sub>6</sub> H <sub>14</sub>	traços – 1	traços – 2
	Heptano	C <sub>7</sub> H <sub>16</sub>	0,0 – 0,5	0,0 – 1,5
Não Hidrocarboneto	Nitrogênio	N <sub>2</sub>	traços – 15	traços – 10
	Dióxido de Carbono	CO <sub>2</sub>	traços – 5	traços – 4
	Sulfeto de hidrogênio	H <sub>2</sub> S	traços – 3	0,0 – 6
	Hélio	He	traços – 5	0

Fonte: Adaptada de McCain (MCCAIN, 2017).

O vazamento de GN é um evento com anatomia e assinatura de propriedades hidráulicas, que requer a adoção de estratégias eficazes de gerenciamento de vazamentos visando minimizar a probabilidade de ocorrências e consequências das mesmas, antevendo possíveis vazamentos e detectando vazamentos existentes. A Figura 2.3 ilustra algumas estratégias de prevenção e detecção de vazamentos de GN (ZAMAN et al., 2020).

O principal enfoque desta tese é a detecção do vazamento de GN, destacado através de um sombreamento na Figura 2.3.

Quanto ao vazamento de GN, na literatura existem vários métodos relacionados à detecção dele. Diversos autores adotam termos distintos para definirem as diferentes metodologias propostas (ADEGBOYE; FUNG; KARNIK, 2019; MONIEN; PREIS; SCHAMBERGER, 2007). Os métodos de monitoramento biológico envolvem o emprego de pessoas e cães (CARLSON, 1993; SCOTT; BARRUFET, 2003). No passado, pássaros eram usados em minas de carvão para detectar vazamentos de gases letais, incluindo CH<sub>4</sub> (RABINOWITZ; SCOTCH; CONTI, 2009). A utilização de humanos, como sistema de prevenção de vazamento de gás, foi descrito por Baroudi, Al-Roubaiey e Devendiran (2019)

2019). Nesse sistema, operadores analisam manualmente os arquivos de vídeo, quadro a quadro, em busca de falhas no duto.

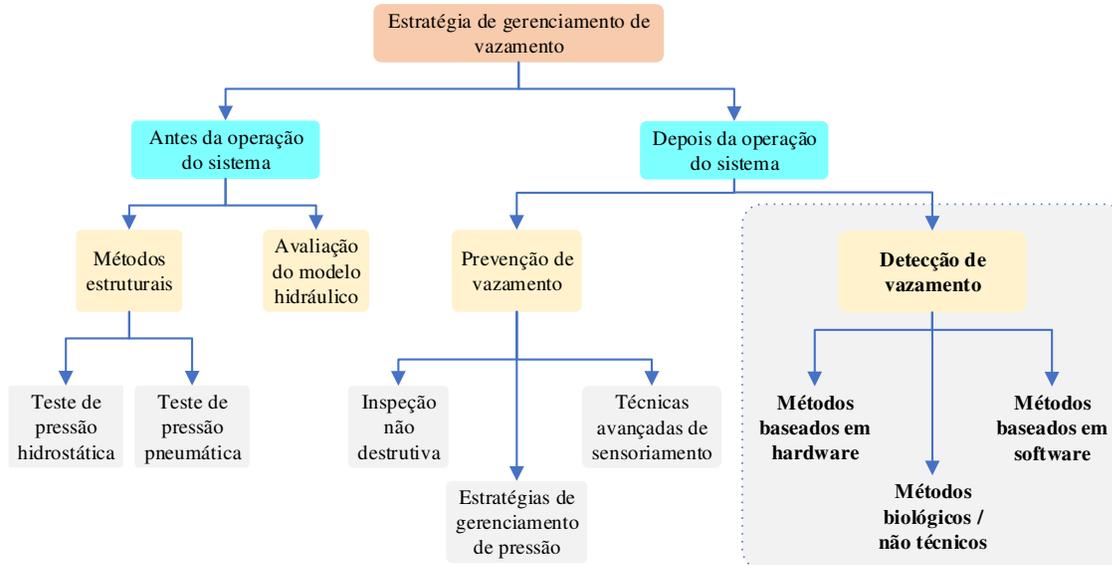


Figura 2.3 - Estratégias de gerenciamento de vazamentos hidráulicos. Fonte: Adaptada e traduzida de Murvay e Silea (2012) e Zaman *et al.* (2020).

Os métodos de detecção baseados em hardware requerem sensores especializados para esse propósito, como sensores acústicos, ópticos, eletroquímicos, ultrassônicos e imageamentos (ADEGBOYE; FUNG; KARNIK, 2019; MURVAY; SILEA, 2012; SACHEDINA; MOHANY, 2018). Os métodos baseados em software, utilizando algoritmos computacionais inteligentes (ADEGBOYE; FUNG; KARNIK, 2019), analisam e classificam os sinais de um sensor ou de uma rede de sensores que monitoram diversas variáveis, como vazão, pressão e temperatura (SACHEDINA; MOHANY, 2018), facilitando assim a detecção de vazamentos, por balanço de massa positivo ou negativo e por anomalias na pressão do sistema de processamento de gás natural (MURVAY; SILEA, 2012).

Estudos sugerem ainda a detecção de vazamentos de gases combustíveis por meio do processamento de imagens, porém, o banco de dados desses trabalhos é formado por imagens térmicas adquiridas por hardware de alto custo como, por exemplo, câmeras infravermelhas (IR, do inglês, *infrared*) específicas para detectar o espectro IR da molécula de metano (JADIN; GHAZALI, 2014; LIU *et al.*, 2018; SHI *et al.*, 2020a; WANG *et al.*, 2020).

As CNN, aplicadas em diversos domínios, com excelentes resultados na resolução de problemas envolvendo imagens e dados gerais, como classificação de nuvens através da análise de imagens de satélite (WOHLFARTH *et al.*, 2018), reconhecimento de gestos manuais

(MOLCHANOV et al., 2015), previsão de GN quanto a sua demanda e oferta (CHEN; XU; KOCH, 2020) e radiação solar (GHIMIRE et al., 2019), avaliação da estabilidade transitória do sistema de potência (SHI et al., 2020b), diagnóstico de doenças em plantas (DARWISH; EZZAT; HASSANIEN, 2020), classificação de nódulos pulmonares (XIE et al., 2019), detecção de intrusão de rede em veículos (SONG; WOO; KIM, 2020), detecção de bifurcação vascular (MIYAGAWA et al., 2019), detecção de incêndio e fumaça (BU; GHARAJEH, 2019), e muitos outros domínios científicos. A Tabela 2.3 resume estes domínios, nos quais os autores utilizaram as CNNs como máquinas de reconhecimento de padrões.

Tabela 2.3 - Algumas tarefas que utilizaram as CNNs que como máquinas de reconhecimento de padrões.

Tarefa	Título	Autores	Ano
Classificação de nuvens através da análise de imagens de satélite	<i>Dense cloud classification on multispectral satellite imagery</i>	WOHLFARTH, Kay. et al.	2018
Detecção de bifurcação vascular	<i>Detecting Vascular Bifurcation in IVOCT Images Using Convolutional Neural Networks With Transfer Learning</i>	MIYAGAWA, M. et al.	2019
Detecção de incêndio e fumaça	<i>Intelligent and vision-based fire detection systems: A survey</i>	BU, F.; GHARAJEH, M. S.	2019
Classificação de nódulos pulmonares	<i>Knowledge-based Collaborative Deep Learning for Benign-Malignant Lung Nodule Classification on Chest CT.</i>	XIE, Y. et al.	2019
Previsão de GN quanto a sua demanda e oferta	<i>Day-ahead high-resolution forecasting of natural gas demand and supply in Germany with a hybrid model</i>	CHEN, Y.; XU, X.; KOCH, T.	2020
Radiação solar	<i>Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms</i>	GHIMIRE, S. et al.	2020
Avaliação da estabilidade transitória do sistema de potência	<i>Convolutional neural network-based power system transient stability assessment and instability mode prediction</i>	SHI, Z. et al.	2020b
Diagnóstico de doenças em plantas	<i>An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis</i>	DARWISH, A.; EZZAT, D.; HASSANIEN, A. E.	2020

Assim, não chega a ser uma surpresa que estas redes também tenham sido aplicadas na predição de vazamentos de GN, com banco de dados de imagens térmicas

Wang *et al.* (2020), abordou os conceitos de CNN para tarefa de predição de vazamentos de hidrocarbonetos analisando imagens digitais. No entanto, o conjunto de dados daquele estudo (imagens) foram obtidas por câmeras infravermelhas especialmente construídas para visualizar moléculas de hidrocarbonetos.

## 2.5 Técnicas contemporâneas de detecção de vazamento de GN

A literatura disponibiliza um vasto volume de técnicas/métodos de detecção de vazamentos de GN. A Figura 2.4 ilustra parte destes métodos.

Os métodos baseados em *hardware* usam sensores (*hardwares*) para detectar diretamente a ocorrência de um vazamento e, em alguns casos, auxiliar na localização. Alguns dos *hardwares* usados nessa tarefa incluem detectores acústicos, fibra óptica, tecnologias ultrassônicas e termógrafo infravermelho (ADEGBOYE; FUNG; KARNIK, 2019; MURVAY; SILEA, 2012; ZAMAN et al., 2020).

Os métodos baseados em software usam inteligência computacional para monitorar constantemente dados das variáveis do processo como pressão, temperatura e taxa de fluxo, no intuito de detectar vazamentos no sistema. Entre as análises utilizadas cita-se: a detecção de mudança de fluxo e pressão, equilíbrio de volume de massa, sistemas baseados em modelo dinâmico e outros (ADEGBOYE; FUNG; KARNIK, 2019; MURVAY; SILEA, 2012; ZAMAN et al., 2020).

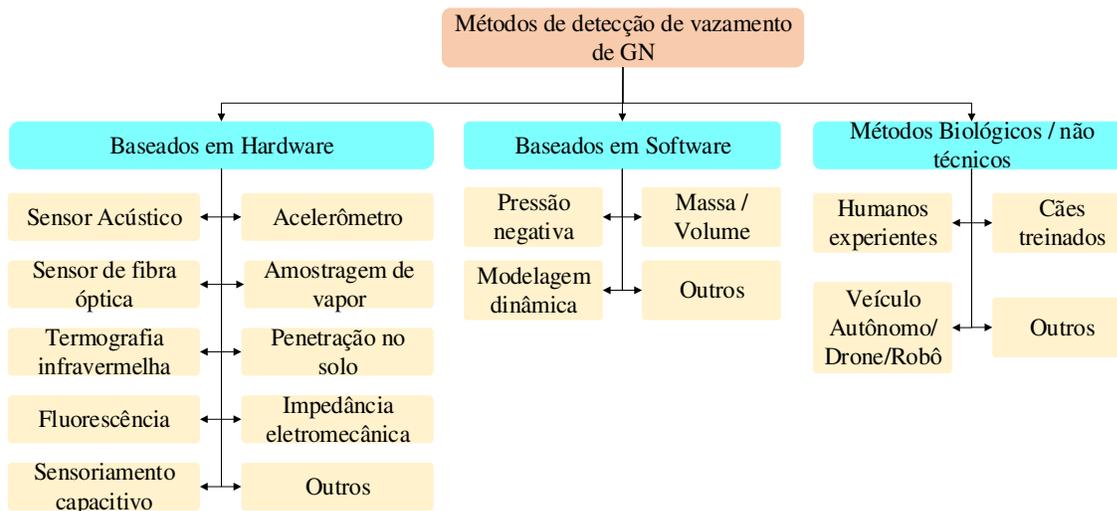


Figura 2.4 - Estratégias de gerenciamento de vazamento hidráulicos. Fonte: Adaptada e traduzida de Murvay e Silea (2012), Adegboye, Fung e Karnik (2019) e Zaman *et al.* (ZAMAN et al., 2020).

Já os métodos visuais/biológicos usam humanos ou cães treinados para detectar e localizar um vazamento por meio de inspeção visual, odor ou som. Estão incluídos nesse grupo a detecção remota por inspeção através de drones ou veículos autônomos.

A seguir detalha-se alguns desses métodos utilizados na detecção de vazamento de GN.

### 2.5.1 Cães treinados

A adoção de cães rastreadores treinados é uma solução não contínua para monitoramento de vazamentos de GN. Estes animais são treinados para reconhecer determinados odores de compostos que são injetados no sistema a ser inspecionado. Após injetar o composto no sistema, o cão então é conduzido ao longo do um oleoduto, farejando o odor específico (ADEGBOYE; FUNG; KARNIK, 2019).

### 2.5.2 Humanos experientes

A capacidade sensorial dos seres humano é usada para detectar vazamento de GN. Inspetores experientes, utilizando sentidos de visão, audição e até olfato, inspecionam extensões de dutos buscando evidências de vazamentos.

### 2.5.3 Veículo Autônomo/Drone/Robô

As inspeções em sistemas e dutos também podem ser realizadas por veículos autônomos, drones e robôs, buscando prever e detectar rachaduras, vazamentos e outras deformações. Estas inspeções podem ser visuais, através de câmeras, na busca de corrosões. Os robôs são equipados com tecnologia moderna, como Raios-X e Ultrassom, ou até câmera térmicas, por exemplo (IBRAHIMOV, 2018).

### 2.5.4 Monitoramento acústico

Como o nome sugere, o monitoramento acústico é baseado nas emissões acústicas nos pontos de vazamento. Os gases, ao se moverem de um ambiente de alta pressão para um outro ambiente de baixa pressão, através de um orifício de vazamento, criam um fluxo turbulento que produz um sinal acústico ultrassônico. O método ultrassônico consiste em detectar estas ondas de pressão sonora (ultrassônicas) através de elementos sensores piezoelétricos, indicando a ocorrência do vazamento. Este método direcional, em que o evento de detecção ocorre imediatamente após seu início, apresenta alta sensibilidade. No entanto, é uma tecnologia sensível a ruído de fundo e incapaz de distinguir qual composto está vazando, por exemplo, se é GN, ar de instrumento ou vapor (KOROTCENKOV, 2013; LI et al., 2019).

Sendo o vazamento descrito como fluxo de gás ou líquido através da parede da tubulação que apresenta a falha, como furo, rachadura ou rupturas, a passagem do gás pelo orifício causa uma diferença de pressão na tubulação, com o fluxo dirigindo-se da região de pressão mais alta para a região de pressão mais baixa, o que gera ondas sonoras. De acordo com Li *et al.* (2019), o nível do ruído do vazamento é definido pela equação (1)

$$L = 80 + 20 \lg \frac{D}{D_0} + 10 \lg \frac{(P - P_0)^4}{P_0^2 (P - 0.5P_0)^2}, \quad (1)$$

em que:

$L$  é o nível de pressão sonora a uma distância de 1 m;

$P$  é a pressão interna ( $kPa$ );

$P_0$  é a pressão absoluta atmosférica ambiente ( $kPa$ );

$D$  é o diâmetro do bico (mm);

$D_0$  é o diâmetro de referência (1 mm).

Assim, é possível, além de detectar o vazamento, localizá-lo, utilizando dois sensores acústicos, através do intervalo de tempo entre os sinais acústicos detectados. A velocidade da onda sonora, dependerá do material e do diâmetro do tubo, e pode ser determinada através da fórmula da velocidade de propagação das ondas de vazamento de fluido monofásico, equação (2) (HOUGH; E, 1988).

$$a = \sqrt{\frac{1}{\rho \left[ \frac{1}{K} + \frac{D}{E} \phi \right]}}, \quad (2)$$

em que:

$a$  = velocidade de propagação da onda ( $m/s$ );

$\rho$  = densidade do fluido ( $kg / m^3$ );

$E$  = módulo de elasticidade de Young do material da parede do tubo ( $N / m^2$ );

$K$  a massa bruta do líquido ( $N / m^2$ );

$\phi$  um fator de restrição dependente do coeficiente de Poisson do material da parede.

O método acústico para detecção de vazamentos pode ser dividido em duas classes: ativo e passivo. Os métodos ativos detectam o som produzido pelo ecos do vazamento, enquanto que os métodos passivos detectam o vazamento auscultando as ondas sonoras no duto (ADEGBOYE; FUNG; KARNIK, 2019).

### 2.5.5 Método baseado em combustão catalítica

Os sensores (Figura 2.5) catalíticos são do tipo termométricos (calorimétricos), pois convertem as alterações de temperatura, geradas por reações químicas, em sinais elétricos. A operação deles consiste em injetar uma corrente elétrica em uma bobina para aquecer a camada de catalisador. O catalisador quente e na presença de um gás inflamável, permite que a oxidação

ocorra em uma reação química semelhante à combustão. A liberação de calor devido a reação catalítica altera a temperatura do catalisador, juntamente com a da bobina, sendo a resistência elétrica da bobina alterada. Essa alteração na resistência da bobina é detectada por uma ponte *Wheatstone*. O método catalítico é versátil, pois detecta uma gama de gases e suas concentrações na atmosfera, sendo de fácil fabricação e custo relativamente barato. No entanto, a vida útil do sensor é de aproximadamente dois anos e podem ficar saturados na presença de grandes concentrações de gases ou ter sua sensibilidade reduzida quando na presença de gases contendo chumbo, óleos, enxofre ou cloro (BASU, 2017; KOROTCENKOV, 2013).

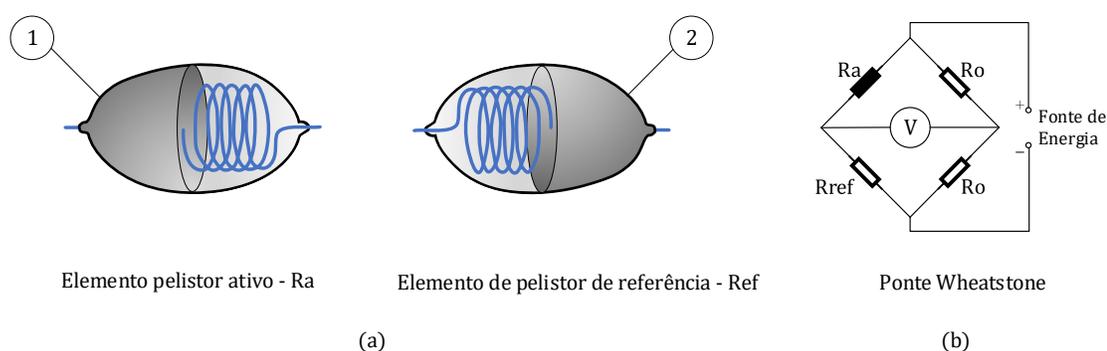


Figura 2.5 - Sensor baseado em combustão catalítica: (a) elementos catalisadores ativo (1) e de referência (2) e (b) Ponte *Wheatstone* com elemento catalítico. Fonte: Adaptada e traduzida de Wang *et al.*(2011).

### 2.5.6 Método baseado em absorção do espectro infravermelho

A detecção de gás por infravermelho (IR) é baseada na capacidade dos gases hidrocarbonetos absorverem a luz em comprimentos de ondas específicos. Esta técnica consiste em medir a absorção de uma região do espectro infravermelho por uma amostra de gás posicionada no caminho de um feixe IR. Em temperaturas acima do zero absoluto, todos os átomos de uma determinada molécula estão em vibração contínua uns em relação aos outros. Quando a frequência de uma vibração específica é igual à frequência da radiação IR direcionada sobre a molécula, a molécula absorve a radiação IR (HSU, 1998). No caso do GN, composto por 70 % da molécula de metano, o sensor é desenvolvido para detectar absorção de onda média de IR no espectro, cujo comprimento de onda é próximo a 3,3  $\mu\text{m}$ .

A Figura 2.6 ilustra o esquema de sensor infravermelho utilizado para detecção de gás combustível.

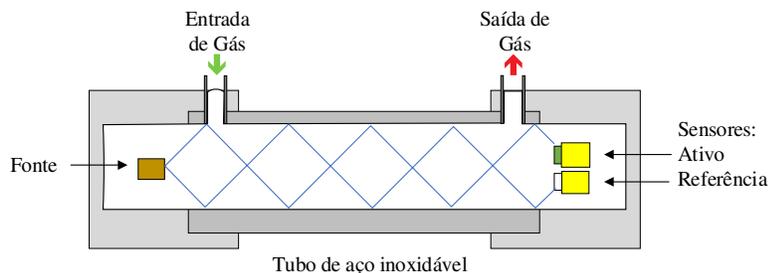


Figura 2.6 - Estratégias de gerenciamento de vazamento hidráulicos. Fonte: Adaptada e traduzida de Murvay e Silea (2012), Adegboye, Fung e Karnik (2019) e Zaman *et al.* (ZAMAN *et al.*, 2020).

### 2.5.7 Balanço de massa / volume

Essa é uma técnica baseada no princípio da conservação de massa, ou seja, o método determina um vazamento quando ocorreu um desequilíbrio na massa ou no volume do gás de entrada e saída, uma vez que no sistema fechado integro, o volume de gás que sai de uma seção do gasoduto é igual ao volume de gás que entra nessa seção. O cálculo dessa diferença é realizado através da análise de variáveis do processo tais como taxa de fluxo, pressão e temperatura (ADEGBOYE; FUNG; KARNIK, 2019; MURVAY; SILEA, 2012).

### 2.5.8 Pressão negativa

Quando ocorre um vazamento no duto, a pressão e velocidade do líquido interno que escoam pelo duto são afetados, resultando na queda instantânea de pressão e redução da velocidade do fluido. Essa mudança brusca de pressão gera uma pressão negativa tanto a montante como a jusante das extremidades do duto. Logo, é possível prever a existência de um vazamento na tubulação (ADEGBOYE; FUNG; KARNIK, 2019; MURVAY; SILEA, 2012).

## 2.6 Técnicas de detecção de vazamento de GN por imagem

O mercado industrial oferece diversos tipos de “câmeras de imagem térmica” que permitem a detecção de plumas de gases combustíveis por imagem, sendo que o fabricante FLIR™ está na vanguarda desse mercado. No entanto, existem outros fabricantes, como OPGAL™, TELOPS™ e SIERRAOLYMPIC™.

O princípio de funcionamento dessas câmeras é baseado na emissão de radiação infravermelha em uma determinada faixa de espectro. Câmeras IR de ondas médias (MWIR, do inglês, *Mid Wave Infrared*), operando na faixa de espectro de 3  $\mu\text{m}$  a 5  $\mu\text{m}$ , com o elemento sensor de antimônio de índio (InSb), detectam gases como o metano, o principal componente do GN (FLIR, 2020; NIEDERER, 2012; XIAO *et al.*, 2020).

Wang *et al.* (2020), propôs a detecção de vazamentos de metano utilizando um conjunto de dados para treinamento, validação e teste composto por 118.060 imagens IR de  $240 \times 320$  pixels, obtidas por câmera MWIR, conforme mostrado na Figura 2.7. Esses autores construíram e aplicaram modelos de CNNs próprias para detectar a classe de vazamento, cuja acurácia geral de detecção foi próxima a 95%.

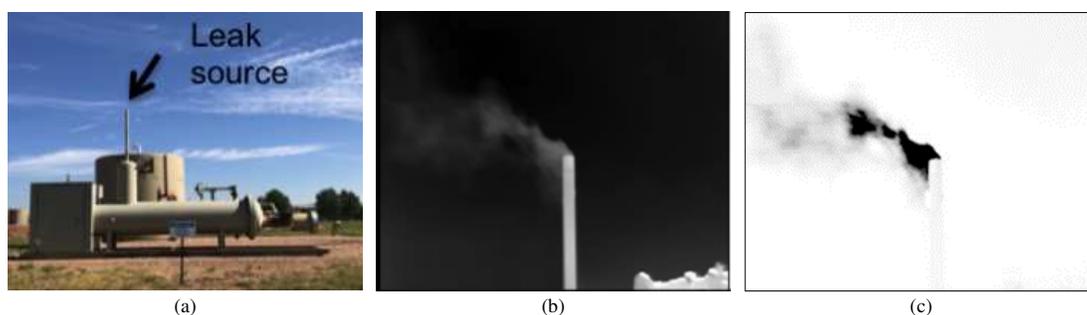


Figura 2.7 - Proposta de Wang *et al.* (2020) para predição de vazamento de GN através de imagens térmicas: (a) instalação do estudo, um separador multifásico; (b) imagem do vazamento de gás natural obtida pela câmera infravermelha e (c) subtração de fundo da figura (b). Fonte: Adaptada de Wang *et al.* (2020).

Shi *et al.* (2020a) também usaram a câmera infravermelha FLIR-320, o mesmo hardware usado por Wang *et al.* (2020), para detectar vazamentos de metano. Os autores utilizaram a técnica *Faster Region-based Convolutional Neural Networks* (FR CNN) para analisar os vídeos, alcançando um valor para a métrica intersecção sobre a união (IOU, do inglês *Intersection-Over-Union*) de 98%. O conjunto de dados utilizado pelos autores é constituído por 3.205 imagens, sendo 2.000 quadros para treinamento, 605 para validação e 600 para teste, com dimensão de  $320 \times 240$  pixels.

Em outro trabalho, os autores treinaram uma rede de segmentação semântica para classificar vazamentos de gás. O conjunto de dados utilizado pelos autores foram imagens obtidas também por uma câmera térmica do tipo MWIR. O valor obtido para a métrica IOU neste estudo foi de 72,6% (MARSHALL; PARK; SONG, 2019).

Em Costa Filho, Melo e Costa (2012), foi descrito o primeiro método de detecção de vazamento de GN utilizando imagens de *wellhead onshore*, baseado em um sistema de vigilância (CFTV). Este sistema não necessitou de instalação de sensores específicos ou câmeras com sensores de imageamento sensibilizados para moléculas de  $\text{CH}_4$ . O método proposto pelo autores utilizou o classificador filtro de novidade (KOHONEN; OJA, 1976), obteve melhores resultados quando a componente R do espaço de cores RGB das imagens do CFTV foi empregada. Obteve-se uma área sob a curva ROC (*Area Under Receiver Operating*

*Characteristic Curve - AUC*), de  $AUC = 98,52\%$ , com erro médio de  $0,37\%$  (MELO; COSTA FILHO; COSTA, 2014). De acordo com a Figura 2.8, os autores também propuseram a segmentação da nuvem de gás com base na novidade de cada pixel.

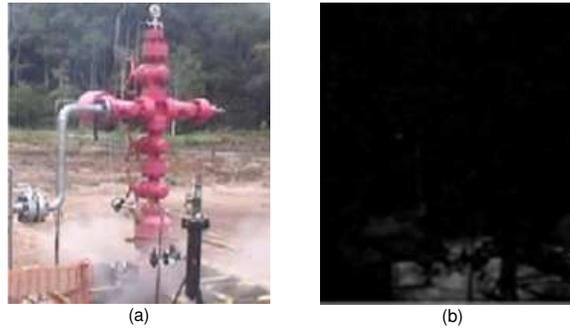


Figura 2.8 - Proposta de Melo et al. (2013) para predição de vazamento de GN através de imagens de CFTV: (a) imagem original com vazamento de gás próximo ao nível do solo e (b) imagem de novidade correspondente da Fig. 3 (a). Fonte: Costa Filho, Melo e Costa (2013).

O conceito de filtro de novidade foi descrito por Kohonen e Oja (1976), baseado no método clássico de ortogonalização de Gram-Schmidt. Seja  $\{x_1, x_2, \dots, x_m\}$  um conjunto de vetores  $n$  euclidianos dimensionais que abrangem um subespaço  $m$  dimensional  $L \subset R^n$ , com  $m < n$ . Considerando o subespaço  $L$ , um vetor arbitrário  $x \in R^n$  que pode ser dividido em duas componentes,  $\hat{x}$  e  $\tilde{x}$ . A Figura 2.9 ilustra essas duas componentes no espaço  $R^3$ , em que:  $\hat{x} \in L$  é a projeção de  $x$  em  $L$  e  $\tilde{x} \perp L$  é a projeção de  $x$  perpendicular a  $L$ . O vetor  $\hat{x} \in L$  é a componente de  $x$  que é "conhecida" pelo subespaço  $L$  e pode ser definida como uma combinação linear de  $\{x_1, x_2, \dots, x_m\}$ . Vector  $\tilde{x} \perp L$  é a informação "novidade", que é desconhecida pelo subespaço  $L$  e não pode ser representada como uma combinação linear de  $\{x_1, x_2, \dots, x_m\}$  (COSTA FILHO; MELO; COSTA, 2012).

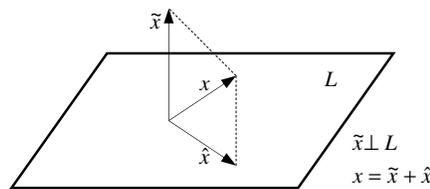


Figura 2.9 - Ilustração do conceito de filtro de novidade no espaço  $R^3$ . Conceito usando o método de ortogonalização de Gram-Schmidt. Fonte: Costa Filho, Melo e Costa(2012).

A Tabela 2.4 sumariza trabalhos da literatura, publicados em jornais (periódicos), que abordaram a detecção de vazamento de gás natural através de *deep learning* e o trabalho que utilizou filtro de novidade como máquina de reconhecimento de padrão (vazamento de gás

natural). A principal distinção entre estes trabalhos, além das técnicas para detecção de vazamento, é o tipo câmara utilizada para capturar imagens (câmara de espectro infravermelho e câmara de espectro visível ao ser humano).

Tabela 2.4 - Conjuntos de dados de imagens de cabeça de poço mostrando o número de imagens com vazamentos e sem vazamentos.

Descritores	Autores		
	Wang, J. et al.	Shi, J.H. et al.	Melo, RO; Costa Filho, CFF; Costa, MGF;
Ano	2020	2020	2014
Título	Machine vision for natural gas methane emissions detection using an infrared camera	Real-time leak detection using an infrared camera and Faster R-CNN technique	Leak detection of natural gas with base on the components of color spaces RGB and HSI using novelty filter
Hardware (fabricante e modelo)	FILR GF-320	FILR GF-320	Yokogawa FC15U
Espectro	Infravermelho	Infravermelho	Visível
Conjunto de dados (imagens)	118.060	3.205	3.060
Método	CNN	FR-CNN	Filtro de Novidade
Desempenho	Acc: 95%	IOU: 98%	AUC= 98,72% TFN = 8,78% *ACC=93,15
Fonte da Publicação	Applied Energy	Computers & Chemical Engineering	IEEE Latin America Transactions

Fonte: Compilação do autor.

## 2.7 Conclusão

A revisão da literatura apresentada neste capítulo mostrou que a detecção de vazamento de GN tem sido amplamente estudada ao longo do tempo. Dentre os métodos apresentados, destaca-se o modelo de detecção de vazamento de GN por imagens, obtidas por câmeras especiais munidas de sensores de infravermelho. Ressalta-se a escassez de estudos voltados para detecção de vazamentos de GN que utilizam imagens obtidas por câmara de CFTV, ou seja, conjuntos de dados formados por imagens oriundas de câmeras de sistemas de vigilância, que não requerem sensores específicos ao espectro da molécula do combustível  $CH_4$ .

Fatos como o custo financeiro de aquisição e instalação de câmeras especiais que detectam a molécula de metano ser superior aos das câmeras de CFTV e além da escassez de estudos com esse tipo de imagem permite inferir que, o desenvolvimento de técnicas para a

detecção de GN com imagens de CFTV, poderia amenizar esses vazamentos em unidades de petrolíferas dotadas de CFTV. Como essas câmeras são comuns nas instalações petrolíferas, monitorando estas áreas diuturnamente, conclui-se, então, que o potencial de utilização, por parte da sociedade, do método proposto nesse trabalho, é enorme.

### 3. FUNDAMENTAÇÃO TEÓRICA

A fundamentação teórica tem por objetivo subsidiar a pesquisa a ser desenvolvida, apresentando os conceitos teóricos essenciais para melhor compreensão desta tese, referentes as CNN, as quais são utilizadas para o desenvolvimento do método quantitativo proposto.

Num primeiro momento (Seção 3.1) é realizada uma contextualização sobre redes neurais artificiais, considerando sua origem e evolução até as atuais CNN. Em seguida, na Seção 3.2, define-se o princípio de operação das redes neurais convolucionais, seguida de subseções que apresentam as principais camadas de uma CNN. Na Seção 3.3 apresenta-se a aprendizagem por transferência de conhecimento, aplicada com redes CNN. Por fim, a Seção 3.4 apresenta-se a técnica de mapeamento de ativação de classe por gradiente ponderado (Grad-CAM), a qual fornece uma identificação da região da imagem responsável pela resposta de uma rede CNN.

#### 3.1 Redes neurais artificiais

As Redes Neurais Artificiais (RNA) passaram por três ondas de evolução. A primeira delas, também conhecida como onda cibernética, teve início em 1944 (GOODFELLOW; BENGIO; COURVILLE, 2016). Seu grande momento ocorreu em 1958, quando Rosenblatt desenvolveu o *perceptron* (ROSENBLATT, 1958). A segunda onda foi o conexionismo, cujo ápice foi o desenvolvimento do algoritmo de treinamento de retropropagação (do inglês, *backpropagation*), proposto por Ruineihart, Hinton e Williams (1986). Em seguida, as redes neurais passaram por um período sombrio e caíram em descrença, até a chegada da terceira, que vigora até o presente momento, conhecida como onda do aprendizado profundo (HEATON, 2018). Esse termo foi proposto em 2006 por Hinton, Osindero e Teh (2006). As principais arquiteturas da terceira onda são as redes convolucionais e as redes recorrentes.

Em 1989, as atuais Redes Neurais Convolucionais já eram esboçadas conforme mostrado na Figura 3.1, por LeCun *et al.* (1989).

A primeira arquitetura de rede neural convolucional constituída, denominada de LeNet-5, Figura 3.2, foi apresentada em 1998 por LeCun *et al.* (1998). Porém, desde o trabalho de Fukushima (1979), o aprendizado profundo já era uma ideia ventilada.

Os esboços da Figura 3.1 foram os passos iniciais que contribuíram para que em 2019 a comunidade científica reconhecesse os esforços de Yann LeCun, Yoshua Bengio, Geoffrey Hinton, denominando-os de “Os Pais da Aprendizagem Profunda” (BENGIO; HINTON; LECUN, 2019), agraciando-os com o Prêmio Turing 2018 pela contribuição destes pesquisadores para a área de inteligência artificial (BENGIO; HINTON; LECUN, 2019).

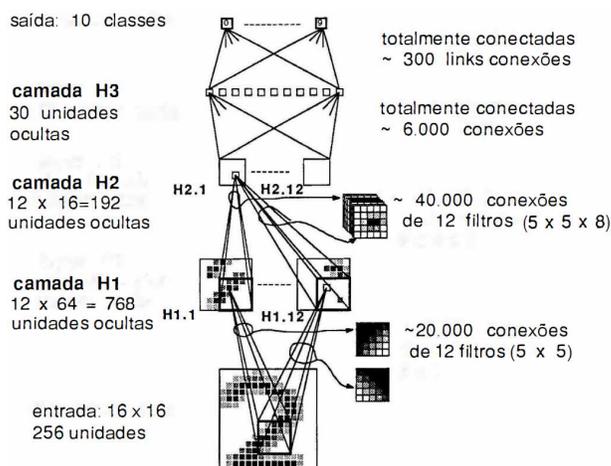


Figura 3.1 - Rede neural convolucional, aplicada às imagens de dígitos manuscritos, proposta por LeCun em 1989, onde o autor já menciona que os coeficientes que compõem o *kernel*  $5 \times 5$  eram determinados automaticamente pela rede. Fonte: LeCun *et al.* (1989).

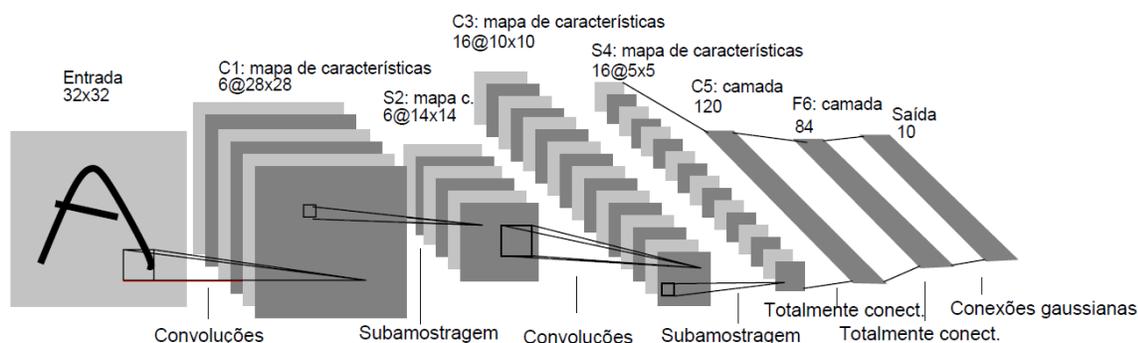


Figura 3.2 - LeNet-5: Rede neural convolucional para reconhecimento de dígitos de 1998. Fonte: LeCun *et al.* (1989).

### 3.2 Redes neurais convolucionais

Notoriamente, foi a partir da repercussão dos resultados alcançados na competição anual do “Desafio de reconhecimento visual em grande escala ImageNet” (ILSVRC, do inglês, *ImageNet Large Scale Visual Recognition Challenge*) de 2012, em que Krizhevsky, Sutskever e Hinton (2012) conseguiram reduzir drasticamente a taxa de erro para reconhecimento de objetos através da aplicação de redes neurais convolucionais, que a comunidade científica começou vislumbrar o real potencial da CNN.

Os motivos que levaram a uma expansão vertiginosa no uso das redes CNN a partir da segunda década do século XXI foram os seguintes: disponibilidade de grandes conjuntos de dados; maior capacidade de processamento computacional, com a introdução das unidades

gráficas de processamento e das unidades tensoriais de processamento, e os desempenhos superlativos obtidos pelas mesmas em um amplo leque de aplicações. A característica diferencial marcante entre as redes RNA (redes inteiramente conectadas da segunda geração) e as redes CNN, é que estas últimas são compostas de diferentes camadas que executam diferentes funções, como a camada convolutiva e a camada de amostragem. Além disso, elas apresentam uma estrutura 3D, com vários planos de neurônios em uma mesma camada. Já as primeiras são constituídas apenas por camadas lineares de neurônios. Outra diferença marcante é que as redes CNN, através das camadas convolutivas, extraem suas próprias características, diferentemente das redes RNA, em que o usuário é que extrai as características, que são disponibilizadas na entrada da rede, em forma de vetores. O processamento da camada convolutiva, como o nome já sugere, utiliza a técnica de convolução para extração de características, técnica essa extensivamente utilizada no processamento digital de imagens. Por último, registra-se que a entrada por excelência das redes convolucionais são imagens e não vetores, como o são nas RNA.

As redes CNNs utilizam, para o seu treinamento, o algoritmo de retropropagação, que foi desenvolvido para o treinamento das RNA da segunda onda.

A arquitetura da CNN é composta, basicamente, pelas camadas de convolução, camadas de amostragem (*pooling*), camadas totalmente conectadas e camadas de classificação (YAMASHITA et al., 2018). No entanto, o número de camadas é mais extenso e a abordagem convolutiva encontra-se em constante expansão.

A seguir são descritas as principais camadas das redes CNN utilizadas nesta pesquisa.

### 3.2.1 Camada convolucional

As CNNs possuem diversos tipos de camadas com funções distintas, porém as camadas que desempenham o papel principal nelas são as camadas convolucionais. Essas camadas criam mapas de ativação a partir dos dados de entrada, através da operação de convolução (GOODFELLOW; BENGIO; COURVILLE, 2016).

A operação de convolução entre dois elementos  $f(t)$  e  $g(t)$  é mostrada na equação (3), em que  $\tau$  representa o deslocamento e  $t$  a variável independente.

$$f(t) * g(t) = \int_{-\infty}^{-\infty} f(t)g(\tau - t)dt \quad (3)$$

Em forma discreta, a equação (3) é definida conforme mostrado na equação (4).

$$f[n] * g[m] = \sum_{m=0}^n f[m]g[m - n] \quad (4)$$

Em aplicação com imagens, dados discretos, a operação de convolução trabalha com duas matrizes. A primeira é a imagem de entrada  $I$ , de dimensão  $i \times j$ , enquanto a segunda é o filtro ou *kernel*  $K$  de dimensão  $m \times n$ . Na equação (5) descreve-se a operação de convolução com esses dois operandos.

$$(K * I)(i, j) = \sum \sum I(i - m, j - n)K(m, n) \quad (5)$$

Devido às características intrínsecas ao processo de aprendizado da rede, ao invés de aplicar diretamente a Equação (5), é implementada a função de correlação cruzada (*cross-correlation*). A diferença entre essa operação e a operação de convolução, é que nela não ocorre a rotação do filtro, conforme mostrado na equação (6).

$$(K * I)(i, j) = \sum \sum I(i + m, j + n)K(m, n) \quad (6)$$

Portanto, a convolução é um tipo de operação linear usado para extração de características, em que uma pequena matriz, denominada de filtro ou *kernel*, é aplicada em uma imagem de entrada, um *tensor*. Na operação de convolução, o *kernel* percorre todo o *tensor* do canto superior esquerdo até o canto inferior esquerdo. Em cada passo da convolução o filtro é superposto a uma pequena região do tensor, ocorrendo o produto entre os elementos do filtro e os elementos do tensor que se encontram subjacentes a ele. O resultado da convolução é a soma desses produtos. A Figura 3.3 ilustra o exemplo de convolução entre o *tensor*  $I^a$  e *kernel*  $K^b$ .

Do resultado da convolução da Figura 3.3 é obtido um mapa de característica de dimensão espacial  $2 \times 2$ , inferior às dimensões do tensor de entrada  $4 \times 4$ . Nesse caso, isso ocorre porque na operação de convolução descrita acima não ocorre o acréscimo de linhas e colunas com zeros nas extremidades do tensor, permitindo que a operação de convolução produza um tensor final com as mesmas dimensões do tensor inicial. Em inglês, essa operação é conhecida como *padding*. A distância entre duas posições sucessivas do kernel sobre o tensor, tanto horizontalmente quanto verticalmente, é conhecido como passo (*stride*). O passo pode assumir um valor de 1 ou valores maiores.

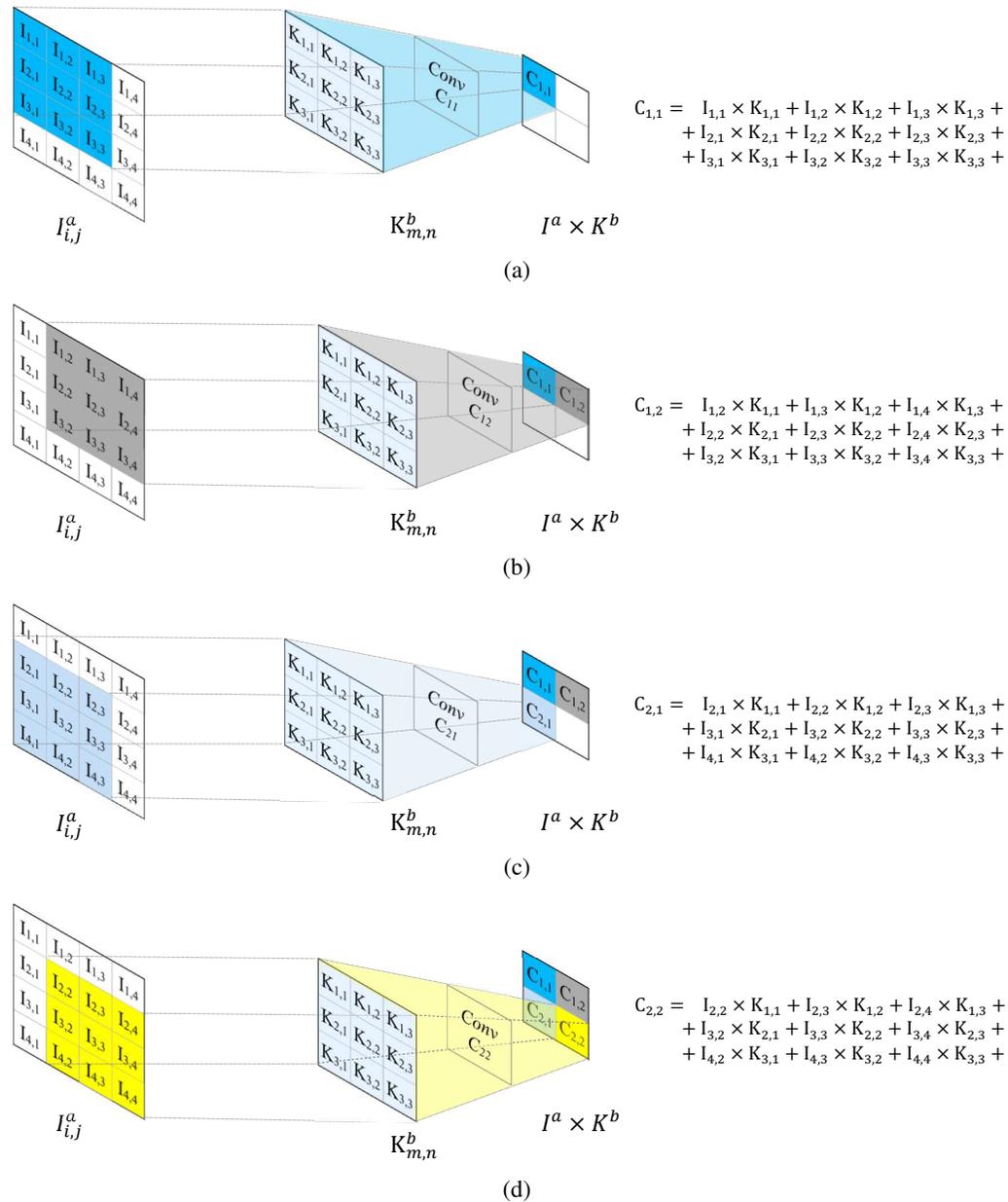


Figura 3.3 - Processo de convolução realizado entre a entrada  $I$  de dimensão  $4 \times 4$  e filtro de dimensão  $3 \times 3$ , passo de  $1 \times 1$  e sem acréscimos de zeros nas extremidades do tensor. Esta configuração resulta na saída da função convolucional  $C$  de dimensão  $2 \times 2$ , a obtenção de cada valor de  $C$  é expressa nas figuras a, b, c e d, respectivamente. Fonte: Compilação do autor.

No entanto, passos maiores do que 1 resultam na diminuição da resolução espacial final da operação de convolução. A Figura 3.4 ilustra o processo de convolução com o acréscimo de linhas e colunas com zeros nas extremidades do tensor e passo (*stride*) de  $1 \times 1$ . Nesse caso, conforme mostrado, o resultado da convolução é um mapa de característica com dimensão  $4 \times 4$ , igual a dimensão do tensor de entrada.

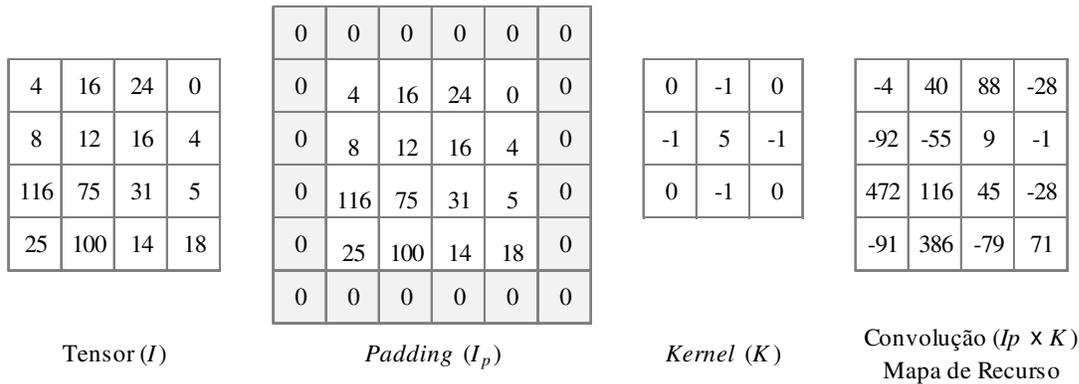


Figura 3.4 - Processo de convolução entre *tensor* I e *kernel* K com acréscimo de zeros de dimensão 1. Fonte: Compilação do autor.

Os principais hiperparâmetros de uma camada convolutiva de uma rede CNN são o tamanho do *kernel*,  $m \times n$ , e o número de *kernels*  $k$  utilizadas pela mesma. Esse último determina a profundidade de uma camada convolutiva da rede CNN, ou seja, a quantidade de mapas de características que será extraído pela camada convolutiva. De acordo com Aggarwal (2018), o *kernel* é geralmente quadrado em termos de dimensões espaciais,  $m \times n$ , e ordem ímpar, sendo muito menor que a imagem de entrada, e comumente sua dimensão é  $3 \times 3$  ou  $5 \times 5$ . Nesta tese, os *kernels* das CNNs propostas possuem dimensão  $3 \times 3$ .

A operação de convolução tem como principal finalidade gerar diferentes características através da aplicação de diferentes filtros. Exemplos de características extraídas são: bordas verticais, bordas horizontais, contornos em forma de hexágonos ou colmeias etc. No último nível de abstração das redes CNN, as características extraídas pelas camadas já correspondem a partes de objetos.

### 3.2.2 Camada de normalização em lote

Muitas vezes confundido como algoritmo de otimização, a camada de normalização em lote (*batch normalization*) é um método para mitigar problemas de desvanecimento ou explosão do gradiente (GOODFELLOW; BENGIO; COURVILLE, 2016). Com essa camada, é possível trabalhar com redes profundas, sem que os valores dos gradientes ao longo da rede se tornem ou muito grandes (explosão do gradiente) ou muito pequenos (desvanecimento do gradiente).

A camada de normalização em lote foi proposta em 2015 por Ioffe e Szegedy (2015), os quais observaram que se pudessem garantir que a distribuição das entradas das funções de ativação permanecessem mais comportadas à medida que a rede era treinada, haveria uma menor tendência das funções de ativação das camadas saturarem, com uma consequente

aceleração do treinamento da rede (GOODFELLOW; BENGIO; COURVILLE, 2016). A solução empregada resolveu o problema denominada pelos autores de variação interna da covariância (*internal covariate shift*).

A solução envolveu uma etapa de normalização, que corrige a média e a variância da saída da camada anterior, adicionando dois parâmetros  $\gamma$  e  $\beta$  a serem aprendidos ao longo do treinamento. A equação (7) expressa o resultado da saída  $y$  da camada de normalização em lote, sendo  $\hat{x}$  a entrada não normalizada.

$$y = \gamma\hat{x} + \beta \quad (7)$$

No entanto, para centralizar e normalizar essa entrada  $\hat{x}$ , é necessário determinar a média  $\mu$  e o desvio padrão  $\sigma$  dos valores de  $\hat{x}$  em um subconjunto dos dados de entrada, ou minilote (*mini-batch*), de tamanho  $m$ . Esses valores são calculados através das equações (8) e (9). Nessas equações, temos que: média do *mini-batch* =  $\mu_B$ , variância do *mini-batch* =  $\sigma_B$ . Cada valor de entrada  $x_i$  é centralizado e normalizado através da equação (10). A constante  $\epsilon$  é um valor extremamente pequeno para evitar uma indeterminação na equação (10).

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i \quad (8)$$

$$\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (9)$$

$$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (10)$$

Portanto, a saída da operação de normalização em lote ( $y_i$ ) é representada na equação (11).

$$y_i = \gamma\hat{x}_i + \beta \quad (11)$$

### 3.2.3 Função de ativação retificadora linear (*Rectifier Linear Unit* - ReLU)

Atualmente, a função ReLU é a função de ativação mais utilizada nas camadas escondidas de uma CNN (GLOROT; BORDES; BENGIO, 2011). A não linearidade da mesma evita que o conjunto das camadas de uma rede CNN produza um mapeamento linear entre entrada e saída. Essa função foi utilizada nas camadas escondidas de todas as redes CNN utilizadas nesse trabalho.

A função de ativação ReLU é expressa pela equação (12). Dado um elemento  $x$ , a saída da função ReLU é definida como o próprio elemento  $x$  para  $x > 0$ , e para os casos que  $x \leq 0$  a saída da ReLU será 0 (GLOROT; BORDES; BENGIO, 2011).

$$ReLU(x) = \max(x, 0) \quad (12)$$

Portanto, a ReLU fornece uma transformação não linear de baixo custo computacional, sendo o resultado de sua derivada os valores 0 ou 1. Como o aprendizado da rede neural dá-se através do algoritmo de *backpropagation*, que utiliza a derivada da função de ativação das camadas da rede CNN, valores de derivada de 0 ou 1, simplificam consideravelmente o processo de aprendizado (AGGARWAL, 2018; GOODFELLOW; BENGIO; COURVILLE, 2016).

O gráfico da Figura 3.5 ilustra o comportamento da função  $ReLU(x)$ .

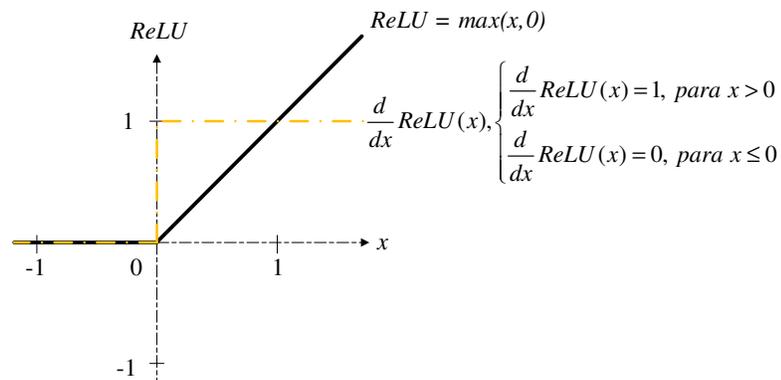


Figura 3.5 - Comportamento da função de ativação ReLU. Fonte: Adaptada de Aggarwal (2018).

Como exemplo, a função de ativação ReLU foi aplicada ao mapa de características da Figura 3.4 (d), sendo o resultado indicado na Figura 3.6.

-4	40	88	-28	ReLU =	0	40	88	0
-92	-55	9	-1		0	0	9	0
472	116	45	-28		472	116	45	0
-91	386	-79	71		0	386	0	71
Camada anterior ( $I_p \times K$ )					Saída da camada ReLU			

Figura 3.6 - Resultado da aplicação da função de ativação ReLU ao mapa de característica da Figura 3.4 (d). Fonte: Compilação do autor.

Ainda, de acordo com os testes realizados por Krizhevsky, Sutskever e Hinton (2012), uma rede neural convolucional de quatro camadas, com funções ReLU de ativação, atingiu uma taxa de erro no treinamento de 25% no conjunto de dados do Instituto Canadense de Pesquisa Avançada (CIFAR-10, do inglês, *Canadian Institute for Advanced Research*), em um tempo seis vezes mais rápido do que uma rede equivalente que utiliza a função de ativação tangente hiperbólica.

### 3.2.4 Camada de amostragem (*pooling*)

Também conhecidas como camadas de agrupamento, as camadas de amostragem são aplicadas posteriormente a uma camada de convolução, com o objetivo de reduzir o tamanho do mapa de características.

Dada a saída de uma camada de convolução, a camada de amostragem realiza uma operação de redução de tamanho, tornando a rede mais invariável a translação através de pequenos deslocamentos (YAMASHITA et al., 2018; ZEIMARANI et al., 2020). A camada de amostragem não contém parâmetros a serem aprendidos através de um treinamento. Os hiperparâmetros dessa camada definidos pelo usuário são: tipo, tamanho e passo do filtro.

Os principais tipos de amostragem são o *max-pooling* (*max*) e o *global average pooling* (*avg*). A Figura 3.7 ilustra esses dois tipos de amostragem com filtros de tamanho  $(2 \times 2)$  e um passo de 2. A dimensão de profundidade do volume de entrada é mantida inalterada, porém a dimensão da saída é reduzida (AL-WAISY et al., 2018). O filtro move-se da esquerda para a direita e de cima para baixo.

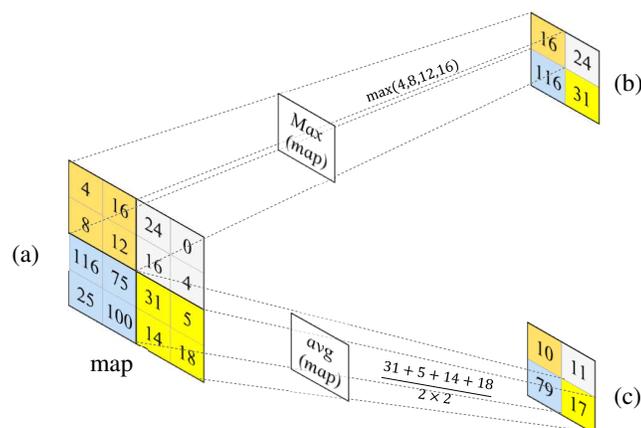


Figura 3.7 - Representação visual de duas camadas de amostragem (*max* e *avg*): (a) mapa de característica da camada anterior que é entrada para camada de amostragem; (b) saída da camada de amostragem *max* e (c) saída da camada de amostragem *avg*. Fonte: Adaptada de Ozkaya, Ozturk e Barstugan (2020).

A cada passo de aplicação do filtro são obtidos: o valor máximo dos elementos do mapa original subjacentes a posição do filtro, através da operação de amostragem *max*. Matematicamente, a amostragem *max* é representado pela equação (13) e a amostragem *avg* pela equação (14) (AL-WAISY et al., 2018; CHEN; XU; KOCH, 2020).

$$y_{max-pooling} = \max_{i,j=1}^{h,w} (\text{map}_{i,j}) \quad (13)$$

$$y_{avg-pooling} = \frac{1}{hw} \sum_{i,j=1}^{h,w} (\text{map}_{i,j}) \quad (14)$$

### 3.2.5 Camada *dropout*

Uma técnica bastante utilizada para melhoria da generalização de uma rede CNN é a técnica de *dropout*. Introduzido por Srivastava *et al.* (2014), o *dropout* é um método de regularização que busca evitar que as CNN se ajustem excessivamente ao conjunto de treinamento.

Essa técnica consiste em, para cada minilote do treinamento, selecionar aleatoriamente um grupo de neurônios de uma camada e zerar a saída dos mesmos. Vem daí o nome de abandono ou *dropout* de neurônios. Assim, durante o treinamento de um minilote, os neurônios possuem uma probabilidade  $p$  de serem eliminados, ou seja, uma parcela desse grupo de neurônios não contribui nem para a passagem direta, nem para a passagem reversa do método de otimização de *backpropagation*. Com isso, diminui-se possíveis coadaptações complexas entre os nós da rede, evitando que interdependências surjam entre eles. A Figura 3.8 ilustra a aplicação da técnica de *dropout* no treinamento de uma rede neural.

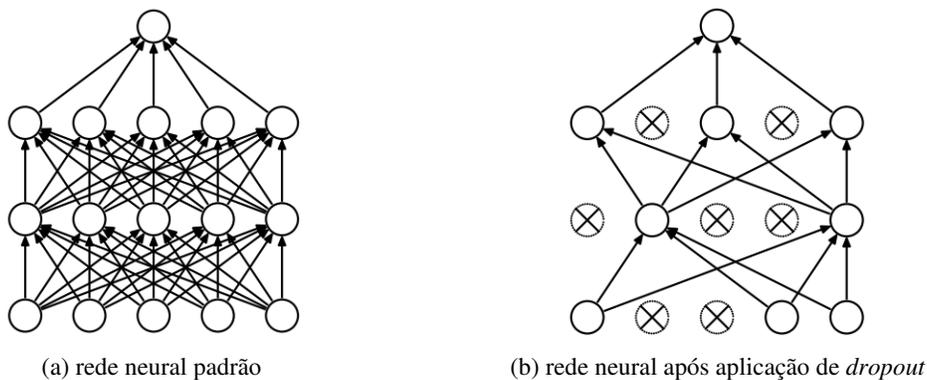


Figura 3.8 - *Dropout* em Redes Neurais: (a) modelo de rede neural; (b) *dropout* aplicado ao modelo com probabilidade de 0,5 de abandono de neurônios (x indica os neurônios abandonados). Fonte: Adaptada de Srivastava *et al.* (2014).

De acordo com Aggarwal (2018), é comum excluir neurônios com probabilidade entre 20 e 50%, pois esta faixa costuma apresentar bons resultados.

### 3.2.6 Camadas totalmente conectadas (*fully-connected layers*)

As camadas totalmente conectadas possuem estruturas idênticas às camadas ocultas utilizadas nas RNA tradicional (SONG; WOO; KIM, 2020). Diferentemente das camadas convolucionais, que possuem conexões esparsas, as camadas totalmente conectadas são densamente conectadas. Cada neurônio da camada FC (*fully-connected*) é conectado a todos os neurônios da camada anterior. Em virtude desse adensamento de conexões, a maioria dos parâmetros da CNN está nas camadas FC, as quais geram efetivamente informações semânticas globais (AGGARWAL, 2018).

A saída de uma camada que constitui a entrada de uma camada FC possui mais de uma dimensão, como por exemplo a saída da camada de agrupamento, sendo então necessário converter essa matriz (2D) em um único vetor (1D), que é utilizado como entrada da FC, conforme ilustrado na Figura 3.2.

## 3.3 *Transfer Learning* (TL)

Aggarwal (2018) afirma que, nas CNNs, as características nas camadas anteriores capturam formas detalhadas, porém primitivas, como linhas ou bordas. Por outro lado, as características nas camadas posteriores capturam formas de maior complexidade, como hexágonos, favos de mel e assim por diante. As últimas camadas da CNN modelam formas semanticamente interpretáveis, geralmente, tendo correlações mais próximas com os rótulos de classe no domínio da imagem.

Logo, as camadas iniciais e medianas de CNNs que foram treinadas com um grande conjunto de dados, retêm as mesmas formas primitivas e intermediárias, mantendo sua usabilidade, independentemente do tipo de tarefa e conjunto de dados em que são aplicados.

A transferência de conhecimento (TL do inglês, *transfer learning*), faz uso da capacidade das CNNs que foram treinadas com grandes conjuntos de dados e que já possuem o conhecimento necessários para modelar as formas básicas e intermediárias, por isso são conhecidas como redes pré-treinadas.

Conforme Yosinski *et al.* (2014), a aprendizagem por transferência de conhecimento pode ser uma ferramenta poderosa para permitir o treinamento de uma rede CNN sem gerar sobreajustamento (*overfitting*), principalmente quando a tarefa de destino tem um conjunto de dados

escasso, na medida em que, no domínio de origem, a rede foi treinada com um grande conjunto de dados, ou seja, já possui um alto nível de aprendizado (PAN; YANG, 2010).

Na resolução de um problema de um novo domínio, um novo modelo é construído mantendo os parâmetros das primeiras  $n$  camadas de um modelo pré-treinado e excluindo as camadas finais da rede pré-treinada (totalmente conectada, *softmax* e classificação). As camadas excluídas são substituídas por novas camadas (totalmente conectada, *softmax* e classificação) que são dimensionadas e treinadas tendo em vista o problema a ser resolvido no novo domínio (GARG et al., 2020).

A aplicação da TL é ilustrada na Figura 3.9. Primeiramente, o usuário seleciona um modelo de CNN pré-treinado e determina quais as camadas da CNN que terão seus recursos, excluindo as camadas subsequentes da arquitetura, conforme mostrado na Figura 3.9 (a). Novas camadas são então adicionadas com pesos inicializados a partir de uma distribuição de probabilidade ou com todos os pesos com valor zeros ou uns, por exemplo, Figura 3.9 (b). Finalmente o modelo é ajustado, através de um treinamento, ao conjunto de dados do novo domínio, sendo a rede posteriormente validada e testada, conforme mostrado na Figura 3.9 (c) (PAN; YANG, 2010).

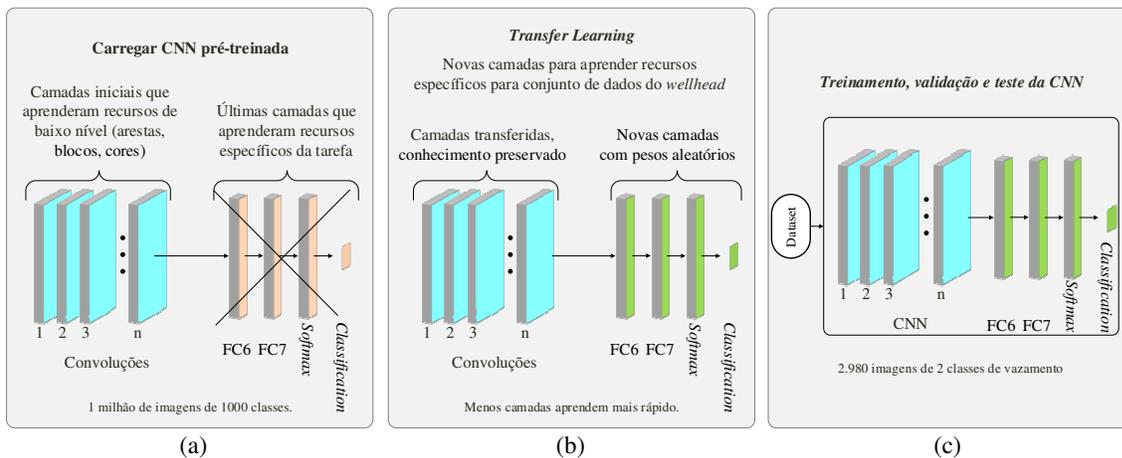


Figura 3.9 - Processo de transferência de aprendizagem aplicado as imagens do *wellhead*, incluindo: (a) carregamento de um modelo pré-treinado e remoção das quatro últimas camadas *fully connected* (FC) FC6 e F7; *softmax* e *classification*); (b) substituição das camadas removidas por camadas novas com valores de pesos aleatórios; (c) ajustar as camadas anteriores e treinar as novas camadas com o conjunto de dados do *wellhead*. Fonte: Adaptada e traduzida de Huang *et al.* (2019).

Outro ponto relevante do TL é que ele é capaz de reduzir o tempo de treinamento, não sendo necessário treinar a rede do zero, pois as camadas iniciais utilizam parâmetros já aprendidos em um treinamento anterior, utilizando uma base de dados maior. Dependendo do

número de camadas que são preservadas do modelo anteriormente, o TL pode ser classificado em ajuste fino (*fine tuning*) ou ajuste raso (*shallow tuning*).

No ajuste raso, também denominado de método de congelamento de pesos, os parâmetros das camadas convolucionais do modelo pré-treinado são preservados, de modo que apenas os parâmetros das últimas camadas (camadas que foram substituídas por novas) são ajustadas durante o treinamento. Quanto ao ajuste fino, os parâmetros do modelo pré-treinado são ajustados durante o treinamento com o novo conjunto de dados da tarefa de destino, o que torna o processo de treinamento mais rápido do que treinar a rede em que os parâmetros seriam iniciados aleatoriamente (ABBAS; ABDELSAMEA; GABER, 2021; DEL RIO et al., 2018; GARG et al., 2020).

### 3.4 Grad-CAM

Embora a previsão correta da ocorrência de um vazamento possa minimizar o impacto negativo de um vazamento, evitando acidentes de proporções maiores, essa inferência de vazamento não fornece nenhuma pista da presença do vazamento na imagem do poço. Para identificação na imagem da região em que o vazamento está ocorrendo, essa tese propõe usar uma técnica de identificação da região da imagem responsável pela resposta positiva de uma máquina de reconhecimento de padrões em relação à presença do vazamento. A visualização fornecida por essa técnica permite identificar qual parte da imagem é relevante para o vazamento. Antes de descrever a técnica Grad-CAM, descreve-se a técnica que a inspirou.

Os autores Zhou *et al.* (2016), propuseram um método que traria mais transparência ao funcionamento das CNNs, realçando as áreas dos mapas de características que foram mais ativadas ao predizer uma determinada classe  $c$ . Os autores definiram a técnica como Mapeamento de Ativação de Classe (CAM, do inglês *class activation mapping*).

O CAM consiste em realizar um agrupamento médio global de todos os  $k$  mapas de uma mesma camada convolucional  $A$ , camada essa anterior a camada de saída final (*softmax*). Então, é realizada uma média ponderada desses agrupamentos para obter o mapa final. Como este mapa final possui dimensão espacial igual à da camada convolucional  $A$ , ele é redimensionado para o tamanho da imagem de entrada  $Im$ , fornecendo assim um mapa de calor que é sobreposto à imagem  $Im$ . A Figura 3.10 ilustra o processo proposto por Zhou *et al.* (2016), apresentando um mapa de calor das áreas da imagem de entrada que foram mais significativas (cor vermelha) para predição da classe  $c$ , terrier australiano.

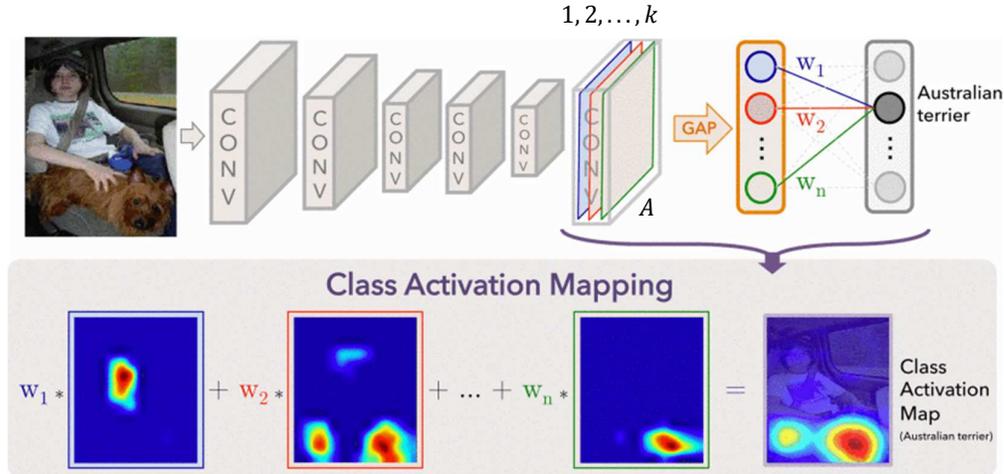


Figura 3.10 - Mapeamento de ativação de classe: a pontuação de classe prevista  $c$  é mapeada na última camada convolucional  $A$ , gerando os mapas de ativação de classe (CAMs) que são redimensionados, destacando as regiões discriminativas específicas da classe  $c$ . Fonte: Adaptada de Zhou *et al.* (2016).

Já o mapeamento de ativação de classe com gradiente ponderado (Grad-CAM, do inglês *gradient-weighted class activation mapping*), proposto por Zhou (2019), utiliza informações de gradiente que fluem para a última camada convolucional da CNN, buscando compreender a relevância de cada neurônio para uma decisão de interesse.

A Figura 3.11, ilustra a aplicação do Grad-CAM para extração do mapa de calor de uma imagem  $Im$  que a CNN classificou como “coelho”, obtendo a região que mais contribuiu para decisão da CNN. Os *pixels* cuja tonalidade mais se aproximam da tonalidade vermelha têm forte contribuição para a resposta positiva da rede, enquanto os *pixels* que mais se aproximam da tonalidade azul, têm baixa relevância para a resposta positiva da rede.

A obtenção do mapa de calor da Figura 3.11 inicia-se calculando o gradiente do *softmax* para classe  $c$  em  $y^c$  em relação às ativações de todos os mapas de características da última camada convolucional  $A^k$ , ou seja,  $\frac{\partial y^c}{\partial A^k}$ .

Então, a estes gradientes é aplicada a operação de agrupamento por média global,  $\alpha_c^k$ , obtendo-se a importância de cada mapa de característica  $k$  para cada classe  $c$ , conforme a equação (15).

$$\alpha_c^k = \overbrace{\frac{1}{Z} \sum_i \sum_j}^{\text{média global}} \underbrace{\left( \frac{\partial y^c(Im)}{\partial A_{ij}^k(Im)} \right)}_{\substack{\text{gradientes} \\ \text{via backprop}}} \quad (15)$$

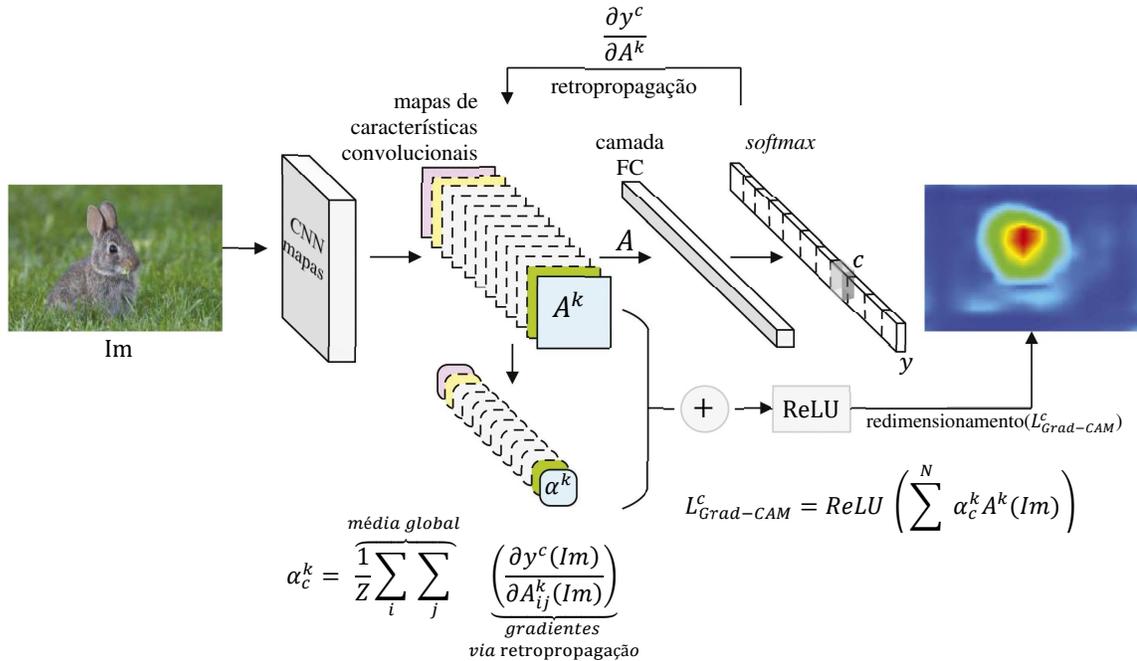


Figura 3.11 - Uma vis\u00e3o geral do m\u00e9todo Grad-CAM. Fonte: Adaptado e traduzido de Selvaraju *et al.* (2017) e Chattopadhyay *et al.* (2018).

Em seguida, utilizando a fun\u00e7\u00e3o de ativa\u00e7\u00e3o ReLU \u00e9 executada uma combina\u00e7\u00e3o ponderada de mapas de ativa\u00e7\u00e3o direta, equa\u00e7\u00e3o (16) (SATTARZADEH *et al.*, 2021),.

$$L_{Grad-CAM}^c = ReLU \left( \sum_{k=1}^N \alpha_c^k A^{lk}(Im) \right) \quad (16)$$

O mapa de calor final do Grad-CAM \u00e9 obtido atrav\u00e9s de uma fun\u00e7\u00e3o de sobre amostragem (interpola\u00e7\u00e3o bilinear) aplicado ao  $L_{Grad-CAM}^c$ , equa\u00e7\u00e3o (16), que o redimensiona para a resolu\u00e7\u00e3o da imagem de entrada  $Im$  (SELVARAJU *et al.*, 2017).

Os estudos de Zhou *et al.* (2019) e Liu, Zhou e Qian (2020) demonstram ser poss\u00edvel tamb\u00e9m aplicar o conceito do Grad-CAM em arquiteturas que operam com o conceito de fus\u00e3o de caracter\u00edsticas (*feature fusion*).

### 3.5 Aumento de dados

Quando a capacidade de representa\u00e7\u00e3o de um modelo \u00e9 muito alta (rede com muitos par\u00e2metros), de tal maneira que ele distingue adequadamente os detalhes contidos no conjunto de treinamento e valida\u00e7\u00e3o. Por\u00e9m, \u00e9 frequente, em modelos muito complexos, que, quando uma nova entrada, que n\u00e3o pertence ao conjunto de treinamento e valida\u00e7\u00e3o, \u00e9 apresentada, o

modelo classificar de maneira errada. Este tipo de problema é conhecido como sobreajustamento (do inglês, *overfitting*). Normalmente, deve-se à pequena quantidade de dados utilizada na etapa de treinamento.

As CNNs foram concebidas para trabalharem com grandes quantidade de dados, como é o caso do ImageNet, que é composta por milhões de imagens (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), o que permite o ajuste adequado de todos os parâmetros da mesma.

Contudo, em muitas aplicações, os conjuntos de dados são limitados e, incrementá-los, com novas amostras, pode ser uma tarefa difícil. Um exemplo disso foi observado nos estudos iniciais, no ano de 2020, para detecção do coronavírus (COVID-19 - *Coronavirus Disease*), através da análise de imagens de raio-X. Nesse ano, o conjunto de imagens de raio-X sobre o tema era escasso, sendo incrementado à medida que novos pacientes eram diagnosticados com a doença (GARG et al., 2020). Neste contexto, observa-se que, conjuntos de dados limitados, podem contribuir para o surgimento do problema de sobreajustamento (*overfitting*) em CNN.

No entanto, o número reduzido de imagens em um conjunto de dados pode ser um desafio, mas não é um impedimento para o desenvolvimento de CNNs eficientes. Bases de dados pequenas é uma característica da área médica. No entanto, para contornar tal problema, técnicas de aumento de dados, como rotação e espelhamento e imagens, podem ser utilizadas. Tais procedimentos foram utilizados no trabalho de Miyagawa *et. al* (2019), onde os autores utilizaram CNN para detecção de bifurcação vascular em imagens de tomografia ótica intravascular (IVOCT - *Intravascular Optical Coherence Tomography*). Nesse trabalho, o conjunto inicial de dados era composto por 1361 imagens. Utilizando técnicas de aumento de dados, os autores obtiveram uma acurácia de detecção de 99,72%. Outro exemplo é reportado no trabalho de Costa *et. al* (2019), onde os autores segmentaram lesões presentes em imagens de ultrassom de mama. O conjunto inicial era composto de apenas 550 imagens. Utilizando técnicas de aumentos de dados, os autores obtiveram uma acurácia de segmentação de 91,80%.

O sobreajustamento pode ser exemplificado através de um experimento hipotético de classificação binária, esferas azuis (verdadeiro positivo) ou esferas vermelhas (verdadeiro negativo). Dado um conjuntos de dados utilizados para ajustar uma CNN, o mesmo é dividido nos subconjuntos “treinamento”, “validação” e “teste” (Figura 3.12).

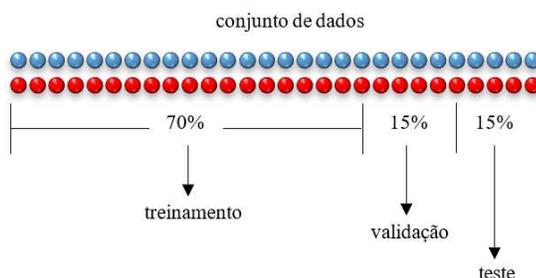


Figura 3.12 - Tarefa de classificação binária: esferas azuis ou esferas vermelhas. Fonte: Compilação do autor.

A CNN então é treinada e validada com os subconjunto “treinamento” e “validação”, respectivamente. Ao final da validação a CNN fornece uma acurácia de 100%, Figura 3.13 (a). Porém, com os pesos da CNN estabelecidos, ao testar essa rede com o subconjunto de dados “teste”, o desempenho da CNN é inferior, fornecendo agora erros de falsos de positivos e falsos negativos, Figura 3.13 (b), ou seja, o modelo está super ajustados para as amostras dos conjuntos “treinamento” e “validação”. No entanto, para elementos que não fizeram parte destes conjuntos dois conjuntos de dados, o modelo apresenta erros de classificação. No âmbito científico, esse tipo de falha é conhecido como sobreajustamento (*overfitting*).

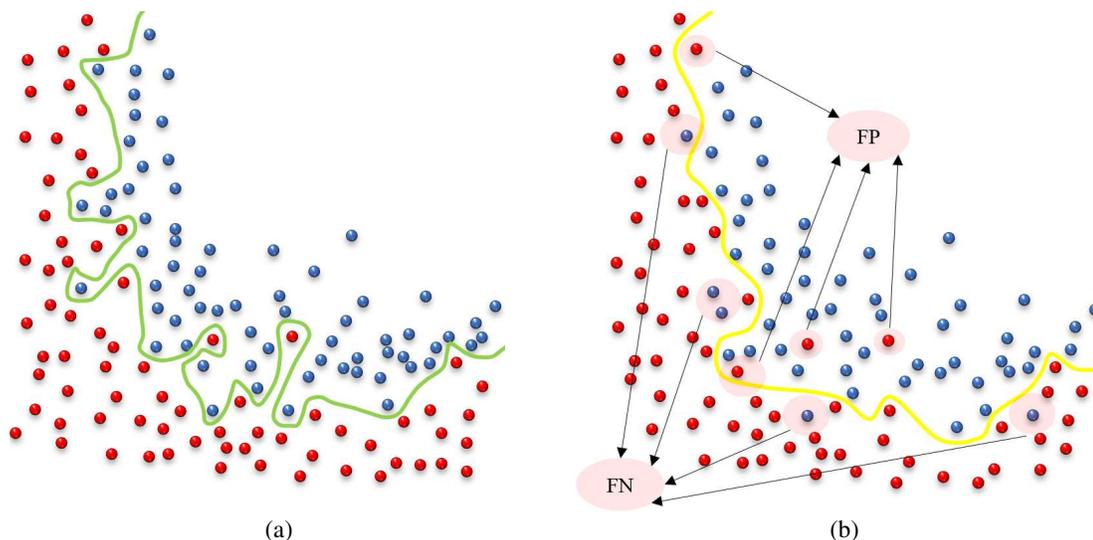


Figura 3.13 – Exemplo de CNN com *overfitting*: (a) desempenho da CNN após treinamento e validação e (b) desempenho da mesma CNN ao ser testada com novo conjunto de dados, resultando em falhas de falso positivo (FP) e falso negativo (FN). Fonte: Compilação do autor.

Portanto, a fim de prevenir o sobre ajustamento relacionado a pequenos conjuntos de dados, é possível incrementar o banco de dados de forma artificial, através da técnica de aumento de dados (*data augmentation*). Essa técnica consiste em realizar transformações em cada elemento do conjunto de dados, gerando novos elementos modificados, possibilitando, assim, o aumento do conjunto de dados. No caso de imagens, as transformações mais

utilizadas são rotação, deslocamento horizontal e vertical, redução do tamanho da cena da imagem, inversão de cor e espelhamento da imagem (TALUKDAR; BISWAS; GUPTA, 2018).

### **3.6 Conclusão**

Neste capítulo, conceitos fundamentais relacionados às redes CNN foram apresentados, como suas principais camadas e hiperparâmetros. A definição desses hiperparâmetros será feita no capítulo seguinte, onde serão apresentadas as arquiteturas de CNN propostas para a detecção de vazamento de GN, detalhando-se o número de camadas, o tamanho dos filtros de cada camada CNN, o tamanho dos filtros das camadas de amostragem etc.

Para o treinamento de arquiteturas de redes CNN, é necessário a disponibilidade de grandes conjuntos de dados. Quando não se dispõe de um grande banco de dados, como é o caso da base disponível nessa tese, a literatura sugere utilizar-se técnicas de aumento de dados ou a transferência de conhecimento. Essa foi a razão pela qual esses dois tópicos foram abordados nesse capítulo.

Por fim, com o objetivo de mapear a ocorrência do vazamento de GN em imagens de CFTV, o capítulo apresentou a técnica de Grad-CAM.

## 4. MATERIAIS E MÉTODOS

Neste capítulo são contextualizados os materiais e os métodos utilizados nas diferentes etapas do desenvolvimento do modelo de detecção de vazamento de GN através da aprendizagem profunda.

### 4.1 Materiais

Nesta seção apresenta-se o ambiente em que ocorreu o desenvolvimento dos modelos de CNN propostos e os conjunto de dados utilizados no treinamento, validação e testes destes modelos.

#### 4.1.1 Ambiente de desenvolvimento

A aquisição das imagens que compõem o conjunto de dados *db1*, utilizadas no treinamento, validação e teste das redes CNN propostas para a detecção de vazamento de GN, foi realizada através da câmera de modelo FC15U CCD da Yokogawa®, conforme apresentado na Figura 4.1 (YOKOGAWA, 2009). Este *hardware* é dedicado ao sistema de vigilância da área do poço, sendo a câmera monitorada pelo circuito fechado de televisão. Na oportunidade, a câmera foi configurada para capturar imagens do tipo RGB com resolução de  $320 \times 240$  *pixels*. As imagens do *db2*, conjunto de dados utilizado para verificar o desempenho dos modelos quando a imagens contendo nuvens do céus, foram capturadas com câmera portátil Panasonic® DMC-FH12 com resolução máxima de  $3456 \times 4608$  *pixels*.

O hardware utilizado nesta pesquisa para o desenvolvimento das CNN possui a seguinte configuração: Processador Intel™ Xeon W-2145 e GPU Nvidia® GV100 e software Matlab® versão 2020b.



Figura 4.1 - Câmera do CFTV: CCD FC15U da Yokogawa®. Fonte: Fabricante Yokogawa (2009).

#### 4.1.2 Conjuntos de dados

Os conjuntos de dados deste estudo é formado por imagens digitais da cabeça de um poço petrolíferos terrestre.

##### 4.1.2.1 Primeiro conjunto de dados: *db1*

O primeiro conjunto de dados de imagens, denominado por *db1*, é composto por imagens obtidas de uma única cabeça de poço (*wellhead*). Tais imagens já foram utilizadas em pesquisas anteriores que propuseram a detecção de vazamento de GN através do método filtro de novidade (COSTA FILHO; MELO; COSTA, 2013; MELO; COSTA; COSTA FILHO, 2014). O conjunto de dados original é composto por 2.810 imagens com dimensão de  $184 \times 160$  *pixels* (MELO; COSTA FILHO; COSTA, 2014). No estudo de 2014 (MELO; COSTA FILHO; COSTA, 2014), esse conjunto de dados original foi complementado com novas imagens da mesma cabeça de poço, totalizando 3.060 amostras.

Nesta tese, a fim de reduzir um possível sobre ajustamento, utilizamos o grupo com maior quantidade de dados (3.060 imagens). Além disso, imagens de cenários e conteúdos semelhantes foram excluídos. Assim, após a exclusão, o número total de imagens no *db1* é de 2.980 imagens, divididas em 2.000 imagens da classe “sem vazamento” e 980 imagens da classe “com vazamento” (MELO; COSTA; COSTA FILHO, 2014). A Figura 4.2 apresenta uma amostra de imagens do *db1*.

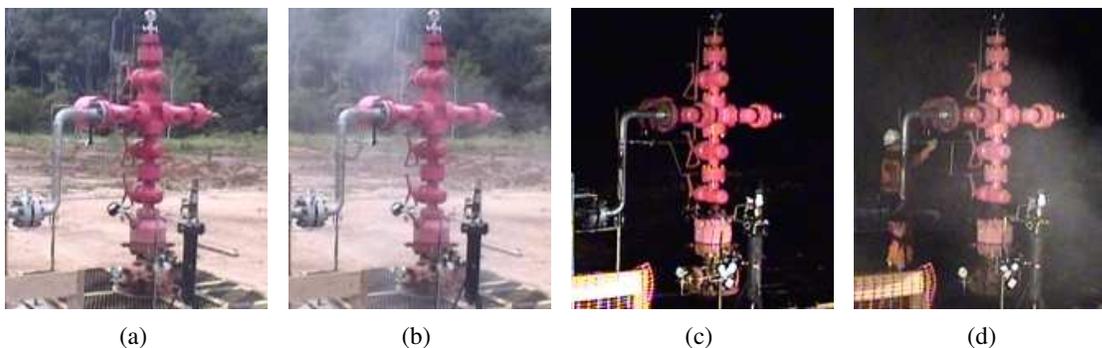


Figura 4.2 - Amostra de imagens do conjunto de dados de imagens *db1*: (a) Imagem diurna sem vazamento de gás (sem a nuvem de hidrocarbonetos), (b) Imagem diurna com vazamento de gás (com a nuvem de hidrocarbonetos), (c) Imagem noturna sem vazamento de gás (sem a nuvem de hidrocarbonetos ), (d) Imagem noturna com vazamento de gás (com a nuvem de hidrocarbonetos). Fonte: Melo, Costa e Costa Filho (2014).

As imagens que compõem o *db1* foram capturadas sobre diversas situações, incluindo imagens diurnas com e sem chuva, imagens noturnas com e sem chuva, além de imagens com

a presença de operadores (humanos). Estas condições e cenários, além de refletirem as situações reais e do cotidiano do poço, buscam dar robustez ao modelo.

#### 4.1.2.2 Segundo conjunto de dados: *db2*

Um segundo conjunto de dados, aqui denominado por *db2*, é constituído por novas imagens de diversos poços, diferentes do poço do *db1*.

O *db2* consiste em 145 imagens de 113 cabeças de poço distintos. Este conjunto contém imagens com dimensões entre  $480 \times 640$  e  $3456 \times 4608$  *pixels*, diferentes ângulos e distâncias das cabeças de poço e, em sua maioria, diferentes planos de fundo, contendo nuvens no céu. A sua captura foi realizada através de câmeras digitais portáteis. Além de permitir uma avaliação do poder de generalização dos modelos de CNN propostos, o objetivo principal deste segundo conjunto de dados foi verificar se o modelo de CNN escolhido consegue distinguir as nuvens no céu (fundo) das nuvens de gás natural. Portanto, todas as 145 imagens no conjunto *db2* são da classe “sem vazamento”, mas contendo nuvens no céu. Uma amostra das imagens das novas cabeças de poço é mostrada na Figura 4.3. Observa-se, na maioria delas, a presença de nuvens no plano de fundo.

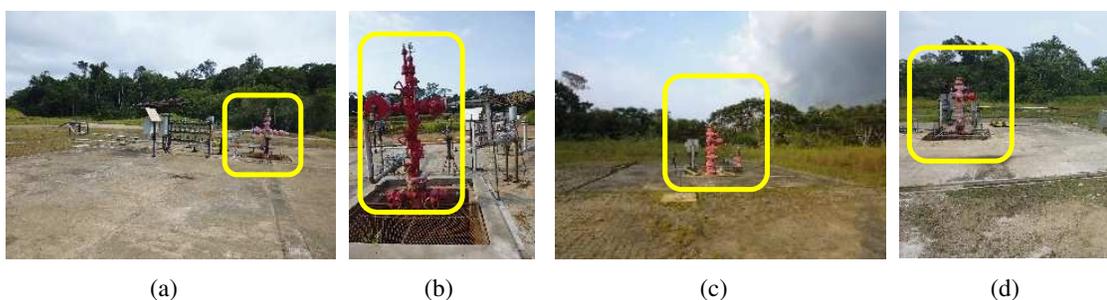


Figura 4.3 - Imagens de amostra do segundo conjunto de dados (*db2*) das novas cabeças de poço, em destaque na cor amarela: (a) Imagem panorâmica da cabeça de poço “X”, (b) Imagem de perfil da cabeça de poço “Y”, (c) Imagem panorâmica da cabeça de poço “W” e (d) imagem de fundo da cabeça do poço “Z”. Fonte: Compilação do autor.

A Tabela 4.1 mostra o número de imagens nos dois conjuntos de dados (*db1* e *db2*), em função do número de imagens por classe.

Ressalta-se que, para obtenção das imagens dos vazamentos, foram realizadas simulações de vazamento de GN na cabeça do poço, por meio da abertura de uma válvula agulha de forma intencional e controlada. Para evitar acidentes ou danos desnecessários ao meio ambiente, o volume de gás liberado durante as simulações de vazamento foi o mínimo necessário para a coleta das imagens utilizadas nos bancos de dados deste trabalho.

Tabela 4.1 - Conjuntos de dados de imagens de cabeça de poço mostrando o número de imagens com vazamentos e sem vazamentos.

Classes de Imagens	1º conjunto de dados (db1)	2º conjunto de dados (db2)	Total
<i>sem vazamento</i>	2.000	145	2.145
<i>com vazamento</i>	980	--	980
Total	2.980	145	3.125

Fonte: Compilação do autor.

## 4.2 Métodos

No diagrama de blocos da Figura 4.4 são apresentadas as etapas da metodologia utilizadas no desenvolvido da pesquisa.

Conforme mostrado na Figura 4.4, inicialmente, as imagens passam por uma etapa de pré-processamento, cujo objetivo é fazer uma adequação nas dimensões das imagens. Em seguida, com o intuito de aumentar o banco de dados e evitar possíveis erros de sobreajustamento, aplica-se a técnica de aumento de dados. O conjunto de dados aumentado é então dividido em três subconjuntos: treinamento, validação e teste, os quais são detalhados na Seção 4.1.2, denominada de Conjuntos de dados.

Em seguida, três arquiteturas de CNN são propostas para a detecção do vazamento de GN. Essas arquiteturas são apresentadas na Seção 4.2.3.1. Quanto ao treinamento delas, com o objetivo de obter-se um melhor desempenho, diferentes algoritmos de otimização e diferentes técnicas de regularização são aplicadas. A combinação entre os algoritmos de otimização e as técnicas de regularização, resultaram na avaliação de 27 modelos de redes CNN. Ao final do treinamento e validação dos vinte sete modelos, determina-se o Modelo com Melhor Desempenho (MMD), o qual é testado com o conjunto de dados de teste. O treinamento, validação e teste das arquiteturas propostas são detalhados na Seção 4.2.4.

Uma vez determinada o MMD, outros cinco modelos de redes CNN são investigados na tarefa de detecção de vazamento de gás natural. Quatro desses modelos correspondem a redes pré-treinadas, descritas na Seção 4.2.3.2. Esses modelos são treinados, validados e testados através da técnica de aprendizagem por transferência de conhecimento.

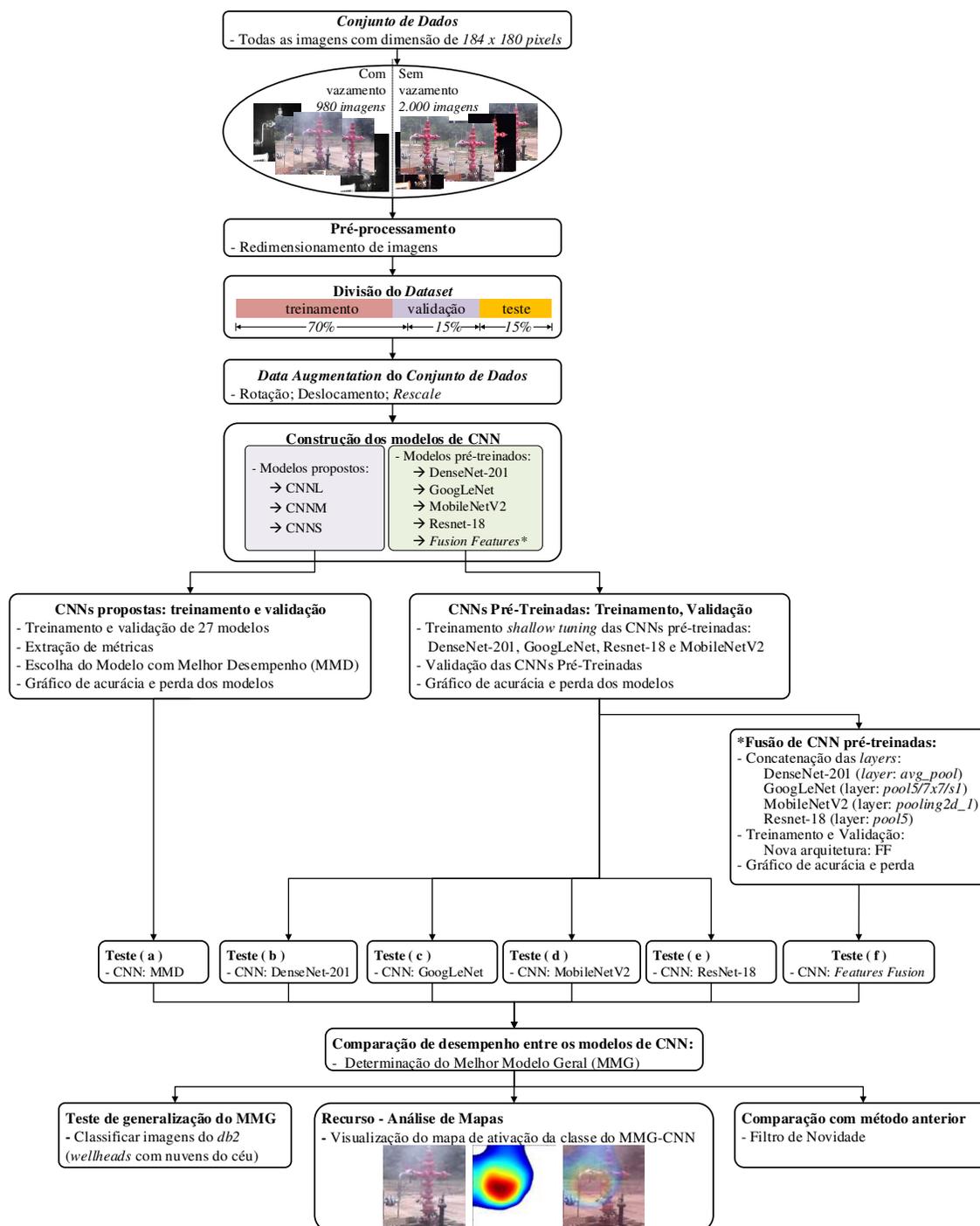


Figura 4.4 - Visão geral da proposta. Fonte: Compilação do autor.

Um último modelo de CNN investigado utiliza a fusão de características resultantes de quatro redes pré-treinadas, já mencionadas no parágrafo anterior. A caracterização desse modelo é feita na Seção 4.2.3.3, denominada de CNN FF baseada em fusão de características.

#### 4.2.1 Pré-processamento de dados

Em relação às dimensões das imagens do conjunto de dados, Aggarwal (2018) afirma que as CNNs trabalham melhor com imagens quadradas. Como os conjuntos de dados *db1* e *db2* são de tamanhos diferentes, todas as imagens dos dois conjuntos de dados foram redimensionadas para um tamanho padrão de  $186 \times 186$  *pixels*. Essas imagens são utilizadas como imagens de entrada dos modelos de CNNs. Para as redes pré-treinadas, o tamanho das mesmas foi ajustado para  $224 \times 224$  *pixels*.

#### 4.2.2 Aumento de dados

Com o objetivo de evitar sobre ajustamento, utilizamos a técnica de aumento de dados (MOLCHANOV et al., 2015; TALUKDAR; BISWAS; GUPTA, 2018). Nesta tese, o aumento de dados foi empregado usando a técnica de aumento tempo real, também conhecida como *on-the-fly*. Essa técnica realiza manipulações aleatórias das imagens dos minilotes de treinamento, por meio de transformações nas imagens, sem essencialmente aumentar o tamanho real do conjunto de dados (MOLCHANOV et al., 2015).

Em nossa proposta, as transformações aplicadas foram rotação de  $\pm 20^\circ$ , translação horizontal e vertical de  $\pm 3$  *pixels* e uma redução do tamanho da cena da imagem, com um fator randômico de entre 0,5 e 1,0, obtendo-se um conjunto de dados aumentados de 10.132 imagens. A Figura 4.5 ilustra essas transformações.

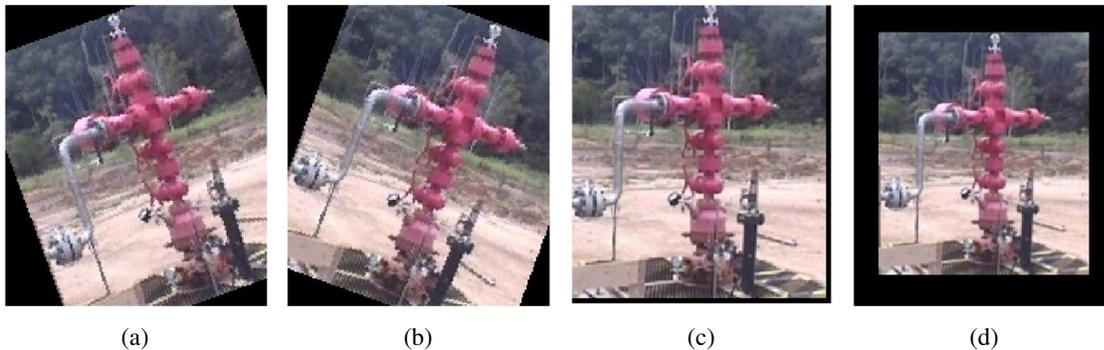


Figura 4.5 - Visualização de exemplos de transformações aleatórias aplicadas a quatro imagens *db1*: (a) rotação de  $+20^\circ$ ; (b) rotação de  $-20^\circ$ ; (c) translação de  $\pm 3$  *pixels*; (d) dimensionamento da cena com fator 0,8. Fonte: Compilação do autor.

#### 4.2.3 Arquitetura das CNN e parâmetros de treinamento

Para efetuar a classificação binária de uma imagem, se a classe da mesma é com vazamento de GN ou sem vazamento de GN, são investigadas as seguintes CNN:

- a) três propostas de arquiteturas de CNN, associadas a 3 métodos de otimização e a três técnicas de regularização. Essas propostas são descritas na Seção 4.2.3.1.
- b) quatro arquiteturas de CNN, anteriormente propostas na literatura, pré-treinadas utilizando a técnica de TL. Essas arquiteturas são descritas na Seção 4.2.3.2.
- c) uma proposta de arquitetura CNN, baseada na fusão de quatro CNN pré-treinadas, utilizando a técnica de fusão de característica (*Feature Fusion*) (KHAN et al., 2019; OZKAYA; OZTURK; BARSTUGAN, 2020; ÖZYURT, 2020). Essa proposta é descrita na Seção 4.2.3.3.

#### 4.2.3.1 Propostas de arquiteturas de CNN

Segundo Aggarwal (2018), o aumento do número de unidades ocultas na CNN aumenta a complexidade da função modelada, o que pode ser útil para modelar problemas difíceis. No entanto, isso pode causar sobreajustamento aos dados, se a arquitetura for complexa ou se o conjunto de dados for pequeno. Uma forma de diminuir a complexidade é aumentando o comprimento (número de filtros das camadas convolutivas) da rede e diminuindo o número de unidades ocultas em cada camada (AGGARWAL, 2018). Este procedimento foi usado por Miyagawa *et al.* (2019).

A primeira arquitetura CNN proposta neste projeto para detecção de vazamento de gás natural, denominada CNN Longa (CNNL), é uma arquitetura serial com 18 camadas convolucionais, baseada no trabalho de Miyagawa *et al.* (2019). A Figura 4.6 (a) ilustra essa arquitetura. De forma geral, todas as redes apresentam a seguinte sequência de camadas: convolutiva 3x3, seguida da camada *batch normalization* e ReLU. Essa sequência será, doravante, denominada de bloco convolutivo.

No caso do modelo CNNL, a extração de características é efetuada através de 18 blocos convolutivos, com camadas convolutivas com filtro 3x3, e quantidade de filtros variando entre 16 e 128, conforme mostrado na Figura 4.6 (a). Ao final de cada 3 blocos convolutivo há uma camada de agrupamento (*max pooling*), para realizar a subamostragem. Após os blocos convolutivos, há uma camada de *dropout*, para melhoria da generalização da rede, seguida por uma camada totalmente conectada, camada *softmax* e uma camada de classificação.

A largura, comprimento e profundidade da primeira camada tem as seguintes dimensões:  $186 \times 186 \times 16$ . À medida que o comprimento da rede aumenta, a largura e o comprimento das camadas são reduzidos, enquanto a profundidade aumenta. A última camada convolutiva é um mapa de características de dimensão  $11 \times 11 \times 128$ .

Para investigar se a CNNL apresenta sobre ajustamento aos dados, mesmo após esse afilamento das unidades nas camadas ocultas, são propostas outras duas arquiteturas, a CNN Média (CNNM) e a CNN Sintética (CNNS). Ambas essas arquiteturas possuem menos camadas convolucionais que a CNNL, porém os blocos convolutivos possuem as mesmas camadas do modelo CNNL.

A CNNM possui 6 blocos convolutivos. A cada três blocos convolutivos temos uma camada de subamostragem. Já a CNNS possui dois blocos convolutivos, ambos seguidos de uma camada de subamostragem. Essas novas arquiteturas são mostradas na Figura 4.6 (b) e Figura 4.6 (c), respectivamente.



Figura 4.6 - Arquiteturas de rede propostas: a) CNNL, com 18 camadas de convolução; b) CNNM, com 6 camadas de convolução; c) CNNS, com 2 camadas de convolução. Fonte: Adaptada de Miyagawa *et al.*(2019).

Nessas três arquiteturas (CNNL, CNNM e CNNS), com o objetivo de evitar os problemas de explosão e de desvanecimento do gradiente, cada camada de convolução é seguida por uma camada de normalização em lote (AGGARWAL, 2018). A função de ativação utilizada nas mesmas é a função ReLU.

Conforme mostrado na Figura 4.6, a fim de reduzir o tamanho espacial do mapa de características, à medida que a profundidade da rede aumenta, foi utilizada a camada de subamostragem com filtros de dimensão  $2 \times 2$ , aplicados com um passo de  $1 \times 1$ . Cada arquitetura possui uma etapa de classificação composta pelas seguintes camadas: camada totalmente conectada, uma camada *softmax* e uma camada de classificação final. A camada *softmax* gera uma distribuição de probabilidade para as duas classes, com vazamento e sem vazamento.

Para otimizar o desempenho das arquiteturas CNNL, CNNM e CNNS, avalia-se a aplicação de diferentes técnicas de regularização e diferentes métodos de otimização, conforme mostrado na Tabela 4.2.

Foram utilizadas as seguintes técnicas de regularização: *dropout* (SRIVASTAVA et al., 2014), que elimina, aleatoriamente, 30% dos neurônios de uma camada. Essa exclusão inibe o processo de retropropagação através dos neurônios eliminados (AGGARWAL, 2018) e  $L_2$  (CHEN et al., 2016); regularização  $L_2$ , com um coeficiente de suavização de 0,0001.

Os seguintes algoritmos de otimização foram avaliados: Estimativa do Momento Adaptativo (*Adaptive Moment Estimation* -ADAM), com taxas de decaimento  $\beta_1$  de 0,9 e  $\beta_2$  de 0,9 (KINGMA; BA, 2014); Propagação da Raiz Quadrática Média (*Root Mean Square Propagation* - RMSProp), com taxas de decaimento  $\beta_2$  de 0,9 (TIJMEN TIELEMAN; HINTON, 2012), Gradiente Estocástico Descendente com Momento (*Stochastic Gradient Descent with Momentum* - SGDM), com momento de 0,9 (HINTON, 2012).

Conforme mostrado na Tabela 4.2, a combinação das três arquiteturas com as duas técnicas de generalização e com os três métodos de otimização, resulta em vinte e sete modelos distintos. Esses modelos são treinados e validados com o conjunto de dados de imagens *db1\_trein* e *db1\_valid*, respectivamente, a serem definidos na Seção 4.2.4, denominada de Treinamento, validação e teste e hiperparâmetros de treinamento.

Tabela 4.2 - 27 modelos de CNN propostos.

Arquitetura	Otimização	Regularização	Modelo
LCNN	ADAM	<i>Dropout</i>	1
		L2	2
		L2 + <i>Dropout</i>	3
	RMSProp	<i>Dropout</i>	4
		L2	5
		L2 + <i>Dropout</i>	6
	SGDM	<i>Dropout</i>	7
		L2	8
		L2 + <i>Dropout</i>	9
MCNN	ADAM	<i>Dropout</i>	10
		L2	11
		L2 + <i>Dropout</i>	12
	RMSProp	<i>Dropout</i>	13
		L2	14
		L2 + <i>Dropout</i>	15
	SGDM	<i>Dropout</i>	16
		L2	17
		L2 + <i>Dropout</i>	18
SCNN	ADAM	<i>Dropout</i>	19
		L2	20
		L2 + <i>Dropout</i>	21
	RMSProp	<i>Dropout</i>	22
		L2	23
		L2 + <i>Dropout</i>	24
	SGDM	<i>Dropout</i>	25
		L2	26
		L2 + <i>Dropout</i>	27

Fonte: Compilação do autor.

#### 4.2.3.2 Quatro CNNs pré-treinadas utilizando a técnica TL

A transferência de conhecimento é usada para alavancar o conhecimento dos domínios de origem para ajudar a construir um novo domínio de destino, conforme discutido na Seção 3.3.

Neste trabalho, com o objetivo de classificar as imagens da cabeça de poço em duas classes, com e sem vazamento, foram avaliadas as seguintes CNN pré-treinadas: DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18. A seguir apresenta-se as mesmas.

- a) **DenseNet-201**: As redes convolucionais densamente conectadas (DenseNet), proposta por Huang *et al.* (2017), melhoram o fluxo de informações entre as camadas da rede, por meio da introdução de conexões diretas entre cada camada e as camadas

subsequentes. Os mapas de características das camadas anteriores são usados como entradas nas camadas subsequentes. A Figura 4.7 ilustra as conexões diretas de qualquer camada para todas as camadas subsequentes.

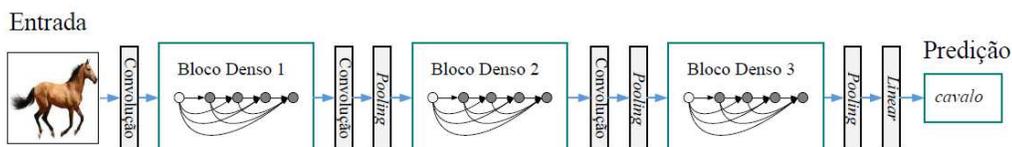


Figura 4.7 - Exemplo de um DenseNet com três blocos densos. As camadas entre dois blocos adjacentes são chamadas de camadas de transição e mudam os tamanhos do mapa de características, por meio de convolução e amostragem. Fonte: Adaptada e traduzida de (HUANG et al., 2017).

- b) **GoogLeNet:** A GoogLeNet, proposta por Szegedy *et al.* (2015), teve o melhor desempenho na ILSVRC14. Ela possui uma arquitetura composta por 22 camadas com cerca de 6,8 milhões de parâmetros. A novidade dessa arquitetura é a inserção de blocos denominados pelos autores de *Inception*, formados por vários filtros de diferentes tamanhos e que ajudam a reduzir o tamanho e melhorar o desempenho da rede. Após o último módulo de *Inception*, o inception5b, existem as seguintes camadas: amostragem por média global (*avg\_pool*), *dropout*, totalmente conectada, *softmax* e camada de classificação final. A Figura 4.8 ilustra um módulo do bloco *Inception*.

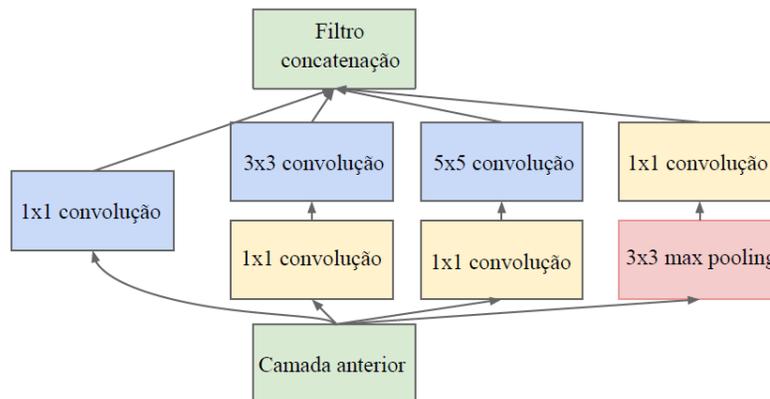


Figura 4.8 - Exemplo da arquitetura de um módulo *Inception* da CNN GoogLeNet. Fonte: Adaptada e traduzida de Szegedy *et al.* (2015).

- c) **ResNet-18:** ResNet-18 é uma acrônimo para *Residual Network*, proposta por He *et al.* (2016). Essa rede obteve excelentes resultados no ILSVRC15. Sua arquitetura é formada por 18 camadas. A novidade dessa arquitetura foi solucionar a dificuldade

em treinar redes tão profundas, já que o gradiente tende a desaparecer ou explodir no treinamento de redes profundas. A solução dada pelos autores é inserir blocos residuais, como o ilustrado na Figura 4.9, os quais são dotados de conexões de atalho, que ignoram uma ou mais camadas do bloco. A saída de conexões curtas é adicionada à saída dessas camadas empilhadas. Após o último bloco residual, *res5b*, existem as seguintes camadas: amostragem por média global (*avg\_pool*), totalmente conectada, *softmax* e camada de classificação.

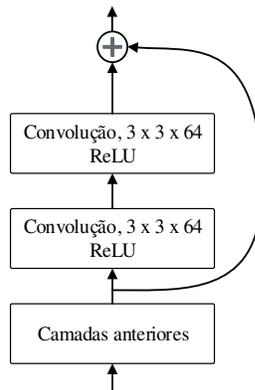


Figura 4.9 - Ideia original dos blocos residuais das CNN do tipo ResNet. Fonte: Adaptada e traduzida de He *et al.* (2016).

- d) **MobileNetV2:** Similar a ResNet-18, porém incorpora módulos derivados da MobileNet (HOWARD et al., 2017), que são constituídos por operações de convoluções separáveis em profundidade, denominada pelos autores de *Depthwise*. O conceito de convolução separada, conforme ilustrado na Figura 4.10, consiste em fatorar uma convolução padrão em uma convolução em profundidade associada a uma convolução 1x1, denominada de convolução pontual. A novidade do MobileNetV2 (SANDLER et al., 2018), quando comparada com sua antecessora, a MobileNetV1 (HOWARD et al., 2017), agrega duas novas funcionalidades: 1) gargalos lineares (*bottleneck*) entre as camadas, e; 2) conexões de atalho entre os gargalos.

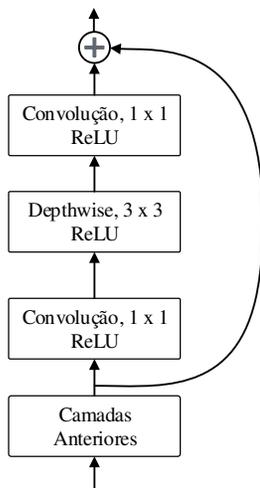


Figura 4.10 - Estrutura do bloco de construção do MobileNetV2. Fonte: Adaptada e traduzida de Sandler *et al.* (2018).

O Quadro 4.1 apresenta uma descrição das arquiteturas DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18, mostrando os tamanhos das saídas das camadas intermediárias. As camadas marcadas em tons de cinza em cada arquitetura são usadas como camadas de entrada para o módulo de concatenação, que será discutido na próxima seção, quando será apresentada uma nova arquitetura de rede, que propõe a fusão de características dessas redes pré-treinadas.

Quadro 4.1 - Descrição das arquiteturas DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18.

<b>DenseNet-201</b> (input 224x224x3)		<b>GoogLeNet</b> (input 224x224x3)	
Layer/Block	Activations	Layer/Block	Activations
conv1	112x112x64	conv1	112x112x64
dense_block_1	56x56x224	conv2	56x56x192
transition layer_1	56x56x128	inception3a	28x28x256
dense_block_2	28x28x512	inception3b	28x28x480
transition layer_2	28x28x256	inception4a...4c	14x14x512
dense_block_3	14x14x1792	inception4d	14x14x528
transition layer_3	14x14x896	inception4e	14x14x832
dense_block_4	7x7x1920	inception5a	7x7x832
avg_pool	1x1920	inception5b	7x7x1024
fully connected	1x2*	avg_pool	1x1024
softmax	1x2*	fully connected	1x2*
Classification(b)		softmax	1x2*
		Classification(c)	

(a)

(b)

MobileNetV2 (input 224x224x3)		ResNet-18 (input 224x224x3)	
Layer/Block	Activations	Layer/Block	Activations
conv1	112x112x32	conv1	112x112x64
bottleneck1	112x112x32	res2a...2b	56x56x64
bottleneck2	112x112x16	res3a...3b	28x28x128
bottleneck3	56x56x24	res4a...4b	14x14x256
bottleneck4	28x28x32	res5a...5b	7x7x512
bottleneck5	14x14x64	avg_pool	1x512
bottleneck5	14x14x96	fully connected	1x2*
bottleneck5	7x7x160	softmax	1x2*
conv2	7x7x1280	Classification(e)	
avg_pool	1x1x1280		
fully connected	1x1x2*		
softmax	1x1x2*		
Classification(d)			

(c) (d)

Fonte: Compilação do autor.

Usando TL, essas redes são treinadas, validadas e testadas com os conjuntos de dados *db1\_trein*, *db1\_valid* e *db1\_test*, respectivamente, para a classificação das imagens de cabeça de poço nas classes com e sem vazamento.

#### 4.2.3.3 CNN FF baseada em fusão de características

Visando aumentar o desempenho das redes pré-treinadas, foi explorada a combinação das mesmas através da fusão de características, resultando na proposta de uma nova arquitetura, uma das contribuições desse trabalho. Essa arquitetura foi denominada de FF (*Feature Fusion*).

A fusão de características já foi aplicada a diversos domínios, tais como, classificação do câncer de mama (KHAN et al., 2019), estudo envolvendo detecção do coronavírus (OZKAYA; OZTURK; BARSTUGAN, 2020), teste de leucócitos (ÖZYURT, 2020).

Nesta pesquisa, as CNN DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18 são treinadas, validadas e testadas individualmente, conforme mostrado no diagrama em blocos da Figura 4.4. Na arquitetura fundida, as camadas de classificação das mesmas são removidas, e a fusão das características extraídas por cada uma delas é feita no nível da camada de concatenação, marcadas em tons de cinza no Quadro 1 apresentado na seção anterior. A Figura 4.11 apresenta um diagrama em blocos da rede com a fusão de características proposta nesse trabalho, denotada por CNN FF.

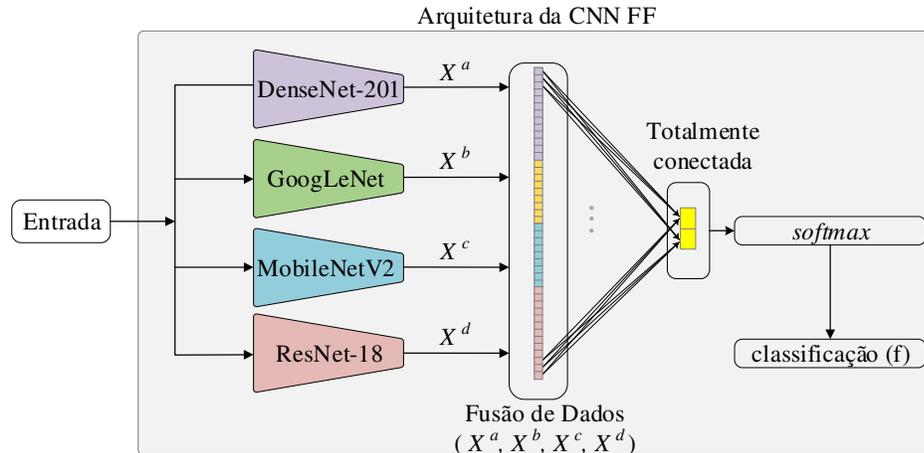


Figura 4.11 - Arquitetura da CNN FF proposta, constituída pela fusão de característica das seguintes redes CNN pré-treinadas: DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18. Fonte: Compilação do autor.

Espera-se que a CNN FF, ao agregar características de quatro redes pré-treinadas, apresente um melhor desempenho que os alcançados pelas redes individuais.

A seguir detalha-se as entradas do módulo de concatenação mostrado na Figura 4.11, provenientes das quatro redes pré-treinadas:

- Última camada de amostragem (*avg\_pool*), denominada de  $X^a$ , da rede DenseNet-20, cuja dimensão é  $1 \times 1.920$ ;
- Última camada de amostragem (*avg\_pool*), denominada de  $X^b$ , da rede GoogLeNet, cuja dimensão é  $1 \times 1.024$ ;
- Última camada de amostragem (*avg\_pool*), da rede MobileNetV2, denominada de  $X^c$ , cuja dimensão é  $1 \times 1.280$ ;
- Última camada de amostragem (*avg\_pool*), da rede ResNet-18, denominada de  $X^d$ , cuja dimensão é  $1 \times 512$ .

As características obtidos  $X^a$ ,  $X^b$ ,  $X^c$  e  $X^d$  são concatenadas, formando uma única camada, denominada de  $Y$ , em que  $Y = f^{\text{cat}}(X^a, X^b, X^c, X^d)$ , como indicado na Figura 4.12. Este tipo de fusão é definida por S. Zhou *et al.* (2020), como fusão de características em cascata.

Após a camada  $Y$ , as seguintes camadas são adicionadas à arquitetura da rede FF: totalmente conectada, *softmax* e classificação, conforme mostrado na Figura 4.11.

Um novo procedimento de treinamento, validação e teste é então realizado com esta nova arquitetura FF, conforme mostrado no diagrama em blocos da Figura 4.4.

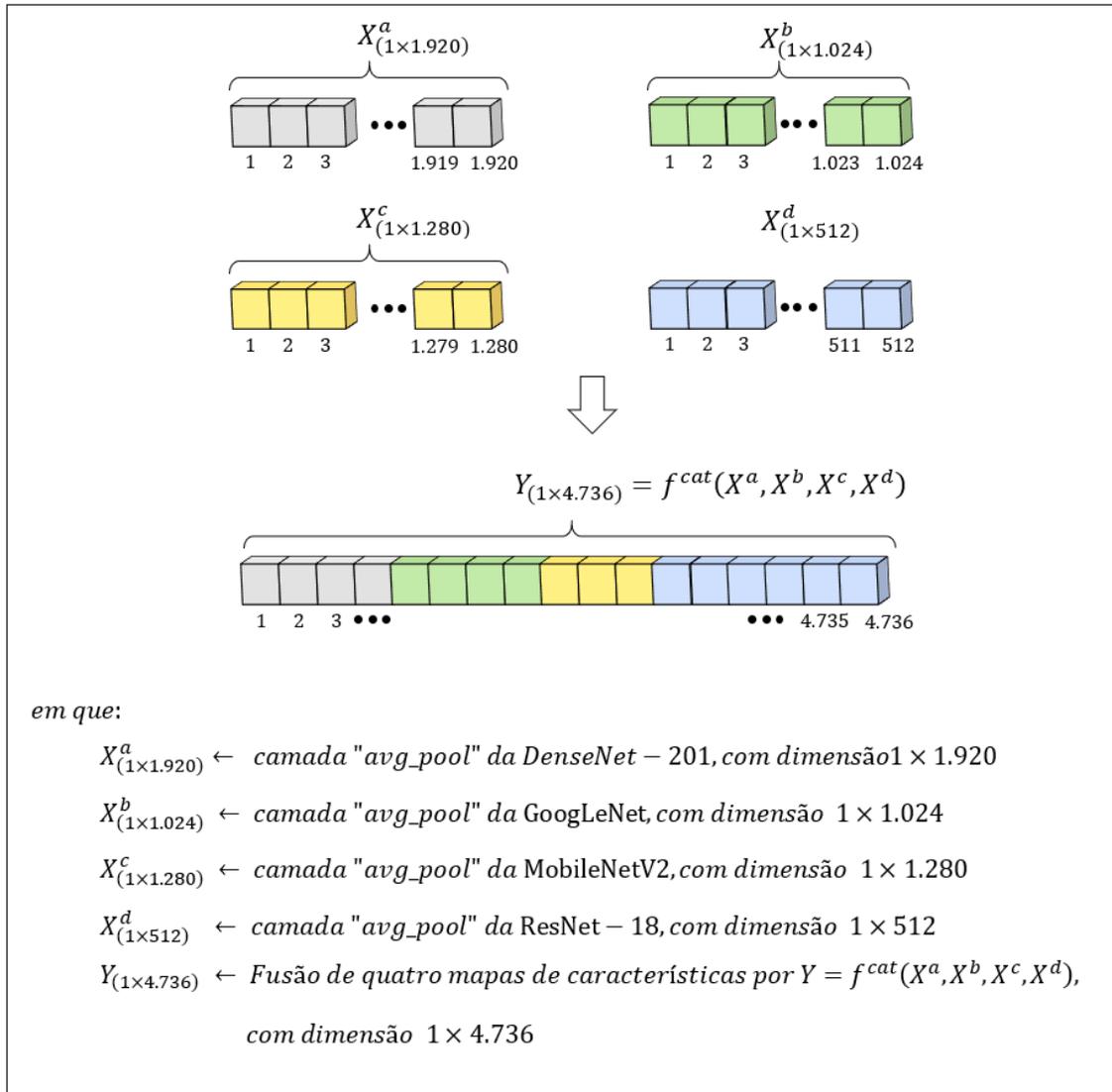


Figura 4.12 - Método proposto de fusão de característica das CNN DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18. Fonte: Compilação do autor.

#### 4.2.4 Treinamento, validação e teste e hiperparâmetros de treinamento

O conjunto de dados de imagem *db1* é dividido na seguinte proporção: 70% treinamento (*db1\_trein*), 15% validação (*db1\_valid*) e 15% teste (*db1\_test*). O conjunto de dados *db2* é usado apenas para teste.

O método usado para treinamento, validação e teste é mostrado na Figura 4.13. Conforme demonstrado na Figura 4.13 (a), os 27 modelos apresentados na Tabela 4.2 são treinados com o *db1\_trein* e avaliados com o conjunto de dados *db1\_valid*. O conjunto de avaliação é usado para escolher o Modelo com Melhor Desempenho (MMD) entre os vinte e sete modelos. Após determinado o MMD, ele é testado com *db1\_test*, conforme mostrado na Figura 4.13 (b).

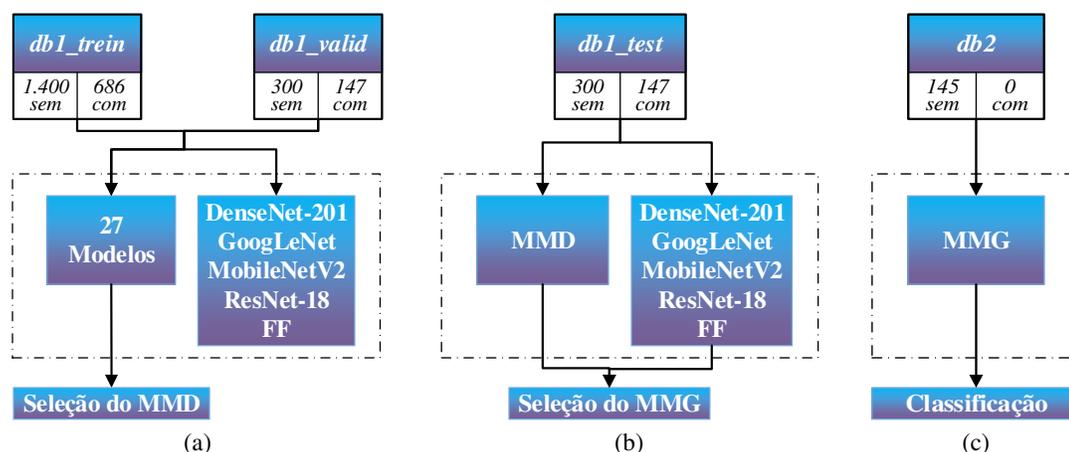


Figura 4.13 - Diagrama de blocos para treinamento, avaliação e teste: (a) Treinamento e validação dos 27 modelos de CNN para obtenção do modelo MMD, treinamento das 4 CNNs pré-treinadas, treinamento da rede FF; (b) Teste do MMD, das 4 CNNs pré-treinadas e da rede FF com o conjunto de dados *db1\_test* e obtenção do Melhor Modelo Geral (MMG); (c) Teste do MMG com o conjunto de dados *db2*. Fonte: Compilação do autor.

Determinado o MMD, os conjuntos de dados *db1\_trein*, *db1\_valid* e *db1\_test* são usadas para treinar, validar e testar os quatro modelos de pré-treinadas DenseNet-201, GoogLeNet MobileNetV2 e ResNet-18, através da técnica de TL indicada na Seção 3.3 (*Transfer Learning*).

Com as CNNs DenseNet-201, GoogLeNet MobileNetV2 e ResNet-18, ajustadas e testadas para tarefa de detecção de vazamento de GN, a CNN-FF é construída combinando as características destas quatro redes, como definido na Seção 4.2.3.3.

Como o conjunto de dados de imagem *db2* não faz parte dos conjuntos de treinamento (*db1\_trein*) ou validação (*db1\_valid*) ou teste (*db1\_test*), ele permitir avaliar o comportamento dos modelos CNN desenvolvidos quanto a distinção entre nuvens de GN e nuvens no céu, uma vez que há similaridade semântica entre elas.

Outros hiperparâmetros adotados para o treinamento supervisionado das CNN são fixados da seguinte maneira. Como neste estudo há CNNs complexas, como por exemplo a CNN FF, um número elevado de épocas de treinamento requeria um longo tempo de treinamento, portanto definiu-se que todas as CNN são treinadas por até 15 épocas. Em virtude do custo computacional, definiu-se o tamanho do minilote como 16. E por sugestão de Goodfellow (2016) e Aggarwal (2016), definiu-se a taxa de aprendizado inicial em 0,001. Finalmente, para a parada do treinamento e validação, usamos a parada antecipada, com uma paciência de validação igual a 4. Isso significa que o treinamento é interrompido quando a perda no conjunto de validação aumenta em quatro mini lote consecutivos (KOFER; MUHR; SPÖCK, 2019; REM et al., 2018). A Tabela 4.3 resume os principais parâmetros e hiperparâmetros adotados no desenvolvimento, treinamento, validação e teste das CNNs.

Tabela 4.3 - Parâmetros e hiperparâmetros adotados no desenvolvimento, treinamento, validação e teste das CNNs.

Divisão do Conjunto de Dados	N.º de épocas	Taxa de aprendizado	Tamanho do mini lote	<i>Dropout</i>	Parada Antecipada	ADAM taxas de decaimento	RMSProp taxas de decaimento	SGDM Momento
70% treinamento 15% validação 15% teste	15	0,001	16	0,3	4	$\beta_1$ de 0,9 $\beta_2$ de 0,9	$\beta_2$ de 0,9	0,9

Fonte: Compilação do autor.

#### 4.2.5 Avaliação das CNN: métricas de desempenho

As métricas utilizadas na validação, teste e comparação dos modelos de CNN foram as seguintes: Acurácia, Taxa de Falsos Negativos, Precisão, Sensibilidade, *F1-Score*. Para verificar a superioridade do desempenho de uma arquitetura de rede sobre a outra, foi utilizado o método do *Chi-quadrado*. Para apresentar essas métricas são definidos previamente os seguintes termos:

- Verdadeiros Positivos (VP): imagens com vazamentos;
- Verdadeiros Negativos (VN): imagens sem vazamentos;
- Falsos Positivos (FP): imagens sem vazamentos de GN que a CNN classificou como “com vazamentos de gás natural”;
- Falsos Negativos (FN): imagens com vazamentos de GN que a CNN classificou como “sem vazamentos de gás natural”.

Através destes termos, foram produzidas matrizes de confusão (Tabela 4.4), que são uma tabulação cruzada das classes de observações previstas e reais do modelo usadas para comparar quais classes estão previstas corretamente.

Tabela 4.4 – Modelo da matriz de confusão.

	Classe prevista	
	com vazamento	sem vazamento
Classe real		
com vazamento	VP	FN
sem vazamento	FP	VN

Fonte: Compilação do autor.

A Acurácia (ACC) é utilizada como a principal métrica para definir o Modelo com Melhor Desempenho (MMD) entre os 27 modelos baseado nas CNNL, CNNM e CNNS.

A Acurácia também é utilizada para comparar o desempenho entre as CNNs MMD, DenseNet-201, GoogLeNet, MobileNetV2, ResNet-18 e FF, sendo então determinado o Melhor Modelo Geral (MMG). O MMG é o modelo, entre as CNNs avaliadas, que melhor prediz a ocorrência ou não de um vazamento GN.

Portanto, a Acurácia é obtida nas etapas de validação com o conjunto de dados *db1\_valid* e de testes com o *db1\_test* e *db2*, sendo definida na equação (17).

$$ACC = \frac{VP + VN}{VP + VN + FP + FN} \quad (17)$$

Em muitos sistemas de aprendizado de máquina, é crucial minimizar a taxa de falsos negativos (KERMANY et al., 2018; RUEHLE, 2019; SONG; WOO; KIM, 2020), como é o caso dos diagnósticos médicos, onde um diagnóstico tardio ou equivocado pode reduzir a eficácia do tratamento (KERMANY et al., 2018; UKIL et al., 2016; XIE et al., 2019). Em veículos autônomos, uma pequena falha pode causar um mau funcionamento momentâneo no veículo, tornando-se uma ameaça (SONG; WOO; KIM, 2020). Além disso, no monitoramento das condições das estradas, uma situação perigosa aparece quando um sistema de navegação indica erroneamente ao usuário rotas mais curtas e sem restrições, levando-o a uma zona de risco (HAMEED; MAZHAR; HASSAN, 2018). Essas circunstâncias também estão na indústria do petróleo: uma indicação de um vazamento falso negativo pode levar a acidentes, uma vez que o sistema classifica erroneamente como sem vazamento, um vazamento real. Por isso, neste trabalho, o escore TFN é uma medida essencial, sendo a segunda variável mais importante do estudo.

A TFN (Taxa de Falsos Negativos) é definida como a razão entre o número de amostras positivas classificadas incorretamente (FN) e o número de amostras positivas (FN + VP), de acordo com a equação (18).

$$TFN = \frac{FN}{FN + VP} \quad (18)$$

Outras métricas também são usadas para avaliar o desempenho das arquiteturas, como:

- 1) Precisão (SOKOLOVA; LAPALME, 2009), mostrada na equação (19). Essa métrica expressa a razão entre o número de amostras positivas que foram classificadas corretamente e o número de amostras positivas (VP+FP). Assim, expressa a capacidade do modelo de evitar falsos positivos.

$$Precisão = \frac{VP}{VP + FP} \quad (19)$$

- 2) *Recall* ou sensibilidade (SOKOLOVA; LAPALME, 2009), mostrada na equação (20). Essa métrica determina a taxa de verdadeiros positivos que o classificador apresenta.

$$Sensibilidade = \frac{VP}{VP + FN} \quad (20)$$

- 3) *F1-Score* (VAN RIJSBERGEN, 1974), mostrada na equação (21). Essa métrica, com valores entre 0 e 1, representa uma média harmônica entre a precisão e a sensibilidade.

$$F1 - Score = \frac{2VP}{2VP + FP + FN} \quad (21)$$

Neste trabalho, há o interesse também em responder se o desempenho das CNNs é melhor que o desempenho obtido no trabalho anterior de Melo, Costa Filho e Costa (2014), ou seja, avalia-se que a hipótese nula de que os resultados entre as CNNs e trabalho anterior serem semelhantes é rejeitada (KING; ECKERSLEY, 2019). Para avaliar essa hipótese nula é utilizado o teste do *Chi-quadrado* ( $\chi^2$ ) (HOWELL, 2010), calculado a partir da tabela de contingência  $2 \times 2$ , mostrada na Tabela 4.5. Essa última contém a quantidade de acerto e erros de dois classificadores, no caso, o classificador baseado em CNN, proposto nesse trabalho, e o classificador baseado no filtro de novidade, proposto no trabalho anterior.

Tabela 4.5 - Tabela de contingência usada para calcular  $\chi^2$ .

<b>Classificador: CNN</b>	<b>Classificador: Filtro de Novidade</b>
Número de classificações corretas	Número de classificações corretas
Número de classificações erradas	Número de classificações erradas

Fonte: Compilação do autor.

Com os dados da Tabela 4.5, o teste de  $\chi^2$  é obtido conforme a equação (22) (HOWELL, 2011; KING; ECKERSLEY, 2019)

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (22)$$

Em que:

$O_{ij}$  = número de casos observados na linha  $i$  da coluna  $j$ ;

$E_{ij}$  = número de casos esperados, sob  $H_0$ , na linha  $i$  da coluna  $j$ ;

$\Sigma_{i=1}^r$  e  $\Sigma_{j=1}^k$  indica somatório sobre todas as linhas  $r$  e todas as colunas  $k$  da tabela de contingência.  $E_{ij}$  pode ainda ser definido pela equação (23).

$$E_{ij} = \frac{R_i \times C_j}{N} \quad (23)$$

Em que:

$R_i$  é o número de todos os casos observados na linha  $i$ ;

$C_j$  é o número de todos os casos observados na coluna  $j$ ;

$N$  é o número de todos os casos da tabela de contingência.

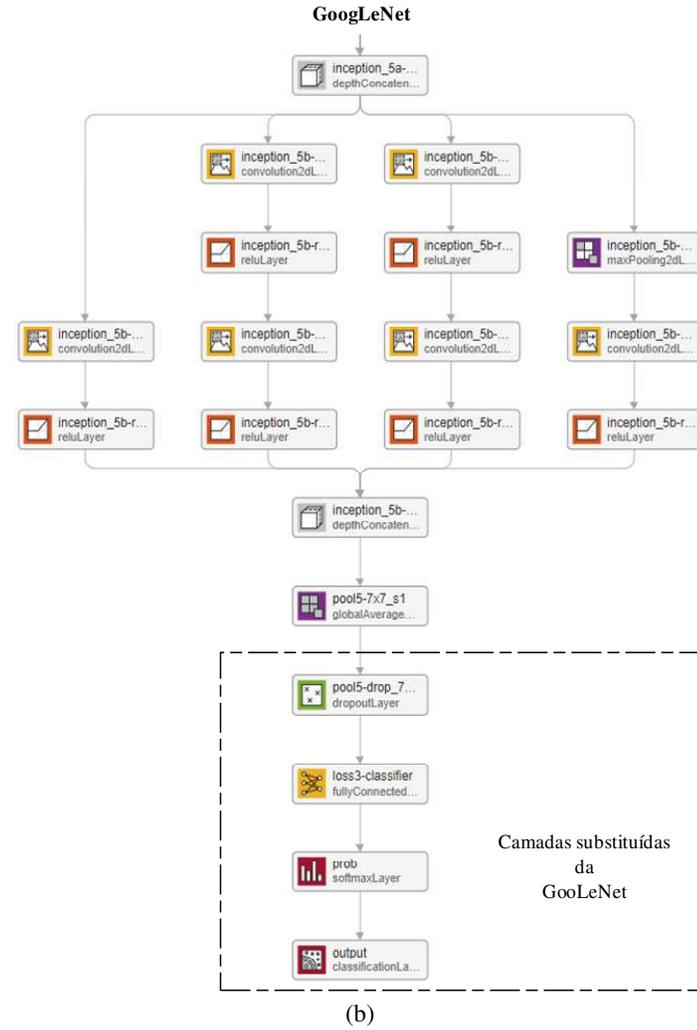
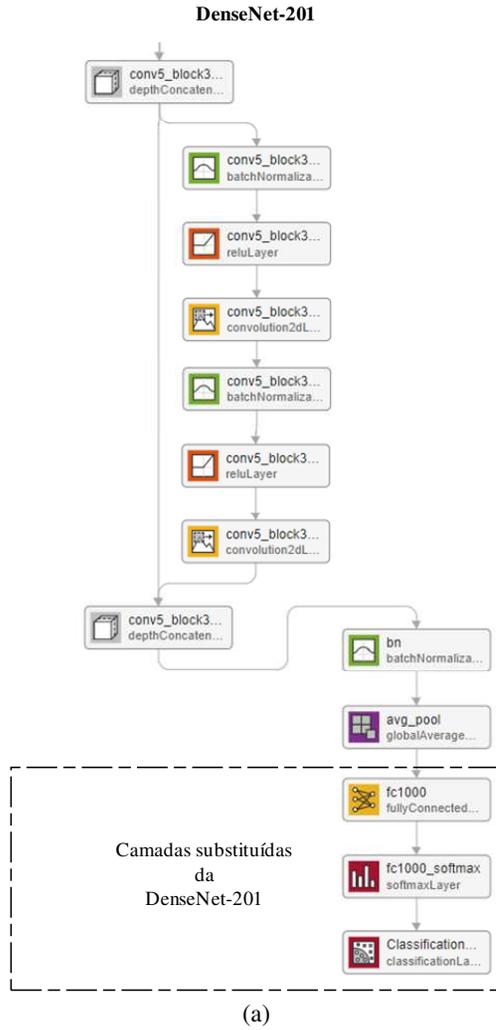
Para concluir sobre a superioridade do resultado apresentado neste trabalho, adotamos um nível de significância de 0,01. A hipótese nula ( $H_0$ ), indica que não existe diferença entre os dois métodos.

Considerando que o conjunto de dados utilizado para comparação entre os dois métodos possui  $N = 447$  amostras (*db1\_valid* e *db1\_test*) e com grau de liberdade 1, com o nível de significância adotado temos que o valor crítico é de  $\chi_{critical}^2 = 6,63$ . Portanto, se o valor de  $\chi_{calculado}^2 > \chi_{critical}^2$ , a hipótese nula é rejeitada.

#### 4.2.6 Transferência de conhecimento

Nesta pesquisa, o método de transferência de conhecimento foi aplicado a quatro modelos de CNN pré-treinadas: DenseNet-201 (HUANG et al., 2017), GoogLeNet (SZEGEDY et al., 2015), MobileNetV2 (SANDLER et al., 2018) e ResNet-18 (HE et al., 2016). A Figura 4.14 ilustra as camadas que são substituídas por novas camadas nas referidas redes pré-treinadas. As camadas remanescentes têm seus parâmetros congelados e, conforme detalhado na Seção 3.3, o treinamento ocorre apenas para ajuste dos parâmetros das novas camadas que são inseridas.

Além dos quatro modelos indicados na Figura 4.14, a fusão das características das redes gera o novo modelo FF. O treinamento do modelo FF também utilizou transferência de conhecimento.





#### 4.2.7 Realce da área com vazamento

Como já descrito, para atender os objetivos propostos na introdução, este estudo é composto por três fases distintas:

- i)* A primeira fase é o desenvolvimento de CNNs com a tarefa de realizar a predição de imagens da cabeça do poço em duas classes: com e sem vazamento de GN;
- ii)* A segunda fase é comparar a capacidade de predição dos classificadores CNN com o classificador Filtro de Novidade, desenvolvido em trabalho anterior;
- iii)* A terceira fase é o realce da região da imagem da cabeça do poço que contribuiu para que o classificador CNN tenha classificada uma determinada imagem como pertencente a classe “com vazamento”.

A metodologia para alcançar as fases *i)* e *ii)* foram descritas nas seções anteriores desta tese. O método para fase *iii)* é descrito de forma resumida a seguir.

Nesta terceira fase, o realce da região com vazamento de GN é realizado através de um mapa de calor que destaca as regiões específicas da imagem que contribuíram para a previsão positiva da CNN. O mapa de calor é obtido através do algoritmo Grad-CAM (*Gradient-Weighted Class Activation Mapping*) proposto por Selvaraju *et al.* (2017). Esse método, que calcula os gradientes entre a última camada convolucional e a saída de classe desejada, foi detalhado na Seção 3.4 (Grad-CAM).

Nesta tese o Grad-CAM é aplicado ao MMG para realçar a região da imagem que contribuiu para uma predição positiva do classificador MMG.

### 4.3 Conclusão

Este capítulo descreveu os materiais e as metodologias aplicados no desenvolvimento dos vinte e sete modelos de CNN, desenvolvidos do zero (*scratch*), dos quatro modelos pré-treinados e do modelo construído através de fusão de características.

Procurou-se detalhar todos os bancos de dados utilizados, as arquiteturas de CNNs propostas para detecção de vazamento e GN, as métricas a serem utilizadas para avaliação do desempenho das CNNs e o método de realce das regiões da imagem de um vazamento que contribui para uma resposta positiva da CNN.

Assim, o capítulo inicia apresentando a câmera utilizada para capturar as imagens do poço. Estas imagens constituem a base de dados *db1*. Em seguida, são detalhados os conjuntos de dados *db1* e *db2* e os métodos de pré-processamento de imagens e aumento de dados aplicados aos conjuntos de dados. Na continuidade, as regras de treinamento, validação e teste, bem como os hiperparâmetros das redes CNNs são definidos, bem como as métricas que são

utilizadas para avaliar o desempenho das mesmas. Quanto às métricas, o capítulo elege a acurácia e a taxa de falsos negativos como primordiais para determinação do desempenho das CNN. Finalizando o capítulo, joga um pouco mais de luz sobre as técnicas de transferência de conhecimento e sobre o realce de regiões na imagem de vazamento através do algoritmo Grad-CAM.

## 5. RESULTADOS E DISCUSSÃO

Neste capítulo, são apresentados e discutidos os resultados das CNN propostas nesta tese para classificação binária de uma imagem em uma entre duas classes: classe com vazamento de GN ou classe sem vazamento de GN.

Inicialmente, investiga-se o desempenho dos diferentes modelos de CNN (vinte e sete modelos), que foram obtidos através da combinação de três arquiteturas (CNNL, CNNM e CNNS) com duas técnicas de generalização, e com três métodos de otimização, conforme descrito no capítulo anterior. Essa combinação fornece vinte e sete modelos distintos de CNN. A partir deles, obtém-se o modelo com melhor desempenho (MMD). Em seguida é analisado o desempenho das quatro CNN pré-treinadas e da CNN FF, cujos resultados são apresentados na Seção 5.1. Portanto, após a fase de treinamento e validação temos seis redes:

- i) CNN MMD;
- ii) DenseNet-201;
- iii) GoogLeNet;
- iv) MobileNetV2;
- v) ResNet-18;
- vi) CNN FF.

Esses seis modelos são testados com o conjunto de dados de teste (*db1\_test*). Nesta etapa de teste, define-se o modelo de CNN com o maior poder de predição de um vazamento de GN. Esse modelo é então denominado de Melhor Modelo Geral (CNN MMG).

Na terceira etapa, Seção 5.3, a análise dos resultados se expande através da comparação da capacidade de predição da CNN MMG com o classificador que utiliza o Filtro de Novidade, desenvolvido em trabalho anterior.

Na quarta etapa, Seção 5.4, a CNN MMG é testada com o conjunto de dados *db2*. Nesta seção, avalia-se a capacidade de generalização da CNN em classificar imagens de novos poços petrolíferos, uma vez que as imagens do *db2* não fizeram parte do conjunto de dados *db1*, conforme definido na seção de metodologia.

Por fim, apresenta-se o mapa de calor, extraído pelo algoritmo Grad-CAM, e discute-se a aplicabilidade do mesmo. Por meio do mapa de calor, é possível analisar a capacidade da CNN MMG em realçar a região da imagem da cabeça do poço que contribuiu para que o classificador tenha classificado uma imagem como verdadeiro positivo (VP) para vazamento de GN.

Essa introdução ao Capítulo 5 é sintetizada na Figura 5.1.

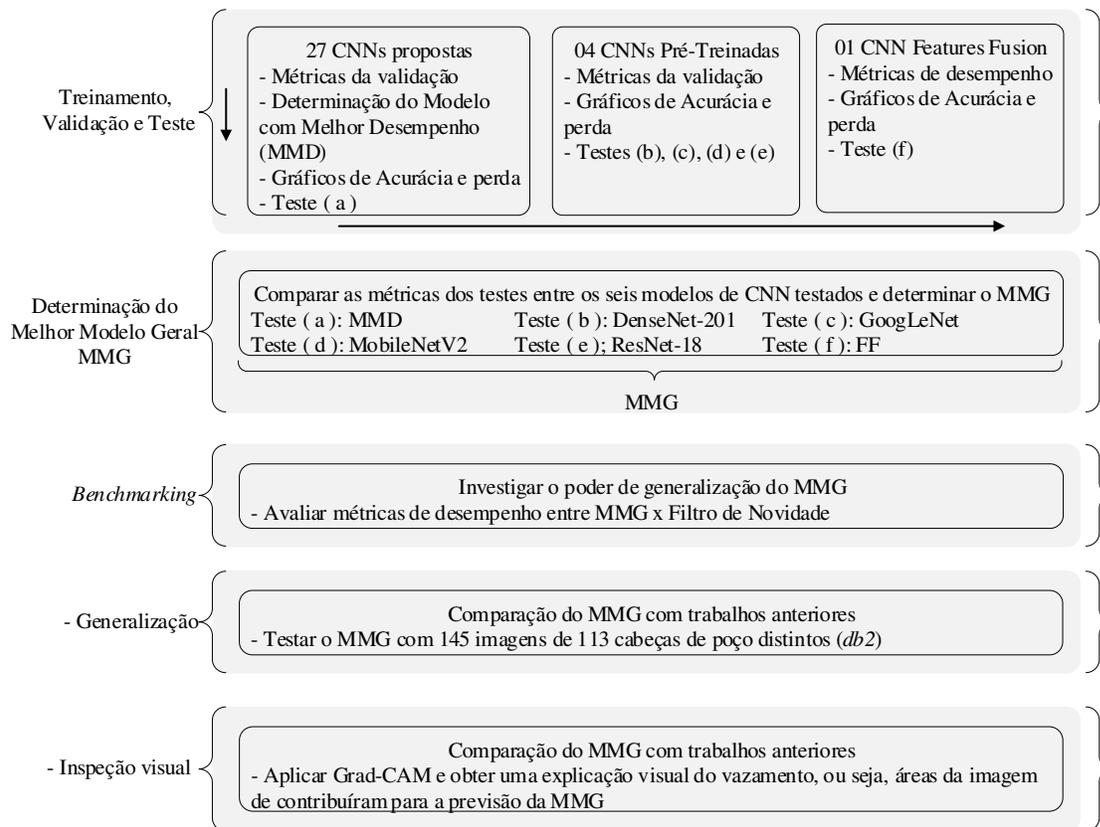


Figura 5.1 - Esquemático das etapas percorridas na seção de resultados e discussão. Fonte: Compilação do autor.

## 5.1 Resultados do treinamento e validação das CNN

Os resultados a seguir foram obtidos após treinamento de cada CNN utilizando o conjunto de dados *db1\_trein* e validados por meio do conjunto de dados *db1\_valid*.

### 5.1.1 Resultados dos 27 modelos de CNN propostos

Nesta seção, apresenta-se os valores das métricas de desempenho dos modelos obtidos através da combinação entre as três arquiteturas de CNN propostas, três métodos de otimização e três métodos para melhoria da generalização, que resultou em vinte e sete modelos de CNN.

A Tabela 5.1 apresenta o desempenho dos vinte e sete modelos de CNN propostos. Os valores das métricas, reportadas na Tabela 5.1, foram obtidos na etapa de validação utilizando o conjunto de dados *db1\_valid*.

O desempenho do modelo 7, marcado em tons de azul na Tabela 5.1, foi construído com uma arquitetura longa (CNNL), algoritmo de otimização SGDM e regularização por *dropout*, destaca-se entre os 27 modelos investigados nesta seção.

Esse modelo obteve uma acurácia de 99,78% e uma TFN de 0%. Esse é o maior valor de acurácia obtido entre todos os vinte e sete modelos, assim como a taxa de TFN foi a menor entre esses modelos. Portanto, o modelo 7 é nosso Modelo de Melhor Desempenho (MMD), conforme discutido a seguir.

Tabela 5.1 - Métricas de desempenho dos 27 modelos CNN obtidas através do conjunto de dados *db1\_valid*: acurácia (ACC), Taxa de Falsos Negativos (TFN), Sensibilidade, F1-Score, Chi-quadrado ( $\chi^2$ ) e tempo do treinamento e validação.

Arquitetura	Otimização	Regularização	Modelo	ACC (%)	TFN (%)	Precisão (%)	Sensibilidade (%)	F1-Score (%)	$\chi^2$	Tempo de Treinamento e validação (min:sec)
LCNN	ADAM	Dropout	1	99,33	2,04	100,00	97,96	98,97	25,57	01:22
		L2	2	99,55	1,36	100,00	98,64	99,32	27,70	03:39
		L2 + Dropout	3	99,55	1,36	100,00	98,64	99,32	27,70	03:34
	RMSProp	Dropout	4	76,06	72,79	100,00	27,21	42,78	104,35	03:15
		L2	5	99,55	1,36	100,00	98,64	99,32	27,70	03:15
		L2 + Dropout	6	99,55	1,36	100,00	98,64	99,32	27,70	03:15
	SGDM	Dropout	<b>7</b>	<b>99,78</b>	<b>0,00</b>	<b>99,32</b>	<b>100,00</b>	<b>99,66</b>	<b>29,95</b>	<b>02:50</b>
		L2	8	99,55	1,36	100,00	98,64	99,32	27,70	01:18
		L2 + Dropout	9	99,55	1,36	100,00	98,64	99,32	27,70	00:57
MCNN	ADAM	Dropout	10	97,32	8,16	100,00	91,84	95,74	10,75	02:28
		L2	11	99,55	1,36	100,00	98,64	99,32	27,70	02:34
		L2 + Dropout	12	97,32	8,16	100,00	91,84	95,74	10,75	02:27
	RMSProp	Dropout	13	95,53	13,61	100,00	86,39	92,70	3,28	02:26
		L2	14	97,32	8,16	100,00	91,84	95,74	10,75	02:26
		L2 + Dropout	15	92,17	23,81	100,00	76,19	86,49	0,51	02:16
	SGDM	Dropout	16	99,55	1,36	100,00	98,64	99,32	27,70	02:21
		L2	17	97,32	8,16	100,00	91,84	95,74	10,75	01:24
		L2 + Dropout	18	95,08	14,97	100,00	85,03	91,91	2,12	02:20
SCNN	ADAM	Dropout	19	97,32	8,16	100,00	91,84	95,74	10,75	00:48
		L2	20	97,32	8,16	100,00	91,84	95,74	10,75	00:42
		L2 + Dropout	21	98,88	3,40	100,00	96,60	98,27	21,61	01:09
	RMSProp	Dropout	22	99,55	1,36	100,00	98,64	99,32	27,70	01:15
		L2	23	99,33	2,04	100,00	97,96	98,97	25,57	01:09
		L2 + Dropout	24	98,66	4,08	100,00	95,92	97,92	19,78	01:06
	SGDM	Dropout	25	94,86	15,65	100,00	84,35	91,51	1,65	01:02
		L2	26	99,55	1,36	100,00	98,64	99,32	27,70	01:22
		L2 + Dropout	27	99,55	1,36	100,00	98,64	99,32	27,70	01:07

Fonte: Compilação do autor.

Na Tabela 5.1, os modelos 2, 3, 5, 6, 8, 9, 11, 16, 22, 26 e 27, forneceram ACC = 99,55%, inferior a 99,78% do MMD. No entanto, apesar da acurácia obtida ser expressiva (99,55%), os experimentos com esses modelos resultaram em classificações do tipo falso negativo (FN), ou seja, seus TFN são diferentes de 0%, alcançando um valor de TFN = 1,36%.

Este tipo de erro pode colocar em risco uma instalação, pois classificar uma imagem verdadeira positiva como negativa, em vazamentos reais, significa negligenciar o potencial risco de um acidente.

Logo, o modelo 7, além de apresentar a maior acurácia (99,78%) entre todos os modelos, é o único livre de erros do tipo FN, ou seja,  $TFN = 0\%$ .

No decorrer da etapa de ajuste do modelo 7, não ocorreu parada antecipada durante as quinze épocas do treinamento e validação. O tempo necessário para realizar o treinamento e validação deste modelo foi de 2 minutos e 50 segundos, sendo o quarto menor custo computacional entre as nove CNN longas (CNL) investigadas.

Continuando a análise do modelo 7, quanto aos valores das demais métricas da Tabela 5.1, constata-se que a capacidade do MMD em inferir os verdadeiros positivos alcançou 100% de exatidão, conforme indica a métrica Sensibilidade. Devido à ocorrência de uma única classificação FP, a precisão, que mede a taxa de FP, ficou ligeiramente abaixo de 100%, o que não é problema, pois não incorre em risco de acidentes. Quanto à média harmônica entre precisão e sensibilidade, *F1-Score*, o modelo 7 também se destacou entre os demais modelos, atingindo 99,66%. A discussão a respeito do teste de significância estatística *Chi-quadrado* será postergada para a Seção 5.3, que trata da comparação entre o MMG e o método que utiliza o Filtro de Novidade, o Classificador de Novidade.

Os gráficos de acurácia e perda durante o treinamento e validação do modelo 7 são mostrados na Figura 5.2 (a) e Figura 5.2 (b), respectivamente. Nestes gráficos, observa-se um comportamento bastante instável desde o início até a oitava época do treinamento e validação, tornando-se suave após esta época. Ambas as curvas, no entanto, apesar do comportamento assintótico, apresentam picos esparsos de pequena magnitude, fato que força o ajuste do modelo até a décima quinta época.

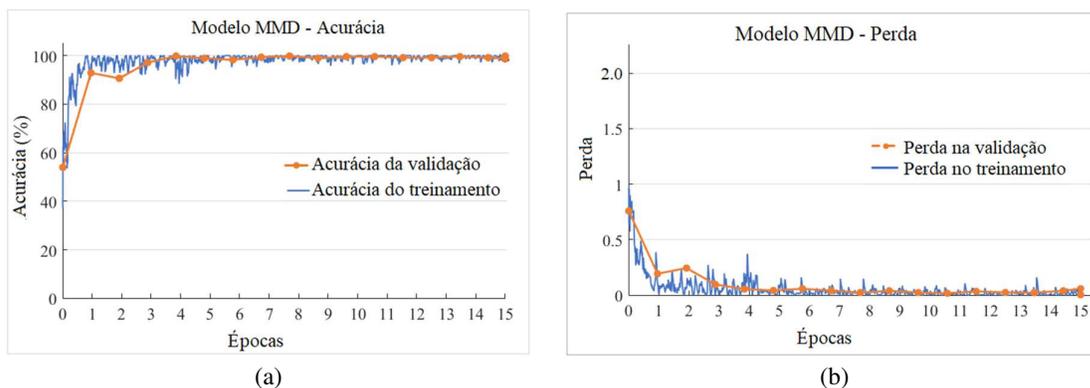


Figura 5.2 - Treinamento e validação do MMD: (a) acurácia durante o treinamento e validação; (b) perda durante o treinamento e validação. Fonte: Compilação do autor.

A classificação do conjunto de dados de teste (*dbl\_test*), submetido ao MMD, ou seja, Teste (a) indicado na Figura 4.4 e na Figura 5.1, resultou em uma acurácia igual àquela obtida no conjunto de dados *dbl\_valid*, ou seja, de  $ACC = 99,78\%$ , e em um caso de classificação do tipo FP. Tal circunstância indica que o modelo MMD não apresenta erro de sobreajustamento, uma vez que o *dbl\_test* é formado por imagens desconhecidas, até então, do MMD, pois essas imagens não fazem parte dos conjuntos de dados de treinamento (*dbl\_trein*) ou validação (*dbl\_valid*).

Considerando que os resultados de acurácia e TFN, foram os mesmos, tanto na fase de validação como na fase de teste, os resultados das métricas do Teste (a) do MMD, indicados na Tabela 5.2, são iguais aos apresentados na Tabela 5.1, os quais foram discutidos nos parágrafos anteriores.

Os valores indicados na Tabela 5.2 serão utilizados na etapa de seleção do MMG, que é discutido na Seção 5.2, que trata da comparação do desempenho dos seis modelos: CNN MMD, DenseNet-201, GoogLeNet, MobileNetV2, ResNet-18 e CNN FF.

Tabela 5.2 - Desempenho do Teste ( a ), referente ao modelo MMD, quando testado com o *dbl\_test*.

Modelo	Acurácia (%)	TFN (%)	Precisão (%)	Sensibilidade (%)	F1-Score (%)	$\chi^2$
CNN MMD (Model 7) Teste ( a )	99,78	0,00	99,32	100,00	99,66	29,95

Fonte: Compilação do autor.

### 5.1.2 Resultados das CNN pré-treinados e CNN FF

Nesta seção, apresenta-se e discute-se os valores das métricas de desempenho das quatro CNN pré-treinadas (DenseNet-201, GoogLeNet, MobileNetV2, e ResNet-18) e da CNN FF construída pela fusão de características destas quatro redes.

As arquiteturas dos vinte e sete modelos, apresentadas na Figura 4.6 e seus resultados experimentais discutidos na Seção 5.1.1, possuem arquiteturas similares, distinguindo-se uma da outra apenas pela profundidade, ou seja, pelo número de camadas.

Por outro lado, as CNN pré-treinadas têm arquiteturas diferentes uma da outra devido a diversos aspectos, tais como: camadas com funções específicas que só existem em determinada rede; a própria estruturação interna da arquitetura e suas interconexões; quantidade de camadas; etc. As diferenças entre as arquiteturas dos 27 modelos e as arquiteturas das CNN pré-treinadas,

motivam-nos a investigar sobre a capacidade das redes pré-treinadas na detecção de vazamento de GN.

Contudo, o que torna mais encorajador investigar as redes pré-treinadas, é a possibilidade de realizar a transferência de conhecimento destas redes, uma vez que elas foram treinadas originalmente com milhões de imagens. Os resultados destas CNN pré-treinadas, aplicadas na tarefa de detecção do vazamento de GN, são apresentados e discutidos a seguir.

A Tabela 5.3 apresenta o desempenho das quatro redes pré-treinadas (DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18) e do modelo FF, com suas respectivas métricas obtidas com o conjunto de dados *dbl\_valid*.

Tabela 5.3 - Métricas de desempenho das CNN DenseNet-201, GoogLeNet, MobileNetV2, ResNet-18 e FF, obtidas com o conjunto de dados *dbl\_valid*: acurácia (ACC), Taxa de Falsos Negativos (TFN), Sensibilidade, F1-Score, Chi-square ( $\chi^2$ ) e tempo do treinamento e validação.

<b>Modelo</b>	<b>ACC</b> (%)	<b>TFN</b> (%)	<b>Precisão</b> (%)	<b>Sensibilidade</b> (%)	<b>F1-Score</b> (%)	<b><math>\chi^2</math></b>	<b>Tempo de Treinamento e validação</b> (min:sec)
DenseNet-201	99,55	1,36	100,00	98,64	99,32	27,70	95:35
GoogLeNet	98,88	3,40	100,00	96,60	98,27	21,61	5:10
MobileNetV2	97,32	8,16	100,00	91,84	95,74	10,75	12:36
ResNet-18	97,54	7,48	100,00	92,52	96,11	12,03	3:52
FF model	100,00	0,00	100,00	100,00	100,00	32,00	130:59

Fonte: Compilação do autor.

As métricas utilizadas para avaliação do modelo MMD (Tabela 5.2) são as mesmas empregadas para avaliação dos modelos pré-treinados. Os valores dessas métricas, obtidas com o conjunto *dbl\_test*, são apresentados na Tabela 5.4.

Tabela 5.4 - Desempenho das CNN DenseNet-201, GoogLeNet, MobileNetV2, ResNet-18 e FF, obtidos com o conjunto de dados *dbl\_test* mostrando métricas: acurácia (ACC), Taxa de Falsos Negativos (TFN), Sensibilidade, F1-Score, *Chi-square* ( $\chi^2$ ) e tempo do treinamento e validação Desempenho dos modelos.

Modelo	ACC (%)	TFN (%)	Precisão (%)	Sensibilidade (%)	F1-Score (%)	$\chi^2$	Tempo de Treinamento e validação (min:sec)
Teste (b) CNN: DenseNet-201	99,55	1,36	100,00	98,64	99,32	27,70	95:35
Teste (c) CNN: GoogLeNet	98,66	4,08	100,00	95,92	97,92	21,61	5:10
Teste (d) CNN: MobileNetV2	97,32	8,16	100,00	91,84	95,74	10,75	12:36
Teste (e) CNN: ResNet-18	98,43	4,76	100,00	95,24	97,56	18,05	3:52
Teste (f) CNN: FF	100,00	0,00	100,00	100,00	100,00	32,00	130:59

Fonte: Compilação do autor.

De acordo com os dados apresentados na Tabela 5.4, a CNN FF possui os melhores resultados para todas as métricas avaliadas, em relação as cinco CNN investigados nesta tabela, com destaque para sua acurácia de  $ACC_{FF} = 100,00\%$ .

Das 447 imagens que compõem o conjunto de dados de teste, conforme divisão indicada na Figura 4.13 (b), as cinco arquiteturas previram corretamente todas as imagens da classe sem vazamento de gás, ou seja, a Precisão das mesmas alcançou o valor de 100%.

Além do excelente desempenho da acurácia do modelo FF em relação aos demais, outra métrica de destaque deste modelo foi o TFN (taxa de imagens com vazamento classificadas erradamente como pertencentes à classe sem vazamento), que resultou na menor taxa de falsos negativos,  $TFN_{FF} = 0,00\%$ .

O modelo DenseNet-201 apresentou dois casos de FNs, resultando em  $TFN_{DenseNet-201} = 1,36\%$  e acurácia de  $ACC_{DenseNet-201} = 99,55\%$ . MobileNetV2 apresentou doze casos de FN, resultando no pior índice de classificação de imagens como FN,  $TFN_{FF} = 8,16\%$  e acurácia de 97,32%. As outras CNNs investigadas GoogLeNet e ResNet-18, obtiveram desempenhos inferiores ao da CNN FF.

A Figura 5.3 ilustra uma das imagens classificadas equivocadamente (FN) pelo DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18. Apesar da imagem possuir uma evidente nuvem de vazamento de gás, tais modelos não conseguiram identificar o vazamento.



Figura 5.3 - Imagem com vazamento real classificada pelas redes DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18, como sem vazamento. Fonte: Compilação do autor.

A discussão a respeito do *Chi-quadrado* será realizada na Seção 5.3, que trata da comparação entre o MMG e o Filtro de Novidade.

O comportamento da DenseNet-201, GoogLeNet, MobileNetV2, e ResNet-18, durante o treinamento e a validação, são mostrados na Figura 5.4 (a), Figura 5.4 (b), Figura 5.4 (c) e Figura 5.4 (d), respectivamente.

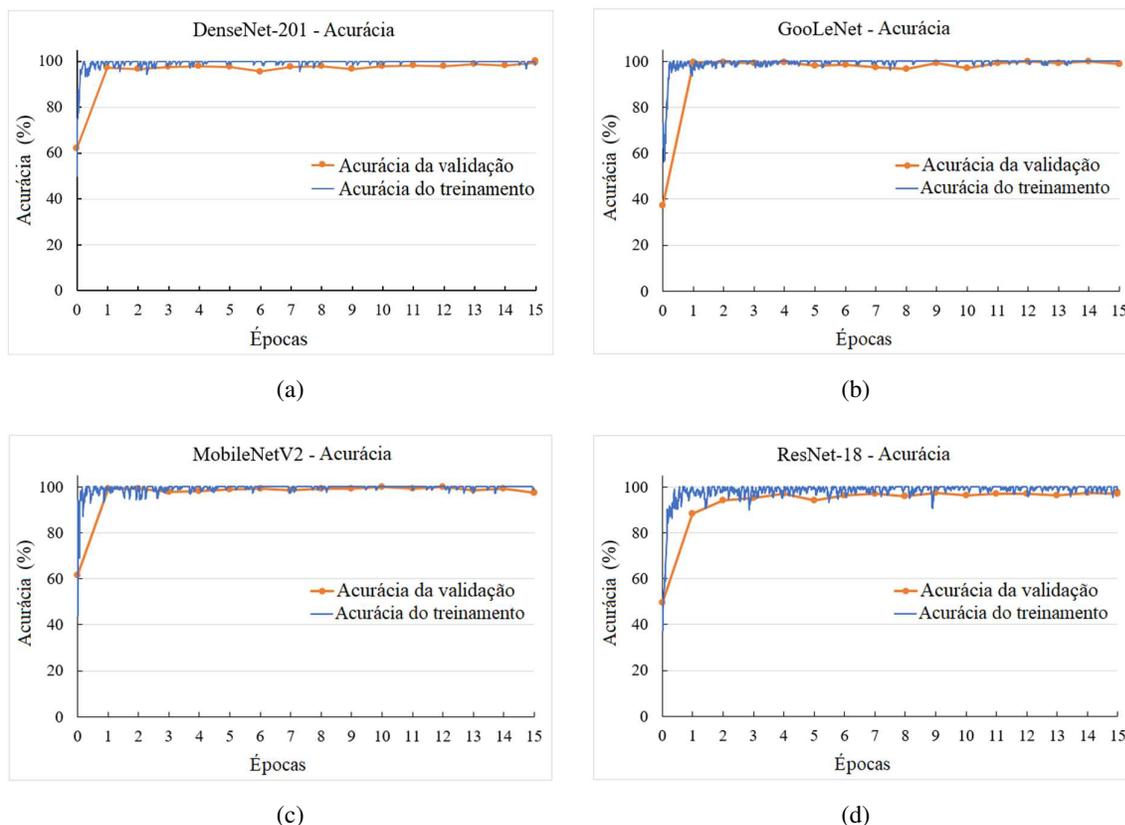


Figura 5.4 - (a), (b), (c) e (d) acurácia durante o treinamento e validação da DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18, respectivamente. Fonte: Compilação do autor.

O comportamento da acurácia do modelo FF, durante o treinamento e validação, é mostrado na Figura 5.5 (a), enquanto o seu comportamento quanto à perda é apresentado na Figura 5.5 (b), com perda de 0,0059 ao final da validação.

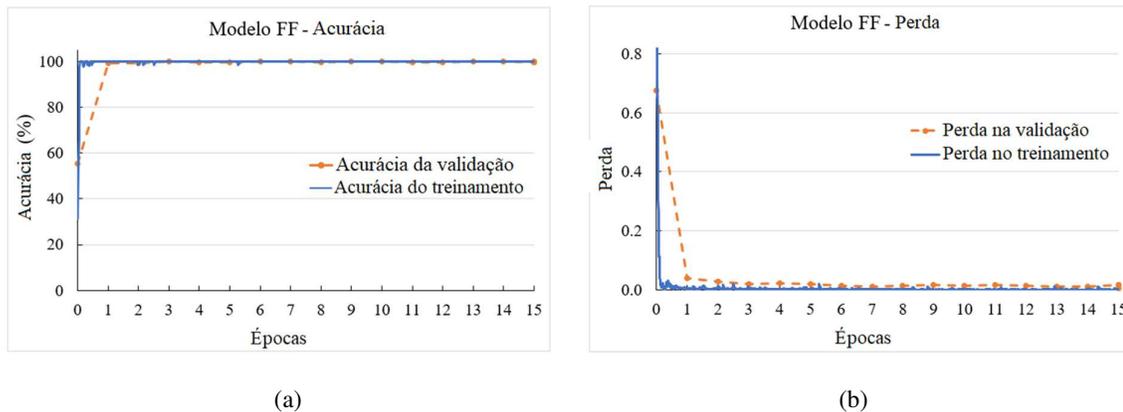


Figura 5.5 - Gráfico de comportamento do modelo FF durante o treinamento e validação: (a) acurácia e (b) perda. Fonte: Compilação do autor.

Ao comparar as curvas de acurácia do treinamento e da validação apresentadas na Figura 5.4 e Figura 5.5, verificou-se que, logo no início do treinamento, o modelo FF apresenta bons resultados e comportamento mais estável ao longo das épocas em relação aos demais modelos pré-treinados.

## 5.2 Determinação do Melhor Modelo Geral (MMG)

Os resultados dos testes apresentados nas Subseções 5.1.1 (Resultados dos 27 modelos de CNN propostos) e 5.1.2 (Resultados das CNN pré-treinados e CNN FF), são resumidos na Tabela 5.5. A partir dos dados desta tabela determina-se o melhor modelo geral (MMG), entre as seis CNN investigadas desta tese.

Com uma acurácia extraordinária, sem falhas na predição entre as duas classes, “com vazamento de GN” e “sem de vazamento de GN”, o modelo CNN FF, que combina as características de outros modelos de CNN pré-treinadas, é definido como Melhor Modelo Geral (MMG) para a tarefa de detecção de vazamentos de gás natural em imagens de CFTV. Além da acurácia, a CNN FF é superior no que tange ao desempenho em todas as demais métricas, conforme destacadas em tons de azul na Tabela 5.5.

Tabela 5.5 - Resumo das métricas finais dos modelos MMD, DenseNet-201, GoogLeNet, MobileNetV2, ResNet-18 e FF obtidas com o conjunto *db1\_test*.

Modelo	ACC (%)	TFN (%)	Precisão (%)	Sensibilidade (%)	F1-Score (%)	$\chi^2$	Tempo de Treinamento e validação (min:sec)	Tempo médio de teste/imagem (sec)
Teste ( a ) CNN: MMD	99,78	0,00	99,32	100,00	99,66	29,95	2:50	9,89E-06
Teste ( b ) CNN: DenseNet-201	99,55	1,36	100,00	98,64	99,32	27,70	95:35	13,10E-03
Teste ( c ) CNN: GoogLeNet	98,66	4,08	100,00	95,92	97,92	21,61	5:10	3,70E-03
Teste ( d ) CNN: MobileNetV2	97,32	8,16	100,00	91,84	95,74	10,75	12:36	4,60E-03
Teste ( e ) CNN: ResNet-18	98,43	4,76	100,00	95,24	97,56	18,05	3:52	2,20E-03
Teste ( f ) CNN: FF	100,00	0,00	100,00	100,00	100,00	32,00	130:59	22,30E-03

Fonte: Compilação do autor.

O tempo de treinamento e validação da CNN FF é superior ao das outras redes, o que era esperado, pois sua arquitetura é a mais complexa entre as CNN, já que esta rede é o resultado da fusão dos quatro modelos de CNN pré-treinados. No entanto, o treinamento e validação do modelo FF é realizado uma única vez a fim de ajustá-la a tarefa de destino. Com a CNN FF treinada e validada, o tempo médio de teste de uma imagem é de 22,3 ms e desvio padrão de  $0,3811E - 3$ , ou seja, o sistema computacional utilizado no desenvolvimento deste trabalho consegue avaliar 44 imagens por segundos que é superior a taxa de quadros por segundos de um vídeo com 30 quadros por segundo. Portanto, a CNN FF consegue testar todos os frames de um vídeo (30 quadros por segundo) com vazamento em tempo real.

As matrizes de confusão dos seis modelos estudados, nesta seção, são apresentadas nas Tabelas 5.6 a 5.11. Nestas matrizes é possível identificar a quantidade e tipos de falhas em cada modelo, assim como constatar, através da matriz de confusão da Tabela 5.11, a inferência perfeita do modelo FF, sem erros de predição durante os testes.

Tabela 5.6 - Matriz de confusão da CNN MMD, obtida com o conjunto *db1\_test*.

	Classe prevista	
	com vazamento	sem vazamento
Classe real		
com vazamento	147	0
sem vazamento	1	299

Fonte: Compilação do autor.

Tabela 5.7 - Matriz de confusão da CNN DenseNet-201, obtida com o conjunto *dbl\_test*.

	Classe prevista	
	com vazamento	sem vazamento
Classe real		
com vazamento	145	2
sem vazamento	0	300

Fonte: Compilação do autor.

Tabela 5.8 - Matriz de confusão da CNN GoogLeNet, obtida com o conjunto *dbl\_test*.

	Classe prevista	
	com vazamento	sem vazamento
Classe real		
com vazamento	141	6
sem vazamento	0	300

Fonte: Compilação do autor.

Tabela 5.9 - Matriz de confusão da CNN MobileNetV2, obtida com o conjunto *dbl\_test*.

	Classe prevista	
	com vazamento	sem vazamento
Classe real		
com vazamento	135	12
sem vazamento	0	300

Fonte: Compilação do autor.

Tabela 5.10 - Matriz de confusão da CNN ResNet-18, obtida com o conjunto *dbl\_test*.

	Classe prevista	
	com vazamento	sem vazamento
Classe real		
com vazamento	140	7
sem vazamento	0	300

Fonte: Compilação do autor.

Tabela 5.11 - Matriz de confusão da CNN FF, obtida com o conjunto *db1\_test*.

	Classe prevista	
	com vazamento	sem vazamento
Classe real		
com vazamento	147	0
sem vazamento	0	300

Fonte: Compilação do autor.

Deste modo, com acurácia de 100%, TFN de 0% e tempo de teste de imagens inferior a  $\frac{1}{24}$  segundos, a CNN FF consolida-se como o MMG.

### 5.3 Comparação dos resultados da CNN MMG e filtro de novidade

Entre os anos de 2012 e 2014, a detecção de vazamento de gás por imagem de CFTV, utilizando o conceito de filtro de novidade, obteve bons resultados. Agora, no ano de 2021, um novo método, utilizando aprendizagem profunda através de CNNs, superou satisfatoriamente os resultados de 2012, conforme dados da Tabela 5.12. O método com filtro novidade apresentou acurácia de 93,15% enquanto o modelo CNN FF obteve acurácia de 100,00%.

Tabela 5.12 - Comparação entre os resultados obtidos com a CNN FF (MMG) e com o Filtro de Novidade.

Método	Acurácia (%)	TFN (%)	Precisão (%)	Sensibilidade (%)	F1-Score (%)	$\chi^2$
Filtro de Novidade	93,15	8,78	88,34	91,22	89,76	--
CNN FF (MMG)	100,00	0,00	100,00	100,00	100,00	32,00

Fonte: Compilação do autor.

A CNN FF apresentou desempenho superior em comparação com o método anterior usando o filtro de novidade (MELO; COSTA FILHO; COSTA, 2014), com uma diferença estatisticamente significativa,  $\chi^2_{calculado} = 32,00$  ( $\chi^2_{calculado} > 6,63$ ). Assim, a hipótese nula é rejeitada. Portanto, o desempenho da CNN FF é admitido como superior ao desempenho do classificador que utiliza o filtro de novidade.

As demais métricas apresentadas na Tabela 5.12, reforçam a superioridade da CNN FF sobre a técnica de filtro de novidade para detecção de vazamentos de gás natural em imagens digitais, pois o modelo MMG obteve valores superiores em todas as métricas.

Os valores do teste  $\chi^2$  apresentados nas tabelas de seções anteriores mostram também que os outros métodos são superiores ao método que utiliza o filtro de novidade, apresentando diferenças de desempenho em relação a esse último que são estatisticamente significantes.

#### 5.4 Avaliação da capacidade de generalização do MMG (FF)

O desempenho do MMG, com o conjunto de dados *db2*, resultou em uma acurácia de 100,00%, ou seja, todas as 145 imagens deste conjunto, as quais são da classe “sem vazamento”, foram classificadas como “sem vazamento” (VN), corretamente.

A matriz de confusão correspondente a classificação do *db2* é mostrada na Tabela 5.13 (a). As Figura 5.6 (a) e Figura 5.6 (b) são exemplos de imagens que formam o *db2*.

Tabela 5.13 - Matriz de confusão da CNN FF (MMG), obtida com o conjunto *db2*. Fonte: Compilação do autor.

	Classe prevista	
	com vazamento	sem vazamento
Classe real		
com vazamento	0	0
sem vazamento	0	145



(a)



(b)

Figura 5.6 – As imagens (a) e (b) são exemplos dos elementos que constituem o conjunto de dados *db2*. Fonte: Compilação do autor.

Nessas imagens, observa-se a presença de muitas nuvens e céu. Tal cenário (nuvens) possui textura similar a um vazamento de gás natural, porém a CNN FF foi capaz de classificá-las adequadamente, conforme Tabela 5.14.

Tabela 5.14 - Desempenho da CNN FF obtido com o conjunto de dados *db2*.

Modelo	Conjunto de dados	Acurácia (%)	TFN (%)
CNN FF (MMG)	<i>db2</i> (imagens sem vazamento, com nuvens)	100,00	0,00

Fonte: Compilação do autor.

## 5.5 Explicação visual

Definido o MMG, ou seja, CNN FF, ele é submetido à ferramenta Grad-CAM, que fornece o mapa de calor das regiões da imagem que estão fortemente relacionadas à decisão da CNN em classificá-la como verdadeiro positivo.

A Figura 5.7 (a) demonstra uma imagem com vazamento de GN, que o MMG rotulou como tal. O Grad-CAM então fornece o mapa de calor desta imagem, conforme mostrado na Figura 5.7 (b). Neste experimento, o mapa de calor está filtrado, a fim de apresentar apenas as pontuações cujas magnitudes são maiores do que o valor médio do mapa original.

A camada de mapa de característica escolhida no MMG para a aplicação Grad-CAM foi a saída do bloco *inception5b*, conforme mostrado no Quadro 4.1.

Para melhorar a visualização dos *pixels* na imagem de entrada que contribuem para a previsão de vazamento de GN, sobrepomos o mapa de calor da Figura 5.7 (b) sobre a imagem de entrada da Figura 5.7 (a), conforme resultado apresentado na Figura 5.7 (c). Embora a textura da nuvem de gás seja lisa, é possível ver na imagem de entrada a silhueta dela na parte esquerda e superior da cabeça do poço, justamente a região onde o mapa de calor é mais intenso.

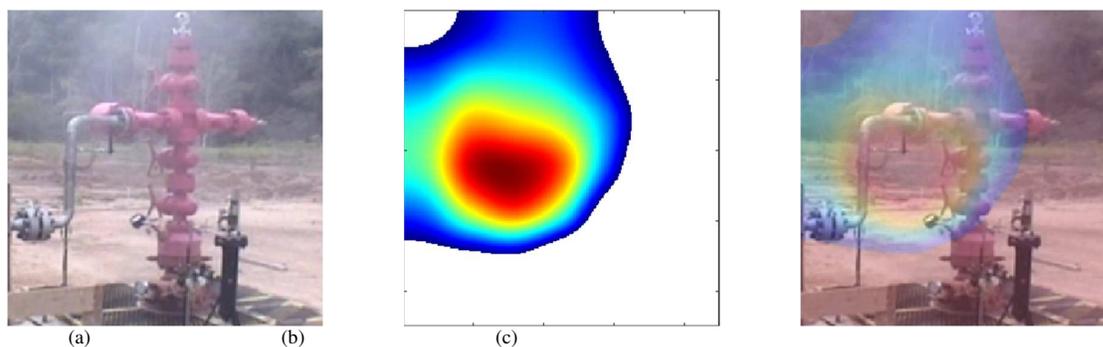


Figura 5.7 - (a) exemplo de uma imagem original com vazamento; (b) mapa de calor com magnitudes pré-processadas; (c) identificação dos pixels que contribuíram para a previsão de vazamento. Os pixels com cor vermelha contribuíram fortemente para a pontuação da classe com vazamento. Fonte: Compilação do autor.

## 5.6 Conclusão

Neste capítulo foi apresentada uma investigação quantitativa que ratificou a proposta de um modelo de detecção de vazamento de GN baseado em aprendizagem profunda. Os melhores resultados obtidos com uma máquina de reconhecimento de padrões baseada em CNN foram muito satisfatórios, com acurácia de 100 % e taxa de falsos negativos de 0%.

Esse resultado foi alcançado após experimentos com trinta e dois modelos de CNN, sendo eles: 27 modelos desenvolvidos do zero (*scratch*), quatro modelos de redes pré-treinadas e um modelo constituído da fusão de características destes quatro modelos pré-treinados.

Dois bancos de imagens foram utilizados (*db1* e *db2*) para ajustar e testar estes modelos. Após análise dos resultados, o modelo baseado em fusão de características (FF) se destacou, apresentando os melhores resultados: acurácia foi de 100 % e taxa de falsos negativos de 0%. Ao mesmo, atribuiu-se o nome de Melhor Modelo Geral – MMG.

Utilizando o teste de significância estatística de *Chi-quadrado*, foi possível evidenciar a capacidade superior da CNN MMG, quando comparado a outro método previamente desenvolvido para detecção de vazamento GN utilizando a mesma base de dados, baseado na utilização do filtro de novidade.

O capítulo ainda destacou a importância prática do sistema de aprendizagem profundo em fornecer uma explicação visual de suas inferências, o que facilita ao operador de vídeo a compreender o que motivou a CNN MMG a prever que uma imagem é da classe com vazamento de GN.



## 6. CONCLUSÕES

O principal objetivo desta tese foi propor um método utilizando redes profundas para detecção de vazamento de gás natural em cabeça de poço petrolíferos *onshore*, mais especificamente utilizando redes neurais convolucionais, classificando imagens digitais de cabeças de poço em duas classes, “com vazamento de gás natural” e “sem vazamento de gás natural”.

Cabe acrescentar que grande parte das imagens utilizadas nesse trabalho foram obtidas por meio de sistemas de vigilância CFTV já existentes nas cabeças de poço, sem necessidade de instalações adicionais ou hardwares que capturam moléculas de metano na atmosfera.

Destaca-se abrangência dos objetivos propostos e atingidos por este estudo, que não se limitou a propor uma ferramenta de reconhecimento de padrões para detecção de vazamento de gás natural, mas, principalmente, analisou diversas arquiteturas para atingir o objetivo principal, assim como comparou os resultados com outros trabalhos já publicados, procurou identificar a região da imagem onde ocorre o vazamento e testou as diversas arquiteturas propostas com banco de dados desafiadoras de cabeças de poço, contendo imagens com fundos diferentes.

Nesse sentido, as seções 6.1 a **Erro! Fonte de referência não encontrada.** descrevem os objetivos específicos propostos nesta tese, os quais foram atingidos na sua integralidade.

### 6.1 Propor e avaliar o desempenho de três diferentes arquiteturas CNN e selecionar aquela que oferece o melhor desempenho na detecção de vazamentos de gás natural

Entre os vinte e sete modelos propostos (*scratch*), o modelo número 7 apresentou os melhores resultados para a detecção de vazamento de GN em cabeça de poços petrolíferos. A arquitetura desse modelo é do tipo CNNL (*Convolutional Neural Network* Longa), arquitetura longa, treinada com o método de otimização SGDM, e *dropout* como técnica de regularização. O desempenho deste modelo alcançou uma acurácia de 99,78% e TFN (Taxa de Falso Negativo) de 0,00%. Os valores destas métricas definem o modelo número 7 como o nosso MMD (Modelo com Melhor Desempenho), entre os vinte e sete modelos investigados.

### 6.2 Avaliar o desempenho de redes pré-treinadas DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18, quanto a detecção de vazamentos de gás natural

Além do MMD, outros modelos de CNN, que aplicam aprendizagem por transferência, foram avaliados, entre eles CNN DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18. Todos os modelos pré-treinados obtiveram acurácia inferior a CNN MMD. Além disso, esses

modelos resultaram em TFN diferentes de 0,00%, ou seja, algumas imagens com vazamento (VP), foram classificadas como sem vazamento (FN).

### **6.3 Propor e avaliar o desempenho de um modelo CNN baseado em fusão de características (CNN FF), obtido por meio da extração simultânea de características das redes DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18, quanto a detecção de vazamentos de gás natural.**

Após o treinamento, validação e teste destas quatro CNNs pré-treinadas, de cada CNN é extraído um vetor de característica das camadas de agrupamento. Após extrair as características, elas são concatenadas em um único vetor. A arquitetura que utilizou essa concepção foi denominada por FF (*Feature Fusion*), e seu desempenho alcançou uma acurácia de 100,00% e TFN de 0,00%.

### **6.4 Comparar o desempenho das CNN MMD, DenseNet-201, GoogLeNet, MobileNetV2 e ResNet-18 e FF, com respeito, principalmente, a acurácia e a taxa de falso negativos fornecidas por essas CNNs, e assim definir o Melhor Modelo Geral (CNN MMG), quanto a detecção de vazamentos de gás natural**

Entre os seis modelos (MMD, DenseNet-201, GoogLeNet, MobileNetV2, ResNet-18 e FF), a arquitetura FF foi definida como o nosso Melhor Modelo Geral (MMG), pois ela se destacou entre todos os modelos estudados, alcançando a acurácia de 100%, ou seja, a TFN = 0,00%. Portanto, a combinação de características de diferentes redes pré-treinadas fornece uma acurácia melhor do que o desempenho individual delas.

Mesmo com o modelo FF necessitando de um tempo de treinamento e validação mais longos do que as redes individuais pré-treinadas e MMD, o seu tempo para testar uma imagem é semelhante àqueles obtidos em todos os seis modelos.

### **6.5 Comparar o desempenho do modelo baseado em CNN com o do modelo baseado em filtro de novidade através do teste Chi-quadrado**

Os valores das métricas do modelo FF são superiores aqueles obtidos em estudos semelhantes encontrados na literatura, que utilizaram a técnica do filtro de novidade (MELO; COSTA FILHO; COSTA, 2014). Por meio da aplicação do teste de hipótese *Chi-quadrado* ( $\chi^2$ ), ficou claro que esse desempenho superior do FF é significativo a um nível de significância de 0,01.

## **6.6 Investigar a capacidade do melhor modelo geral em distinguir nuvens do céu de plumas de GN através da avaliação de imagens de diferentes cabeças de poço, obtidas em diversas condições de ambiente, desconhecidas do modelo**

Para investigar a capacidade da CNN FF (MMG) em classificar imagens contendo nuvens do céu que possuem textura similar ao vazamento de GN, seu desempenho foi avaliado com outro conjunto de dados de imagens obtidos de novas cabeças de poço em que o fundo das imagens da cabeça dos poços contém nuvens no céu. O desempenho do FF nesse novo conjunto de dados foi  $ACC = 100\%$ .

## **6.7 Limitações e desafios**

A principal limitação deste trabalho está relacionada a baixa representatividade do conjunto de dados *db2*, referente as imagens da classe “com vazamento” de GN, visto que ele é formado apenas por imagens da classe “sem vazamento” de GN. Cabe enfatizar que a ação de simular vazamentos de GN nos 113 *wellheads*, afim de obter imagens da classe “com vazamento”, seria um risco adicional para a equipe que conduzia os experimentos *in loco*, dado o potencial risco de ocorrer um acidente, além da emissão de moléculas de  $CH_4$  para atmosfera.

No entanto, as imagens do *db2* possuem objetos semânticos (nuvens do céu) com textura similar ao vazamento GN, e a abordagem proposta neste estudo, usando redes convolucionais, foi consistente, uma vez que a CNN MMG forneceu excelentes resultados, sendo capaz de classificar corretamente todas as imagens deste conjunto de dados.

## **6.8 Trabalhos futuros**

Como propostas de trabalhos futuros, desenhamos as seguintes hipóteses de trabalho:

- 1) Validar o método proposto com uma maior quantidade de imagens;
- 2) Aplicar os conceitos desenvolvidos na detecção de GN em outros tipos de vazamentos industriais de gás;
- 3) Avaliar a possibilidade de transformar o método proposto em um produto que seja útil a indústria petrolífera.



## REFERÊNCIAS

**2011-2019 Greenhouse Gas Reporting Program Industrial Profile: Petroleum and Natural Gas Systems. Environmental Protection Agency.** [s.l: s.n.]. Disponível em: <[https://www.epa.gov/sites/production/files/2020-11/documents/subpart\\_w\\_2019\\_industrial\\_profile.pdf](https://www.epa.gov/sites/production/files/2020-11/documents/subpart_w_2019_industrial_profile.pdf)>. Acesso em: 30 jan. 2021.

ABBAS, Asmaa; ABDELSAMEA, Mohammed M.; GABER, Mohamed Medhat. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. **Applied Intelligence**, [s. l.], v. 51, n. 2, p. 854–864, 2021. Disponível em: <<https://doi.org/10.1007/s10489-020-01829-7>>

ADEGBOYE, Mutiu Adesina; FUNG, Wai Keung; KARNIK, Aditya. Recent advances in pipeline monitoring and oil leakage detection technologies: Principles and approaches. **Sensors (Switzerland)**, [s. l.], v. 19, n. 11, 2019.

AGGARWAL, Charu C. **Neural Networks and Deep Learning**. Cham: Springer International Publishing, 2018. Disponível em: <<http://link.springer.com/10.1007/978-3-319-94463-0>>. Acesso em: 18 out. 2019.

AL-WAISY, Alaa S. et al. A multi-biometric iris recognition system based on a deep learning approach. **Pattern Analysis and Applications**, [s. l.], v. 21, n. 3, p. 783–802, 2018.

ARAUJO RODRIGUES, Larissa; LUÍS SAUER, Ildo. Exploratory assessment of the economic gains of a pre-salt oil field in Brazil. **Energy Policy**, [s. l.], v. 87, p. 486–495, 2015.

ATKINSON, Graham; HALL, Jonathan; MCGILLIVRAY, Alison. **Review of Vapour Cloud Explosion Incidents**. Harpur Hill.

BAROUDI, Uthman; AL-ROUBAIEY, Anas A.; DEVENDIRAN, Abdullah. Pipeline Leak Detection Systems and Data Fusion: A Survey. **IEEE Access**, [s. l.], v. 7, p. 97426–97439, 2019.

BASU, Swapan. Instrumentation Safety Implementation and Explosion Protection. In: BASU, Swapan (Ed.). **Plant Hazard Analysis and Safety Instrumentation Systems**. [s.l.] : Academic Press, 2017. p. 699–806.

BENGIO, Yoshua; HINTON, Geoffrey; LECUN, Yann. FATHERS OF THE DEEP LEARNING REVOLUTION RECEIVE ACM A.M. TURING AWARD Bengio, Hinton and LeCun Ushered in Major Breakthroughs in Artificial Intelligence. [s. l.], p. 1–5, 2019. Disponível em: <<https://awards.acm.org/binaries/content/assets/press-releases/2019/march/turing-award-2018.pdf>>. Acesso em: 22 out. 2019.

BOARD; INVESTIGATION, Buncefield Major Incident. **Buncefield Major Incident Investigation: Initial Report to the Health and Safety Commission and the Environment Agency of the Investigation Into the Explosions and Fires at the Buncefield Oil Storage and Transfer Depot, Hemel Hempstead, 11 December 2005**. Hertfordshire.

BP. **Statistical Review of World Energy**. 2021. Disponível em: <<https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2021-full-report.pdf>>. Acesso em: 19 jul. 2021.

BU, Fengju; GHARAJEH, Mohammad Samadi. Intelligent and vision-based fire detection systems: A survey. **Image and Vision Computing**, [s. l.], v. 91, p. 103803, 2019.

CALABRÒ, Paolo S. Greenhouse gases emission from municipal waste management: The role of separate collection. **Waste Management**, [s. l.], v. 29, n. 7, p. 2178–2187, 2009. Disponível

em: <<https://www.sciencedirect.com/science/article/pii/S0956053X09000713>>. Acesso em: 20 fev. 2020.

CARLSON, Brian N. Selection and use of pipeline leak detection methods for liability management into the 21st century. **Proceedings of the International Conference on Pipeline Infrastructure II**, [s. l.], p. 369–383, 1993.

CHATTOPADHAY, Aditya et al. Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks. **Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018**, [s. l.], v. 2018- Janua, p. 839–847, 2018.

CHEN, Ying; XU, Xiuqin; KOCH, Thorsten. Day-ahead high-resolution forecasting of natural gas demand and supply in Germany with a hybrid model. **Applied Energy**, [s. l.], v. 262, p. 114486, 2020.

CHEN, Yushi et al. Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. **IEEE Transactions on Geoscience and Remote Sensing**, [s. l.], v. 54, n. 10, p. 6232–6251, 2016.

COSTA FILHO, Cícero Ferreira Fernandes; MELO, Roberlanio de Oliveira; COSTA, Marly Guimarães Fernandes. Detecting natural gas leaks using digital images and novelty filters. In: (Mohamed Kamel, Fakhri Karray, Hani Hagra, Eds.) **LECTURE NOTES IN COMPUTER SCIENCE (INCLUDING SUBSERIES LECTURE NOTES IN ARTIFICIAL INTELLIGENCE AND LECTURE NOTES IN BIOINFORMATICS) 2012**, Berlin, Heidelberg. **Anais...** Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

COSTA FILHO, Cícero Ferreira Fernandes; MELO, Roberlanio De Oliveira; COSTA, Marly Guimarães Fernandes. An improved system for detecting natural gas leaks. In: **PROCEEDINGS OF 2013 SCIENCE AND INFORMATION CONFERENCE, SAI 2013 2013**, **Anais...** [s.l: s.n.]

COSTA, Marly Guimarães Fernandes et al. Evaluating the performance of convolutional neural networks with direct acyclic graph architectures in automatic segmentation of breast lesion in US images. **BMC Medical Imaging**, [s. l.], v. 19, n. 1, p. 85, 2019. Disponível em: <<https://bmcmedimaging.biomedcentral.com/articles/10.1186/s12880-019-0389-2>>

DARWISH, Ashraf; EZZAT, Dalia; HASSANIEN, Aboul Ella. An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis. **Swarm and Evolutionary Computation**, [s. l.], v. 52, p. 100616, 2020.

DEL RIO, Felipe et al. **Do Better ImageNet Models Transfer Better... for Image Recommendation?**, 2018.

FARAMAWY, S.; ZAKI, T.; SAKR, A. A. E. Natural gas origin, composition, and processing: A review. **Journal of Natural Gas Science and Engineering**, [s. l.], v. 34, p. 34–54, 2016.

FLIR. **The Science Behind Optical Gas Imaging | FLIR Systems**. 2020. Disponível em: <<https://www.flir.com/discover/instruments/gas-detection/the-science-behind-optical-gas-imaging/>>. Acesso em: 18 set. 2020.

FORSTER, P. et al. Changes in atmospheric constituents Physical, and in radiative forcing. In: **Climate Change 2007: the Fourth, Science Basis. Contribution of Working Group I to the Climate, Assessment Report of the Intergovernmental Panel on Change**. [s.l: s.n.]. p. 129–234.

FUKUSHIMA; K. Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position-Neocognitron-. **IEICE Technical Report, A**, [s. l.], v. 62, n. 10, p. 658–

665, 1979. Disponível em: <<http://ci.nii.ac.jp/naid/10026558204/en/>>. Acesso em: 28 out. 2019.

FULTON, Mark et al. **Comparing Life-Cycle Greenhouse Gas Emissions from Natural Gas and CoalDB Climate Change Advisors**. [s.l.: s.n.]. Disponível em: <[http://www.worldwatch.org/system/files/pdf/Natural\\_Gas\\_LCA\\_Update\\_082511.pdf](http://www.worldwatch.org/system/files/pdf/Natural_Gas_LCA_Update_082511.pdf)>.

GARG, Tanmay et al. Convolutional Neural Networks with Transfer Learning for Recognition of COVID-19: A Comparative Study of Different Approaches. **AI**, [s. l.], v. 1, n. 34, p. 586–606, 2020. Disponível em: <<https://doaj.org/article/6884860a3439448797a7b4505576b01a>>. Acesso em: 9 jul. 2021.

GHIMIRE, S. et al. Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. **Applied Energy**, [s. l.], v. 253, p. 113541, 2019.

GLASBY, Geoffrey. Abiogenic Origin of Hydrocarbons: An Historical Overview. **Resource Geology**, [s. l.], v. 56, p. 83–96, 2006.

GLOROT, Xavier; BORDES, Antoine; BENGIO, Yoshua. Deep Sparse Rectifier Neural Networks. In: (Geoffrey Gordon, David Dunson, Miroslav Dudík, Eds.) PROCEEDINGS OF THE FOURTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND STATISTICS 2011, Fort Lauderdale, FL, USA. **Anais...** Fort Lauderdale, FL, USA: PMLR, 2011. Disponível em: <<http://proceedings.mlr.press/v15/glorot11a.html>>

GOLD, Thomas. The Origin of Natural Gas and Petroleum, and the Prognosis for Future Supplies. **Annual Review of Energy**, [s. l.], v. 10, n. 1, p. 53–77, 1985. Disponível em: <<https://doi.org/10.1146/annurev.eg.10.110185.000413>>

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. [s.l.] : The MIT Press, 2016.

HAMEED, Hadia; MAZHAR, Suleman; HASSAN, Naufil. Real-time road anomaly detection, using an on-board data logger. In: IEEE VEHICULAR TECHNOLOGY CONFERENCE 2018, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2018.

HAUGLAND, Torleif. **Best Practice Guidance for Effective Methane Management in the Oil and Gas Sector: Monitoring, Reporting and Verification (MRV) and Mitigation**. 65. ed. Genebra. Disponível em: <[https://unece.org/DAM/energy/images/CMM/CMM\\_CE/Best\\_Practice\\_Guidance\\_for\\_Effective\\_Methane\\_Management\\_in\\_the\\_Oil\\_and\\_Gas\\_Sector\\_\\_Monitoring\\_\\_Reporting\\_and\\_Verification\\_\\_MRV\\_\\_and\\_Mitigation-FINAL\\_\\_with\\_covers\\_.pdf](https://unece.org/DAM/energy/images/CMM/CMM_CE/Best_Practice_Guidance_for_Effective_Methane_Management_in_the_Oil_and_Gas_Sector__Monitoring__Reporting_and_Verification__MRV__and_Mitigation-FINAL__with_covers_.pdf)>. Acesso em: 30 jan. 2021.

HE, Kaiming et al. Deep residual learning for image recognition. In: PROCEEDINGS OF THE IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2016, **Anais...** : IEEE Computer Society, 2016. Disponível em: <<http://image-net.org/challenges/LSVRC/2015/>>. Acesso em: 25 ago. 2020.

HEATON, Jeff. Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning. **Genetic Programming and Evolvable Machines**, [s. l.], v. 19, n. 1–2, p. 305–307, 2018.

HINTON, Geoffrey E. A Practical Guide to Training Restricted Boltzmann Machines. In: MONTAVON, Grégoire; ORR, Geneviève B.; MÜLLER, Klaus-Robert (Eds.). **Neural Networks: Tricks of the Trade: Second Edition**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. p. 599–619.

HINTON, Geoffrey E.; OSINDERO, Simon; TEH, Yee-Whye. A Fast Learning Algorithm for Deep Belief Nets. **Neural Computation**, [s. l.], v. 18, n. 7, p. 1527–1554, 2006. Disponível

em: <<http://www.mitpressjournals.org/doi/10.1162/neco.2006.18.7.1527>>. Acesso em: 28 out. 2019.

HOUGH; E, J. Leak testing of pipelines uses pressure and acoustic velocity. **Oil Gas J.; (United States)**, [s. l.], v. 86:47, 1988.

HOWARD, Andrew G. et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. [s. l.], 2017. Disponível em: <<http://arxiv.org/abs/1704.04861>>. Acesso em: 20 jan. 2021.

HOWELL, David C. **Stastical Methods for Psychology**. 7th. ed. [s.l.] : Cengage Wadsworth, 2010.

HOWELL, David C. Chi-Square Test: Analysis of Contingency Tables. In: **International Encyclopedia of Statistical Science**. [s.l.] : Springer Berlin Heidelberg, 2011. p. 250–252.

HSU, Sherman. Infrared Spectroscopy. In: **Handbook of Instrumental Techniques for Analytical Chemistry**. [s.l.] : American Chemical Society, 1998. v. 120p. 6633.

HUANG, Gao et al. Densely connected convolutional networks. In: PROCEEDINGS - 30TH IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, CVPR 2017 2017, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2017.

HUANG, Lvwen et al. A Multi-Feature Fusion Based on Transfer Learning for Chicken Embryo Eggs Classification. **Symmetry 2019, Vol. 11, Page 606**, [s. l.], v. 11, n. 5, p. 606, 2019. Disponível em: <<https://www.mdpi.com/2073-8994/11/5/606/htm>>. Acesso em: 10 jul. 2021.

IBRAHIMOV, Bahadur. A cost-oriented robot for the Oil Industry. **IFAC-PapersOnLine**, [s. l.], v. 51, n. 30, p. 204–209, 2018. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S2405896318329604>>

ICF INTERNATIONAL. Economic Analysis of Methane Emission Reduction Opportunities in the Canadian Oil and Natural Gas Industries. [s. l.], n. October, p. 118, 2015. Disponível em: <<http://www.pembina.org/reports/edf-icf-methane-opportunities.pdf>>

IOFFE, Sergey; SZEGEDY, Christian. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: 32ND INTERNATIONAL CONFERENCE ON MACHINE LEARNING, ICML 2015 2015, **Anais...** [s.l.: s.n.] Disponível em: <<https://arxiv.org/abs/1502.03167>>. Acesso em: 15 out. 2019.

JADIN, Mohd Shawal; GHAZALI, Kamarul Hawari. Gas leakage detection using thermal imaging technique. In: PROCEEDINGS - UKSIM-AMSS 16TH INTERNATIONAL CONFERENCE ON COMPUTER MODELLING AND SIMULATION, UKSIM 2014 2014, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2014.

KERMANY, Daniel S. et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. **Cell**, [s. l.], v. 172, n. 5, p. 1122- 1131.e9, 2018.

KHALIL, Y. F. A probabilistic visual-flowcharting-based model for consequence assessment of fire and explosion events involving leaks of flammable gases. **Journal of Loss Prevention in the Process Industries**, [s. l.], v. 50, p. 190–204, 2017.

KHAN, Sana Ullah et al. A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. **Pattern Recognition Letters**, [s. l.], v. 125, p. 1–6, 2019.

KING, Andrew P.; ECKERSLEY, Robert J. Inferential Statistics III: Nonparametric Hypothesis Testing. In: **Statistics for Biomedical Engineers and Scientists**. [s.l.] : Elsevier,

2019. p. 119–145.

KINGMA, Diederik P.; BA, Jimmy. Adam: A Method for Stochastic Optimization. [s. l.], 2014. Disponível em: <<http://arxiv.org/abs/1412.6980>>. Acesso em: 15 out. 2019.

KOFLER, Corinna; MUHR, Robert; SPÖCK, Gunter. Classifying Image Stacks of Specular Silicon Wafer Back Surface Regions: Performance Comparison of CNNs and SVMs. **Sensors (Basel, Switzerland)**, [s. l.], v. 19, n. 9, p. 2056, 2019. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/31052579>>

KOHONEN, T.; OJA, E. Fast adaptive formation of orthogonalizing filters and associative memory in recurrent networks of neuron-like elements. **Biological Cybernetics**, [s. l.], v. 21, n. 2, p. 85–95, 1976.

KOROTCENKOV, Ghenadii. Introduction. In: **Handbook of Gas Sensor Materials: Properties, Advantages and Shortcomings for Applications Volume 1: Conventional Approaches**. New York, NY: Springer New York, 2013. p. 1–45.

KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. ImageNet classification with deep convolutional neural networks. **Advances in Neural Information Processing Systems**, [s. l.], v. 2, p. 1097–1105, 2012.

LECUN, Y. et al. Backpropagation Applied to Handwritten Zip Code Recognition. **Neural Computation**, [s. l.], v. 1, n. 4, p. 541–551, 1989. Disponível em: <<http://www.mitpressjournals.org/doi/10.1162/neco.1989.1.4.541>>. Acesso em: 22 out. 2019.

LECUN, Yann et al. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, [s. l.], v. 86, n. 11, p. 2278–2323, 1998. Disponível em: <<http://ieeexplore.ieee.org/document/726791/>>. Acesso em: 22 out. 2019.

LI, Lei et al. A gas leakage localization method based on a virtual ultrasonic sensor array. **Sensors (Switzerland)**, [s. l.], v. 19, n. 14, 2019. Disponível em: <</pmc/articles/PMC6679241/>>. Acesso em: 2 mar. 2021.

LIU, Binglu et al. Monitoring and detection of combustible gas leakage by using infrared imaging. In: IST 2018 - IEEE INTERNATIONAL CONFERENCE ON IMAGING SYSTEMS AND TECHNIQUES, PROCEEDINGS 2018, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2018.

LIU, Xiaoming; ZHOU, Xiaobo; QIAN, Xiaohua. Transparency-guided ensemble convolutional neural network for the stratification between pseudoprogession and true progression of glioblastoma multiform in MRI. **Journal of Visual Communication and Image Representation**, [s. l.], v. 72, p. 102880, 2020. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1047320320301267>>

MAGEE, Charles. **Survey of Natural Gas Leakage Abatement Best Practices** California/USAReport Number: R.15-01-008 CEK/ek4, , 2015.

MARSHALL; PARK, Jang Sik; SONG, Jong Kwan. FCN Based Gas Leakage Segmentation and Improvement Using Transfer Learning. In: 2019 IEEE STUDENT CONFERENCE ON ELECTRIC MACHINES AND SYSTEMS, SCEMS 2019 2019, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2019.

MCCAIN, William D. **Properties of Petroleum Fluids**. 3rd. ed. Tulsa, Oklahoma: Pennwell Books, 2017.

MELO, Roberlanio De Oliveira; COSTA FILHO, Cícero Ferreira Fernandes; COSTA, Marly Guimarães Fernandes. Leak detection of natural gas with base on the components of color

spaces RGB and HSI using novelty filter. **IEEE Latin America Transactions**, [s. l.], v. 12, n. 8, p. 1560–1565, 2014. Disponível em: <<http://ieeexplore.ieee.org/document/7014528/>>

MELO, Roberlânio de Oliveira; COSTA, Marly Guimarães Fernandes; COSTA FILHO, Cícero Ferreira Fernandes. Using digital image processing and a novelty classifier for detecting natural gas leaks. In: CHEN, Liming; KAPOOR, Supriya; BHATIA, Rahul (Eds.). **Studies in Computational Intelligence**. Cham: Springer International Publishing, 2014. v. 542p. 409–422.

MIYAGAWA, M. et al. Detecting Vascular Bifurcation in IVOCT Images Using Convolutional Neural Networks With Transfer Learning. **IEEE Access**, [s. l.], v. 7, p. 66167–66175, 2019.

MOKHATAB, Saeid; POE, William A.; MAK, John Y. Handbook of natural gas transmission and processing: Principles and practices. **Handbook of Natural Gas Transmission and Processing: Principles and Practices**, [s. l.], p. 1–826, 2018.

MOLCHANOV, Pavlo et al. Hand gesture recognition with 3D convolutional neural networks. **IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops**, [s. l.], v. 2015- Octob, p. 1–7, 2015.

MONIEN, Burkhard; PREIS, Robert; SCHAMBERGER, Stefan. Approximation algorithms for multilevel graph partitioning. In: HANDBOOK OF APPROXIMATION ALGORITHMS AND METAHEURISTICS 2007, USA. **Anais...** USA: Curran Associates Inc., 2007. Disponível em: <<https://doi.org/10.1016/j.neunet.2019.05.010>>

MURVAY, Pal Stefan; SILEA, Ioan. A survey on gas leak detection and localization techniques... **Journal of Loss Prevention in the Process Industries**, [s. l.], v. 25, n. 6, p. 966–973, 2012.

MYHRE, G. et al. Anthropogenic and natural radiative forcing. In: STOCKER, T. F. et al. (Eds.). **Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change**. Cambridge, UK: Cambridge University Press, 2013. p. 659–740.

NIEDERER, Johannes M. G. The Infrared Spectrum of Methane. [s. l.], 2012. Disponível em: <<https://www.research-collection.ethz.ch/443/handle/20.500.11850/40595>>. Acesso em: 18 set. 2020.

OZKAYA, Umut; OZTURK, Saban; BARSTUGAN, Mucahid. Coronavirus (COVID-19) Classification using Deep Features Fusion and Ranking Technique. [s. l.], 2020. Disponível em: <<http://arxiv.org/abs/2004.03698>>. Acesso em: 26 ago. 2020.

ÖZYURT, Fatih. A fused CNN model for WBC detection with MRMR feature selection and extreme learning machine. **Soft Computing**, [s. l.], v. 24, n. 11, p. 8163–8172, 2020. Disponível em: <<https://link.springer.com/article/10.1007/s00500-019-04383-8>>. Acesso em: 26 ago. 2020.

PAN, Sinno Jialin; YANG, Qiang. A survey on transfer learning. **IEEE Transactions on Knowledge and Data Engineering**, [s. l.], v. 22, n. 10, p. 1345–1359, 2010.

PSYRRAS, Nikolaos K.; SEXTOS, Anastasios G. Safety of buried steel natural gas pipelines under earthquake-induced ground shaking: A review. **Soil Dynamics and Earthquake Engineering**, [s. l.], v. 106, p. 254–277, 2018. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0267726117301768>>

RABINOWITZ, Peter; SCOTCH, Matthew; CONTI, Lisa. Human and animal sentinels for shared health risks. **Veterinaria italiana**, [s. l.], v. 45, n. 1, p. 23–4, 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20148187%0Ahttp://www.pubmedcentral.nih.gov/arti>>

clerender.fcgi?artid=PMC2818012>

REM, Benno S. et al. Identifying Quantum Phase Transitions using Artificial Neural Networks on Experimental Data. **Nature Physics**, [s. l.], v. 15, n. 9, p. 917–920, 2018. Disponível em: <<http://arxiv.org/abs/1809.05519>>. Acesso em: 27 fev. 2020.

ROSENBLATT, Frank. The perceptron: A probabilistic model for information storage and organization in the brain. **Psychological Review**, [s. l.], v. 65, n. 6, p. 1–23, 1958. Disponível em:

<<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.335.3398&rep=rep1&type=pdf%0Apapers://d471b97a-e92c-44c2-8562-4efc271c8c1b/Paper/p322>>

RUEHLE, Fabian. **Data science applications to string theory** *Physics Reports* Elsevier B.V., 2019.

RUINEIHART, David E.; HINTON, Geoffrey E.; WILLIAMS, Ronald J. **Learning Internal Representations Error Propagation** *Cognitive Science*, 1986. Disponível em: <<https://doi.org/10.21236/ADA164453>>. Acesso em: 23 out. 2019.

SACHEDINA, K.; MOHANY, A. A review of pipeline monitoring and periodic inspection methods. **Pipeline Science and Technology**, [s. l.], v. 2, n. 3, p. 187–201, 2018.

SANDLER, Mark et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: PROCEEDINGS OF THE IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2018, **Anais...** : IEEE Computer Society, 2018.

SAPART, C. J. et al. Natural and anthropogenic variations in methane sources during the past two millennia. **Nature**, [s. l.], v. 490, n. 7418, p. 85–88, 2012. Disponível em: <<https://doi.org/10.1038/nature11461>>

SATTARZADEH, Sam et al. Integrated Grad-Cam: Sensitivity-Aware Visual Explanation of Deep Convolutional Networks Via Integrated Gradient-Based Scoring. [s. l.], p. 1775–1779, 2021.

SCOTT, S. L.; BARRUFET, M. A. Worldwide assessment of industry leak detection capabilities for single & multiphase pipelines. **Project Report for the Offshore Technology Research Center**, [s. l.], p. 1–4, 2003. Disponível em: <<http://www.celou.com/res/icelou/medicalres/201011/20101116203055479.pdf>>

SELVARAJU, Ramprasaath R. et al. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In: 2017 IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV) 2017, **Anais...** : IEEE, 2017. Disponível em: <<http://ieeexplore.ieee.org/document/8237336/>>. Acesso em: 27 out. 2019.

SHI, Jihao et al. Real-time leak detection using an infrared camera and Faster R-CNN technique. **Computers and Chemical Engineering**, [s. l.], v. 135, p. 106780, 2020. a.

SHI, Zhongtuo et al. Convolutional neural network-based power system transient stability assessment and instability mode prediction. **Applied Energy**, [s. l.], v. 263, p. 114586, 2020. b.

SOKOLOVA, Marina; LAPALME, Guy. A systematic analysis of performance measures for classification tasks. **Information Processing & Management**, [s. l.], v. 45, n. 4, p. 427–437, 2009. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0306457309000259>>. Acesso em: 27 out. 2019.

SONG, Hyun Min; WOO, Jiyong; KIM, Huy Kang. In-vehicle network intrusion detection

using deep convolutional neural network. **Vehicular Communications**, [s. l.], v. 21, p. 100198, 2020.

SOVACOO, Benjamin K. The costs of failure: A preliminary assessment of major energy accidents, 1907–2007. **Energy Policy**, [s. l.], v. 36, n. 5, p. 1802–1820, 2008. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0301421508000529>>. Acesso em: 20 fev. 2020.

SPEIGHT, James G. **The Chemistry and Technology of Petroleum**. Boca Raton: CRC Press, 2006.

SPEIGHT, James G. Natural gas: A basic handbook. **Natural Gas: A Basic Handbook**, [s. l.], p. 1–462, 2018.

SRIVASTAVA, Nitish et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. **Journal of Machine Learning Research**, [s. l.], v. 15, p. 1929–1958, 2014. Disponível em: <<http://jmlr.org/papers/v15/srivastava14a.html>>

SZEGEDY, Christian et al. Going deeper with convolutions. In: PROCEEDINGS OF THE IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2015, **Anais...** : IEEE Computer Society, 2015. Disponível em: <<https://arxiv.org/abs/1409.4842v1>>. Acesso em: 25 ago. 2020.

TALUKDAR, Jonti; BISWAS, Ayon; GUPTA, Sanchit. Data Augmentation on Synthetic Images for Transfer Learning using Deep CNNs. **2018 5th International Conference on Signal Processing and Integrated Networks, SPIN 2018**, [s. l.], p. 215–219, 2018.

THOMAS, José Eduardo. **Fundamentos de Engenharia e Petróleo**. Rio de Janeiro: Interciência, 2001.

TIJMEN TIELEMAN; HINTON, Geoffrey. **Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude**COURSERA: Neural Networks for Machine Learning, , 2012. Disponível em: <[https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture\\_slides\\_lec6.pdf](https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf)>

UKIL, Arijit et al. IoT healthcare analytics: The importance of anomaly detection. In: PROCEEDINGS - INTERNATIONAL CONFERENCE ON ADVANCED INFORMATION NETWORKING AND APPLICATIONS, AINA 2016, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2016.

VAN RIJSBERGEN, C. J. FOUNDATION OF EVALUATION. **Journal of Documentation**, [s. l.], v. 30, n. 4, p. 365–373, 1974. Disponível em: <<https://www.emerald.com/insight/content/doi/10.1108/eb026584/full/html>>

WANG, Jingfan et al. Machine vision for natural gas methane emissions detection using an infrared camera. **Applied Energy**, [s. l.], v. 257, p. 113998, 2020.

WANG, Ying et al. Improving the performance of catalytic combustion type methane gas sensors using nanostructure elements doped with rare Earth cocatalysts. **Sensors (Basel, Switzerland)**, [s. l.], v. 11, n. 1, p. 19–31, 2011. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/22346565>>

WOHLFARTH, Kay et al. Dense cloud classification on multispectral satellite imagery. In: 2018 10TH IAPR WORKSHOP ON PATTERN RECOGNITION IN REMOTE SENSING, PRRS 2018 2018, **Anais...** : Institute of Electrical and Electronics Engineers Inc., 2018.

XIAO, Chunlei et al. Detecting the Sources of Methane Emission from Oil Shale Mining and Processing Using Airborne Hyperspectral Data. **Remote Sensing**, [s. l.], v. 12, n. 3, p. 537,

2020. Disponível em: <<https://www.mdpi.com/2072-4292/12/3/537>>. Acesso em: 18 set. 2020.
- XIE, Y. et al. Knowledge-based Collaborative Deep Learning for Benign-Malignant Lung Nodule Classification on Chest CT. **IEEE Transactions on Medical Imaging**, [s. l.], v. 38, n. 4, p. 991–1004, 2019.
- YAMASHITA, Rikiya et al. Convolutional neural networks: an overview and application in radiology. **Insights into Imaging**, [s. l.], v. 9, n. 4, p. 611–629, 2018. Disponível em: <<https://insightsimaging.springeropen.com/articles/10.1007/s13244-018-0639-9>>. Acesso em: 22 out. 2019.
- YOKOGAWA. **FC15U CCD Color Camera**. 2009. Disponível em: <[https://web-material3.yokogawa.com/GS85A7C01-11EN\\_003.pdf](https://web-material3.yokogawa.com/GS85A7C01-11EN_003.pdf)>. Acesso em: 18 ago. 2021.
- YOSINSKI, Jason et al. How transferable are features in deep neural networks? [s. l.], 2014. Disponível em: <<https://arxiv.org/abs/1411.1792>>. Acesso em: 21 fev. 2020.
- ZAMAN, Dina et al. A review of leakage detection strategies for pressurised pipeline in steady-state. **Engineering Failure Analysis**, [s. l.], v. 109, p. 104264, 2020. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1350630719300068>>
- ZEIMARANI, Bashir et al. Breast Lesion Classification in Ultrasound Images Using Deep Convolutional Neural Network. **IEEE Access**, [s. l.], v. 8, p. 133349–133359, 2020.
- ZHOU, Bolei et al. Learning Deep Features for Discriminative Localization. In: 2016 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR) 2016, **Anais...** [s.l: s.n.]
- ZHOU, Fei et al. A Generic Automated Surface Defect Detection Based on a Bilinear Model. **Applied Sciences** 2019, Vol. 9, Page 3159, [s. l.], v. 9, n. 15, p. 3159, 2019. Disponível em: <<https://www.mdpi.com/2076-3417/9/15/3159/htm>>. Acesso em: 9 jul. 2021.
- ZHOU, Shangli et al. A Cotton and Flax Fiber Classification Model Based on Transfer Learning and Spatial Fusion of Deep Features. In: **Lecture Notes in Networks and Systems**. [s.l.] : Springer, 2020. v. 96p. 152–162.



## APÊNDICE 1

ARTIGO PUBLICADO NO PERIÓDICO IEEE ACCESS, ISSN 2169-3536, EM  
OUTUBRO DE 2020

<https://doi.org/10.1109/ACCESS.2020.3031683>

IEEE Access  
Multidisciplinary | Rapid Release | Open Access Journal

Received September 18, 2020, accepted October 13, 2020, date of publication October 16, 2020, date of current version October 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3031683

# Applying Convolutional Neural Networks to Detect Natural Gas Leaks in Wellhead Images

ROBERLÂNIO OLIVEIRA MELO, M. G. F. COSTA<sup>✉</sup>, (Member, IEEE),  
AND CICERO F. F. COSTA FILHO<sup>✉</sup>, (Member, IEEE)

Department of Electronic and Computation, Federal University of Amazonas, Manaus 69080-900, Brazil

Corresponding author: Roberlânio Oliveira Melo (roberlaniomelo@gmail.com)

This work, according to Article 48 of Decree n° 6.008/2006, was supported in part by the Samsung Electronics of Amazonia Ltd., through the terms of Federal Law n° 8.387/1991, signed with the Center for Research and Development in Electronics and Information from the Federal University of Amazonas-CETELI/UFAM under Agreement 004; and in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil, under Grant 001.

**ABSTRACT** Detecting natural gas leaks is one of the most important measures in the oil industry for preventing accidents. The literature provides different techniques for detecting natural gas leaks. However, except for previous studies by the authors on this topic, there remains a gap in the literature on leak detection of natural gas using digital images, without the need for sensors or special cameras calibrated for the spectrum of methane molecules. These previous studies used image-processing techniques associated with a novelty filter classifier to detect the presence or absence of visible cloud of hydrocarbon vapors, that is, a natural gas plume in Closed Circuit Television (CCTV) frames installed in onshore wellheads. In this article, we present a new method for detecting natural gas leaks in oil facilities that enhances the results obtained previously, along with the Gradient-weighted Class Activation Mapping Algorithm (Grad-CAM) to identify natural gas leaks. In this new method, convolutional neural networks (CNN) are applied to classify images (CCTV frames) as belonging to classes with or without natural gas leaks in onshore wellheads. Experimental results showed that the best performance model presented an accuracy of 99.78% and false negative rate of 0.00%.

**INDEX TERMS** Gas leak, gas detection, convolutional neural network, onshore wellhead image, gas and oil industry.

## I. INTRODUCTION

Natural Gas (NG), extracted from oil wells, is composed of methane, ethane, propane, butane, pentane and other hydrocarbons, in addition to nonhydrocarbon compounds (water, carbon dioxide, etc.), but 70 to 98% of the NG is composed of methane [1], being a major contributor to the greenhouse effect [2]. In addition to the environmental impacts, the effects of natural gas leak accidents are commonly lethal to humans and capable of destroying oil and gas processing plants [3].

Considering industrial processes involving natural gas, whether in production, processing, transmission and distribution of this fuel, the sources of methane emission to the atmosphere are due to well completions, blowdowns, workovers, reciprocating compressor rod packing, pipe leaks,

compression station leaks, and oil wellhead leaks [4]. Therefore, to mitigate the effects of natural gas leaks on the greenhouse effect, it is mandatory to develop efficient prevention systems, such as natural gas leak detection systems.

Natural gas leaks, besides contributing to global climate change, have catastrophic consequences in industrial facilities, especially when such leaks are followed by an explosion [3], causing significant loss of human lives and financial losses [5]. To mitigate events involving methane leaks and explosions, the options range from area classification to monitoring and leak detection of this gas in industrial processing plants.

From the underground rocks, wellheads extract oil, a gaseous mixture of hydrocarbons and other non-hydrocarbons under high pressure. Leaking gas under high pressure will produce a visible vapor cloud emanating from a wellhead leak site. This cloud is the result of hydrocarbon and water vapors that condense the humidity of the surrounding air [6], [7].

The associate editor coordinating the review of this manuscript and approving it for publication was Zhenhua Guo<sup>✉</sup>.

In the literature, several methodologies have been published regarding leak detection. The biological monitoring methods involve employing people and dogs [8]. The hardware-based methods require specialized sensors for this purpose [9], such as acoustic, optical, electrochemical, and ultrasonic sensors [8]–[10]. Software-based methods, using intelligent computational algorithms [8], analyze, and classify sensor signals that monitor variables, such as flow, pressure, and temperature [9].

Other studies suggest the detection of leakage of combustible gases using image processing, however, the images of these works are thermal images acquired by high-cost hardware, which are infrared cameras specifically for detecting the infrared spectrum of the methane molecule [11]–[14].

Only two previously works, that we published in literature [15], [16], used visible image processing to detect gas leakage. These works employed the novelty filter classifier.

In this article, we also intend to detect gas leakage using visible images. Motivated by the excellent performance of CNNs in image classification tasks and given the lack of studies applying CNN to image classification regarding the presence or not of natural gas leakage (two-class problem), using visible images, we aim at investigating this problem. Therefore, this article proposes a new model for natural gas leak detection, using visible images, based on Convolutional Neural Networks.

The great difference between our work and others that use the infrared spectrum is that in our study we only use images obtained by simple cameras of CCTV systems to detect visible gas clouds.

Another objective of the work is to provide a location map that highlights the most important regions in the image that contributed to the detection of natural gas leakage, that is, this work will generate visual explanations of the leak, which would help untrained users to discern the leak location. For this we used the Gradient-weighted Class Activation Mapping (Grad-CAM) technique [17].

To compare the results of this new method with those obtained by previous studies that used the novelty classifier [15], [16], this work uses the same image dataset of these studies.

To minimize the overfitting, we enrich the data set, through a data augmentation technique and also used the Transfer Learning (TL) technique. In TL, the following pre-trained deep networks were evaluated: AlexNet [18], DenseNet-201 [19], GoogLeNet [20], MobileNet-v2 [21], ResNet-18 [22] and VGG-16 [23].

Therefore, the main objectives of this article are as follows:

- Propose and evaluate the performance of different CNN architectures and choose the one that provides the best performance in natural gas leak detection.
- Compare the best performance model with the results obtained with TL, using AlexNet, DenseNet-201, GoogLeNet, MobileNet-v2, ResNet-18 and VGG-16.
- Compare the performance of the CNN based model with the previous novelty filter based model [15], [16].

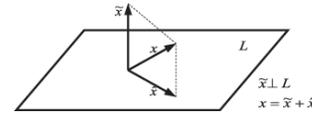


FIGURE 1. Illustration of novelty filter concept in  $R^3$  space. Concept using Gram-Schmidt Orthogonalization Method [16].

- Provide a visual explanation of gas leakage, illustrating the regions of the image that were relevant to the “with leak” class label.

The paper is organized as follows: section II introduces the related; section III provides details of the image datasets used in this work and presents the proposed CNN models and performance evaluation metrics; section IV presents and discusses the results and the last section provides a conclusion.

## II. RELATED WORKS

In [15], [16] the authors described a first natural gas leak detection system, using production head images, based on a surveillance system (CCTV). This system did not need installation of specific sensors. The proposed method was based on the use of the RGB color space component R and the novelty filter concept [24]. The novelty filter classifier obtained an AUC (Area Under Curve ROC - Receiver Operating Characteristic) of 98.53%, with an average error of 0.37%.

The novelty filter concept was described by Kohonen and Oja [24], based on the classic Gram-Schmidt orthogonalization method. Let  $\{x_1, x_2, \dots, x_m\}$  be a set of  $n$  dimensional Euclidian vectors which span an  $m$  dimensional subspace  $L \subset R^n$ , with  $m < n$ . Considering the subspace  $L$ , an arbitrary vector  $x \in R^n$  can be divided into two components,  $\hat{x}$  and  $\tilde{x}$  and Figure 1 illustrates these two components in  $R^3$  space, where:  $\hat{x} \in L$  is the projection of  $x$  on  $L$  and  $\tilde{x} \perp L$  is the projection of  $x$  perpendicular to  $L$ . Vector  $\hat{x} \in L$  is the component of  $x$  that is “known” by the subspace  $L$  and can be defined as a linear combination of  $\{x_1, x_2, \dots, x_m\}$ . Vector  $\tilde{x} \perp L$  is the new information, that is unknown by the subspace  $L$  and cannot be represented as a linear combination of  $\{x_1, x_2, \dots, x_m\}$  [15].

In this method, the magnitude of the vector  $\tilde{x}$ ,  $\|\tilde{x}\|$ , is considered a measure of dissimilarity, that is, the higher the value  $\|\tilde{x}\|$ , the greater the probability that an image  $X$  is not part of subspace  $L$ . Therefore, considering subspace  $L$ ,  $\tilde{x}$  is named novelty and the system that extracts this component from  $x$  and shows it as an output can be named the Novelty Filter. In terms of pattern recognition, the novelty filter novelty rule can be applied to solve problems in which it is desired to classify a pattern as belonging or not to a class, subspace  $L$ .

Convolutional Neural Networks (CNN) [25] have been delivering excellent results in solving problems of different image processing areas that have a texture similar to the natural gas clouds used in this work. The first one work is this of cloud classification [26] and other is fire and smoke detection [27].



**FIGURE 2.** (a) image used by [26] to detect clouds and (b) image used by [27] to detect smoke.

We highlight the studies of cloud classification [26] and fire and smoke classification [27], as their images have a similar texture to the natural gas leak clouds. Figures 2(a) and 2(b), show clouds and smoke images, respectively, while Figures 4(b) and 4(d) show gas leak clouds.

To detect and distinguish thin clouds from thick clouds, Xie *et al.* [26] combine the superpixel approach with deep learning. The experiment dataset consists of 81 satellite images. From these images, patches were extracted with thick clouds, thin clouds and cloudless, totaling 26,000, 22,000 and 156,000 patches, respectively. They obtained a hit rate of 0.9454 and an error rate of 0.0330.

Namozov and Cho [27] applied the CNN concepts to fire and smoke detectors, but instead of using traditional rectified linear units or tangent functions, they used an adaptive piecewise linear unit as activation function of the hidden layers. The original dataset of this study is composed of 2,440 images, divided into 1,220 smoke images and 1,220 fire images. To overcome the overfitting problem, they used a combination of traditional image transformation methods with Generative Adversarial Networks (GANs) to create a large dataset. An accuracy and a false alarm rate of 0.9485 and 0.63, respectively, were obtained.

Wang *et al.* [14] proposed the detection of methane leaks using a database composed of infrared images obtained by a special camera, FLIR GF-320 infrared camera, whose imaging sensor is specifically sensitized to detect methane molecules. These authors applied CNNs for leak detection. The overall detection accuracy was close to 95%.

Shi *et al.* [13] also used the FLIR-320 infrared camera, the same hardware used by Wang *et al.* [14], to detect methane leaks. The authors used the Faster Region-Convolutional Neural Network technique to analyze the videos, achieving an average accuracy of 98%.

### III. MATERIALS AND METHODS

The block diagram of Figure 3 presents the main steps of the methodology employed in this work to detect natural gas leak. Initially, at the data preparation block, the dataset images go through a preprocessing stage. After, we used a data augmentation technique. This augmented dataset is divided into three subsets: training, validation and testing. Section III.A presents the materials of this work and the steps just described.

In the model building block, three CNN architectures are proposed. Section III.B presents these architectures. In training these architectures, to obtain a network with better performance, we evaluated different optimization algorithms and

regularization techniques. The training, validation and testing of the proposed architectures are detailed in section III.C. Once the best performance network and its parameters have been chosen, pre-trained networks are implemented through TL to perform the classification task. TL networks are presented in section III.D. Finally, we used the best performance network to visualize the class activation map. This theme is presented in section III.E.

#### A. DATASET

The original image dataset used in previous studies with the novelty filter is comprised of 2,810 images [16]. The size of these images is  $184 \times 160$  pixels. However, in a recent study [28], this set was supplemented with images from the same wellhead, totaling 3,060 samples.

In this work, to reduce possible overfitting [29], we used the largest dataset. In addition, images with similar content were excluded. These excluded images were obtained from consecutive frames, in which the time between two consecutive images was  $1/24$  seconds. In this period the shape of the vapor cloud did not change. After exclusion, the total number of images was 2,980. This dataset is divided into 2,000 images of the “without leak” class and 980 of the “with leak” class. Figure 4 illustrates some of these images.

To obtain leak images, natural gas leak simulations were performed at the wellhead, by opening a needle valve intentionally and in a controlled way. To prevent accidents or needless damage to the environment, the volume of gas released during the leak simulations was the minimum necessary to collect the images to be used in this work.

Regarding the dimensions of the dataset, Aggarwal [25] states that CNNs should work with square images. Because the images have the dimensions of  $184 \times 160$  pixels, we chose to resize all images to a standard size of  $186 \times 186$  pixels. Aiming to prevent overfitting, we used the technique of on-the-fly data augmentation, also known as online augmentation [30]. In our proposal, the applied transformations were the following: rotated  $\pm 20^\circ$ , horizontal and vertical translation of  $\pm 3$  pixels, and a reduction of the image size with a factor between 0.5 and 1.0 [31]. These operations resulted in an augmented dataset with 10,132 images. Figure 5 illustrates such transformations

#### B. CNN ARCHITECTURE AND TRAINING PARAMETERS

According with Aggarwal [25], the increase in the number of hidden units in CNN increases the complexity of the modeled function, which can be useful for modeling difficult problems. However, this can cause overfitting if the problem is not so complex or when the dataset is small. A way to decrease complexity is to increase the length of the network by decreasing the number of units hidden in each layer [25]. This procedure was used by Miyagawa *et al.* [32]. The first CNN architecture proposed in this study for leak detection of natural gas, referred to as Long CNN (LCNN), is a serial architecture with 18 convolutional layers, based on the work of Miyagawa *et al.* [32]. Figure 6(a) illustrates this architecture. The width  $\times$  length and depth of the first convolution

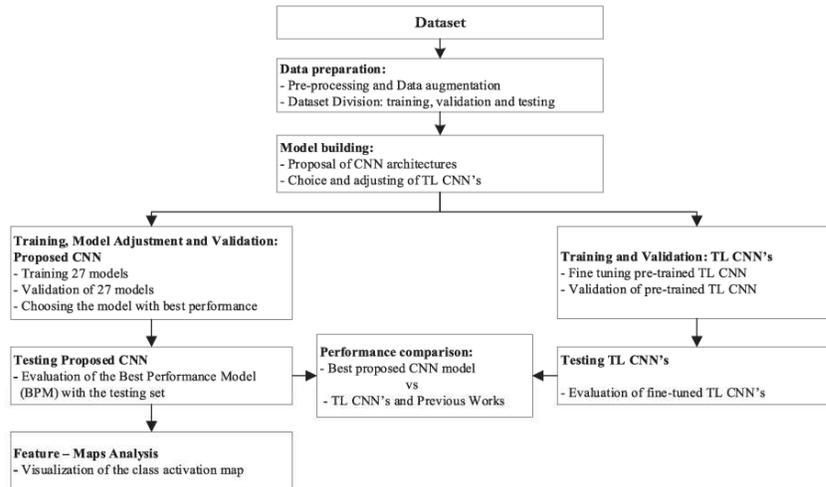


FIGURE 3. Overview of the proposed approach.

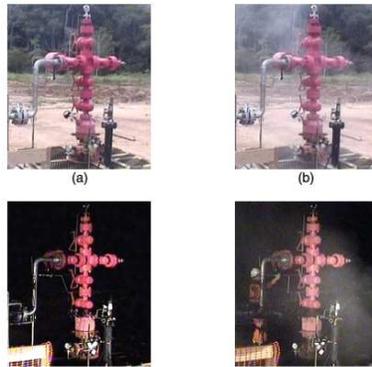


FIGURE 4. Sample images from the image dataset: (a) Daytime image without gas leak (without the hydrocarbon cloud), (b) Daytime image with gas leak (with the hydrocarbon cloud), (c) Nighttime image without gas leak (without the hydrocarbon cloud), (d) Nighttime image with gas leak (with the hydrocarbon cloud).

layer has the following dimensions:  $186 \times 186 \times 16$ . Depth convolution layers have shorter width and length. The last convolution layer has a feature map with dimension of  $11 \times 11 \times 128$ .

To investigate whether the LCNN was complex enough to solve the problem, with possible occurrence of overfitting, even after this tapering of units in the hidden layers, we proposed two other architectures with a lower number of convolution layers. These new architectures, also shown in Figure 6, are the following: Medium architecture (MCNN) with 6 convolution layers. And Short architecture (SCNN) with 2 convolution layers.

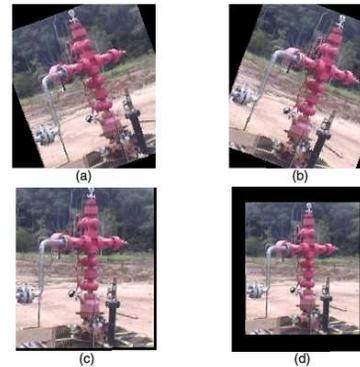
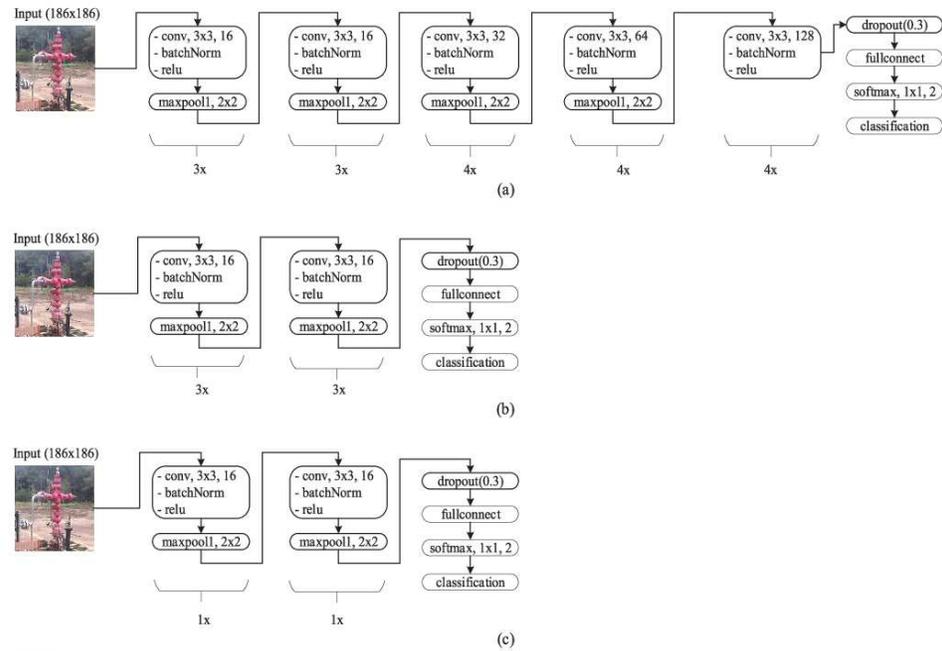


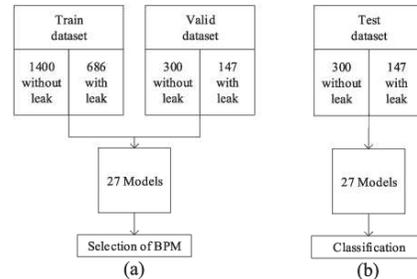
FIGURE 5. Visualization of examples of random transformations applied to four images of the dataset: (a) rotation of  $+20^\circ$ ; (b) rotation of  $-20^\circ$ ; (c) translation of  $-3$  by  $-3$  pixels; (d) image resizing with factor 0.8.

Each convolution layer is followed by a batch normalization layer [33], as a method to address the vanishing and exploding gradient problems [25], and a ReLU layer as the activation function. As shown in Figure 6, to reduce the spatial size of the feature map in depth layers, we applied max-pooling operations with  $2 \times 2$  kernels, with stride  $1 \times 1$ . Each architecture has a classification step composed of the following layers: fully connected and softmax layer, which outputs a probability distribution of two classes, and a final classification layer.

To optimize the performance of the architectures, we evaluated the application of the following technique for improving generalization: Dropout [34], which randomly eliminates neurons of a given CNN layer, at a rate of 30%.



**FIGURE 6.** Proposed network architectures: a) LCNN, with 18 convolution layers; b) MCNN, with 6 convolution layers; c) SCNN, with 2 convolution layers.



**FIGURE 7.** Block diagram for training, evaluation and testing: (a) CNN training and evaluation to obtain the best performance model (BPM model); (b) BPM test with test dataset.

and L2 regularization [35], that regularizes the weights in the loss function with a smoothing coefficient of 0.0001.

The following optimization algorithms were evaluated: Adaptive Moment Estimation (ADAM) with decay rates  $\beta_1$  of 0.9 and  $\beta_2$  of 0.9 [36]; Root Mean Square Propagation (RMSProp) with decay rates  $\beta_2$  of 0.9 [37], and Stochastic Gradient Descent with Momentum (SGDM), with momentum of 0.9 [38].

Other training hyperparameters are fixed as follows: maximum number of epochs is 15, mini-batch size is 16 and

**TABLE 1.** Contingency table used to calculate  $\chi^2$ .

CNN	Novelty filter
Number of correct classifications	Number of correct classifications
Number of wrong classifications	Number of wrong classifications

learning is 0.001. Finally, for training stop, we used early stop with a validation patience of 4. This means that the training is interrupted when the loss in the validation set increases in four consecutive mini-batches [39], [40].

The combination of three architectures, three optimization algorithms, and two techniques for improving generalization resulted in twenty-seven distinct network models being analyzed.

### C. TRAINING, VALIDATION AND TESTING

To evaluate the 27 models, the dataset is divided into the following proportions: 70% training (*train* dataset), 15% validation (*valid* dataset), and 15% testing (*test* dataset).

The methodology used for training, validation and testing is shown in Figure 7. As shown in Figure 7(a), the *valid* dataset is used to choose the Best Performance Model (BPM). After choosing the BPM, it is tested with the *test* dataset as shown in Figure 7(b).

The following metrics were used in evaluating the performance of the CNN models: accuracy, false negative rate,

**TABLE 2.** Performance of the 27 CNN models with the valid dataset showing the metrics accuracy (ACC), false negative rate (FNR) precision, recall, F1-score and chi-square ( $\chi^2$ ) test performed comparing the results of the CNN with the results of [16].

Architecture	Optimization	Regularization	Model	ACC (%)	FNR (%)	Precision (%)	Recall (%)	F1-Score (%)	$\chi^2$
LCNN	ADAM	Dropout	1	99.33	2.04	100.0	97.96	98.97	25.57
		L <sub>2</sub>	2	99.55	1.36	100.0	98.64	99.32	27.70
		L <sub>2</sub> + Dropout	3	99.55	1.36	100.0	98.64	99.32	27.70
	RMSProp	Dropout	4	76.06	72.79	100.0	27.21	42.78	104.35
		L <sub>2</sub>	5	99.55	1.36	100.0	98.64	99.32	27.70
		L <sub>2</sub> + Dropout	6	99.55	1.36	100.0	98.64	99.32	27.70
	SGDM	Dropout	7	<b>99.78</b>	<b>0.00</b>	99.32	<b>100.0</b>	<b>99.66</b>	<b>29.95</b>
		L <sub>2</sub>	8	99.55	1.36	100.0	98.64	99.32	27.70
		L <sub>2</sub> + Dropout	9	99.55	1.36	100.0	98.64	99.32	27.70
MCNN	ADAM	Dropout	10	97.32	8.16	100.0	91.84	95.74	10.75
		L <sub>2</sub>	11	99.55	1.36	100.0	98.64	99.32	27.70
		L <sub>2</sub> + Dropout	12	97.32	8.16	100.0	91.84	95.74	10.75
	RMSProp	Dropout	13	95.53	13.61	100.0	86.39	92.70	3.28
		L <sub>2</sub>	14	97.32	8.16	100.0	91.84	95.74	10.75
		L <sub>2</sub> + Dropout	15	92.17	23.81	100.0	76.19	86.49	0.51
	SGDM	Dropout	16	99.55	1.36	100.0	98.64	99.32	27.70
		L <sub>2</sub>	17	97.32	8.16	100.0	91.84	95.74	10.75
		L <sub>2</sub> + Dropout	18	95.08	14.97	100.0	85.03	91.91	2.12
SCNN	ADAM	Dropout	19	97.32	8.16	100.0	91.84	95.74	10.75
		L <sub>2</sub>	20	97.32	8.16	100.0	91.84	95.74	10.75
		L <sub>2</sub> + Dropout	21	98.88	3.40	100.0	96.60	98.27	21.61
	RMSProp	Dropout	22	99.55	1.36	100.0	98.64	99.32	27.70
		L <sub>2</sub>	23	99.33	2.04	100.0	97.96	98.97	25.57
		L <sub>2</sub> + Dropout	24	98.66	4.08	100.0	95.92	97.92	19.78
	SGDM	Dropout	25	94.86	15.65	100.0	84.35	91.51	1.65
		L <sub>2</sub>	26	99.55	1.36	100.0	98.64	99.32	27.70
		L <sub>2</sub> + Dropout	27	99.55	1.36	100.0	98.64	99.32	27.70

precision, recall and F1 score [41]. To present these metrics, we need the following definitions: True-Positives (TP): images with natural gas leak that CNN has classified as such; True-Negatives (TN): images without natural gas leak that CNN has classified as such; False-Positives (FP): images without natural gas leak that CNN classified as with natural gas leak and False-Negatives (FN): images with natural gas leak that CNN has classified as without natural gas leak.

The accuracy (ACC), shown in Equation (1), was used to define the BPM from the 27 proposed models.

In many machine learning applications, it is crucial to minimize the False Negatives Rate (FNR), as is the case with medical diagnoses, where a late or mistaken diagnosis can reduce the effectiveness of treatment [42]. Also, in autonomous vehicles, a small failure could cause a momentary malfunction in the car, rendering it a hazard [43]. This also applies to the oil industry: a false negative, when an actual leakage is classified as non-leakage, can lead to serious accidents. FNR is defined according to Equation (2). Other metrics are also used in this work to evaluate model performance, such as: precision [41], which determines the rate of positive samples that were correctly classified, that is, the ability of the model to avoid false positive; *recall* or sensitivity [41], which determines

the rate of true positives that the classifier can provide and *F1-Score* [44] with values between 0 and 1, which represents a harmonic mean of precision and sensitivity. These metrics are shown in Equations (3)-(5).

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$FNR = \frac{FN}{FN + TP} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (5)$$

In this work, we may be interested in answering whether the performance of the BPM is better than the performance obtained in the previous work of Melo and Costa [16]. To answer this question, we formulate a null hypothesis that the results are similar. To evaluate this null hypotheses we employed the Chi-Square Test [45], using the contingency table shown in Table 1. The values of Table 1 were collected in the *valid* dataset.

To confirm the results presented in this work, we adopted a significance level of 99% ( $\alpha = 0.01$ ). The null hypothesis ( $H_0$ ) proposes that no difference exists between the two methods. Considering that the dataset used for comparison between the two methods has  $N = 447$  samples (*valid* dataset) and with a degree of freedom of 1, we have the value of  $\chi^2_{critical} = 6.63$ . If the value of  $\chi^2_{calculated} > \chi^2_{critical}$ , the null hypothesis is rejected.

The hardware used in this work has the following configuration: Intel<sup>TM</sup><sub>i7</sub> processor 8th generation and GeForce MX150 GPU of 2 GB and software used was Matlab<sup>TM</sup> version 2019b.

#### D. TRANSFER LEARNING

As stated by Yosinski *et al.* [46], when the dataset is limited, transfer learning can be a powerful tool to allow training the network without generating overfitting. In this work, six transfer learning models were fine-tuned: AlexNet [18], DenseNet-201 [19], GoogLeNet [20], MobileNet-v2 [21], ResNet-18 [22] and VGG-16 [23]. Transfer learning consists of using a pre-trained model and adapting it to a new dataset [47]. In this work, these models are shallow tuned (only the parameters of the last fully connected layer are tuned) to identify one of two classes: with or without a natural gas leak.

For training stop, both the proposed and transfer learning architectures used the early-stop method, with a patience of 4. A maximum number of 15 epochs was used.

#### E. VISUAL EXPLANATION

The gas leak clouds have slight nuances that are difficult for operators to perceive, leaving them uncertain that they are facing a real leak. Therefore, just identifying an image with a gas leak and informing the operator of its presence may not resolve all doubts. To improve the quality of the information provided to the operator, we employed the Gradient-weighted Class Activation Mapping (Grad-CAM) [17]. This tool provides a visual explanation, showing which region of an input image was relevant for the classifier to make the decision that there was a leak.

Grad-CAM produces heatmaps that highlight specific regions of the image that contributed to CNN prediction. This heat map is obtained by calculating gradients between the last convolution layer and the output of the desired class.

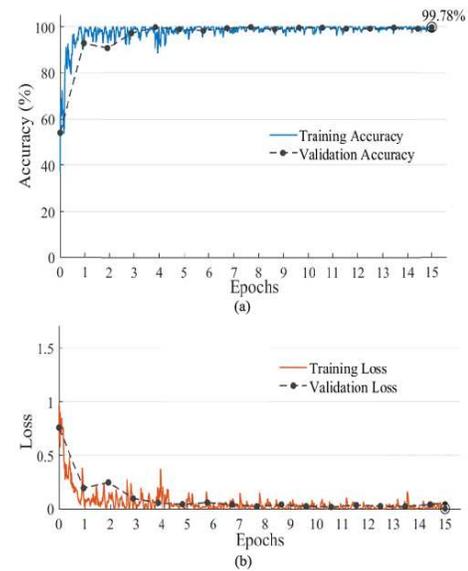
#### IV. RESULTS AND DISCUSSION

Table 2 presents results of the twenty-seven CNN models, showing the metrics presented in the last section. These results were obtained using the *valid* dataset. Among the 27 CNN models, model 7 showed the highest accuracy. This model, which uses the LCNN architecture, the SGDM optimization algorithm and  $L_2$  regularization, obtained an accuracy of 99.78% and an FNR of 0%. Therefore, model 7 is our Best Performing Model (BPM).

In Table 2, models 2, 3, 5, 6, 8, 9, 11, 16, 22, 26, and 27 provided ACC = 99.55%, lower than 99.78% of BPM. However, despite the high accuracy obtained (99.55%), these

**TABLE 3.** Performance comparison between the classifier BPM and novelty filter classifier [16].

Method	ACC (%)	FNR (%)	Precision (%)	Recall (%)	F1 Score (%)
Novelty Filter	93.15	8.78	91.22	91.22	91.22
CNN (BPM)	99.78	0.00	99.32	100.00	99.66

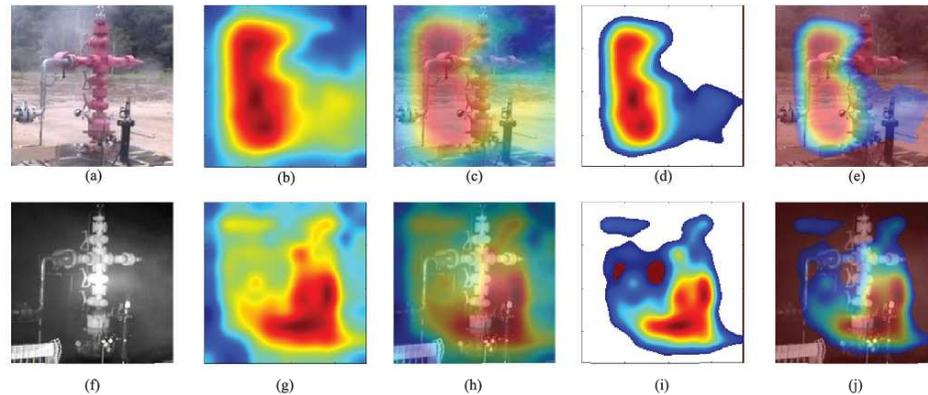


**FIGURE 8.** Training and validation of BPM: (a) accuracy during training and validation; (b) loss during training and validation.

listed models presented FN classifications. This type of error can put an installation at risk, as classifying a positive image as negative for natural leakage means neglecting the potential risk of an accident. The BPM, besides presenting the highest accuracy (99.78%) among all models, is the only one free of FNs, that is, FNR = 0%.

Most of the models shown in Table 2 presented a higher performance compared with the previous method using the novelty filter [16], with a statistically significant difference ( $\chi^2_{calculated} > 6.63$ ). The method using the novelty filter presented an accuracy of 93.15%. In this work, as shown in Table 2, BPM achieved an accuracy of 99.78%.

As shown in Table 2, the value of  $\chi^2_{calculated} = 29.95$  for the BPM is greater than the  $\chi^2_{critical}$ . Thus, the null hypothesis is rejected. Therefore, CNN performed better than the novelty filter classifier, and this difference is statistically significant. Table 3 presents a comparison between the metrics obtained with the BPM model and the metrics obtained with the novelty filter technique.



**FIGURE 9.** (a) and (f) examples of original images “with leak”; (b) and (g) heatmap of the original images; (c) and (h) locating the class discriminatory regions through Grad-CAM; (d) and (i) new heatmap with pre-processed magnitudes; (e) and (j) identification of the pixels that contributed to the “with leak” prediction. The pixels with red color contribute strongly to the score of the “with leak” class.

The accuracy and loss graphs during training and validation of the BPM model are shown in Figure 8(a) and Figure 8(b) respectively. The training accuracy was 100% and the validation accuracy was 99.78%. Classification of the *test* dataset (not used during training), also resulted in an accuracy of 99.78%, but with one FP.

#### A. PRE-TRAINED CNNs PERFORMANCE IN RELATION TO BPM

Among the pre-trained CNNs, DenseNet-201 and AlexNet provided the best results. As shown in Table 4, the accuracies obtained with pre-trained networks were  $ACC_{DenseNet} = 99.55\%$  and  $ACC_{AlexNet} = 99.33\%$ . Therefore, the performance of the two pre-trained networks are lower than BPM, which presented an accuracy of 99.78%. This demonstrates the superiority of the BPM model compared with DenseNet-201 and AlexNet, and over other pre-trained networks.

According to Table 4, MobileNet-v2 presented the worst value for recall, 91.84%. The lower accuracy of MobileNet-v2 compared to BPM is due to the number of FNs. The accuracy of DenseNet-201 is 99.55%, close to the BPM, but the DenseNet-201 presented an FNR of 1.36%. This type of failure can be catastrophic in the oil industry, delaying or obstructing decisions that could prevent the explosion of a natural gas cloud.

With AlexNet, using the *test* dataset, an FNR of 2.04% was obtained, with three FNs and zero FP. With DenseNet-201, an FNR of 1.36% was obtained, with two FNs and zero FP.

#### B. VISUAL EXPLANATIONS

After defining the BPM, it was used to evaluate images with a gas leak using the Grad-CAM tool, aiming to identify the gas cloud. Two original images with a gas leak are shown in Figure 9(a) and 9(f). With Grad-CAM a heatmap is obtained that shows, in the original images, the pixels that have more influence in the CNN prediction of the “with leak” class.

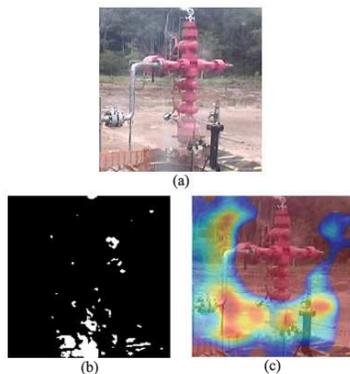
**TABLE 4.** Performance comparison between the classifier BPM with the pre-trained networks: AlexNet, DenseNet-201, GoogLeNet, MobileNet-v2, ResNet-18 and VGG-16.

Method	ACC (%)	FNR (%)	Precision (%)	Recall (%)	F1 Score (%)	Time training and validation (min:sec)	Average time test (sec)
<b>BPM</b>	<b>99.78</b>	<b>0.00</b>	<b>99.32</b>	<b>100.00</b>	<b>99.66</b>	<b>11:46</b>	<b>0.0055</b>
DenseNet-201	99.55	1.36	100.00	98.64	99.32	397:29	0.0484
AlexNet	99.33	2.04	100.00	97.96	98.97	11:22	0.0049
GoogLeNet	98.66	4.08	100.00	95.92	97.92	21:27	0.0096
ResNet-18	98.43	4.76	100.00	95.24	97.56	16:04	0.0082
VGG-16	97.76	6.80	100.00	93.20	96.48	73:19	0.0228
MobileNet-v2	97.32	8.16	100.00	91.84	95.74	52:26	0.0180

Figure 9(b) and 9(g) show the heatmaps corresponding to original images shown in Figure 9(a) and 9(f), respectively. In these heatmaps, Grad-CAM shows that pixels in red color have a strong correlation with the “with leak” class. In Figure 9(c) and 9(h) we superimposed these red regions over the original images. As shown in these figures, there is a high coincidence between the red pixels and the gas cloud. To better highlight the regions strongly correlated with the prediction of the class “with gas”, we calculated the average value of this heatmap and excluded activations with scores below this average value, resulting in the new heatmap shown in Figure 9(d) and 9(i).

Finally, in Figure 9(e) and 9(f) we superimposed the heatmaps of Figure 9(d) and 9(j) over the input images, providing better visualization of the pixels of the input image that contribute to the prediction of gas leakage.

In Figure 10(a) we show an original image with gas leak that was incorrectly classified as with no gas leak in a previous work [15] and correct classified as with gas leak



**FIGURE 10.** (a) Image classified as FN in previous work [15] and properly classified (TP) by BPM; (b) novelty region obtained by the technique proposed in [15]; (c) heatmap of the original image in this work.

in present work. Figure 10(b) shows the region detected as a novelty in [15], while Figure 10(c) shows the heatmap obtained with the BPM of the present work. Visually, the novelty region obtained by the novelty filter is lower than the region of high magnitude scores provided by the Grad-CAM heat map, thus suggesting that our technique has greater sensitivity than the previous one [15].

#### V. CONCLUSION

This study proposed a methodology using deep learning tools, more specifically convolutional neural networks, for digital image classification of a wellhead into two classes, “with natural gas leak” and “without natural gas leak”. It is noteworthy that these images were obtained through CCTV surveillance systems already existing in the wellhead site, with no demand for additional facilities. To achieve our goal, we evaluated the performance of 27 different CNN models. The best performing model, BPM, had the following characteristics: 18 convolution layer architecture, SGDM optimization algorithm and dropout regularization technique. The accuracy obtained in the test set, 99.78%, is higher than similar studies found in the literature, which used the novelty classifier technique [16]. After applying a Chi-square ( $\chi^2$ ) hypothesis testing, it became clear that this higher performance is significant at a 99% significance level. The aforementioned accuracy obtained in the test dataset is close to the accuracy obtained in the train set, 100%, suggesting that the generalization techniques employed avoid overfitting. We also showed that the BPM outperforms pre-trained networks AlexNet, DenseNet-201, GoogLeNet, MobileNet-v2, ResNet-18 and VGG-16. Additionally, these pre-trained networks result in FNs, while BPM presented no FNs. The Grad-CAM tool proved to be of great importance for energy facility operators, as it shows, with heatmaps, the location of the gas cloud in original images. For future works we aim at evaluating the generalization of the model on unseen images of different types of wellheads.

#### ACKNOWLEDGMENT

Academic English Solutions (<https://www.academicenglishsolutions.com>) revised this article.

#### REFERENCES

- [1] W. D. McCain, *Properties of Petroleum Fluids*. Tulsa, OK, USA: Pennwell Books, 2017.
- [2] C. MAgee, “Survey of natural gas leakage abatement best practices,” California Public Utilities Commission, San Francisco, CA, USA, Tech. Rep. R.15-01-008 CEK/ek4, 2015.
- [3] Y. F. Khalil, “A probabilistic visual-flowcharting-based model for consequence assessment of fire and explosion events involving leaks of flammable gases,” *J. Loss Prevention Process Industries*, vol. 50, pp. 190–204, Nov. 2017.
- [4] M. Fulton, N. Mellquist, S. Kitasei, and J. Bluestein, “Comparing life-cycle greenhouse gas emissions from natural gas and coal,” Deutsche Bank Group, DB Climate Change Advisors, Frankfurt, Germany, Tech. Rep., 2011.
- [5] M. T. Dröge and R. Kenter, “Gas pipeline incidents: 10th report of the european gas pipeline incident data group,” Tech. Rep. VA 17.R.0395, Mar. 2018, no. 10, pp. 1–50.
- [6] G. Atkinson, J. Hall, and A. McGillivray, “Review of vapour cloud explosion incidents,” Health Saf. Executive, Harpur Hill, U.K., RR1113 Res. Rep., 2017.
- [7] Board and B. M. I. Investigation, “Buncefield major incident investigation: Initial report to the health and safety commission and the environment agency of the investigation into the explosions and fires at the buncefield oil storage and transfer depot, Hemel Hempstead, 11 December 2005,” Health Safety Commission, Hertfordshire, U.K., Tech. Rep., 2006.
- [8] M. A. Adegboye, W.-K. Fung, and A. Karnik, “Recent advances in pipeline monitoring and oil leakage detection technologies: Principles and approaches,” *Sensors*, vol. 19, no. 11, p. 2548, Jun. 2019.
- [9] K. Sachedina and A. Mohany, “A review of pipeline monitoring and periodic inspection methods,” *Pipeline Sci. Technol.*, vol. 2, no. 3, pp. 187–201, Sep. 2018.
- [10] P.-S. Murvay and I. Silea, “A survey on gas leak detection and localization techniques,” *J. Loss Prevention Process Industries*, vol. 25, no. 6, pp. 966–973, Nov. 2012.
- [11] M. S. Jadin and K. H. Ghazali, “Gas leakage detection using thermal imaging technique,” in *Proc. UKSim-AMSS 16th Int. Conf. Comput. Modeling Simulation*, Mar. 2014, pp. 302–306.
- [12] B. Liu, H. Ma, X. Zheng, L. Peng, and A. Xiao, “Monitoring and detection of combustible gas leakage by using infrared imaging,” in *Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST)*, Oct. 2018, pp. 1–6.
- [13] J. Shi, Y. Chang, C. Xu, F. Khan, G. Chen, and C. Li, “Real-time leak detection using an infrared camera and faster R-CNN technique,” *Comput. Chem. Eng.*, vol. 135, Apr. 2020, Art. no. 106780.
- [14] J. Wang, L. P. Tchappmi, A. P. Ravikumar, M. McGuire, C. S. Bell, D. Zimmerle, S. Savarese, and A. R. Brandt, “Machine vision for natural gas methane emissions detection using an infrared camera,” *Appl. Energy*, vol. 257, Jan. 2020, Art. no. 113998.
- [15] C. F. F. C. Filho, R. de O. Melo, and M. G. F. Costa, “Detecting natural gas leaks using digital images and novelty filters,” in *Proc. 3rd Int. Conf. AIS*, vol. 7326 LNAI. London, U.K.: Springer, 2012, pp. 242–249.
- [16] R. O. de Melo, M. G. F. Costa, and C. F. F. C. Filho, “Using digital image processing and a novelty classifier for detecting natural gas leaks,” in *Studies in Computational Intelligence*, vol. 542, L. Chen, S. Kapoor, and R. Bhatia, Eds. Cham, Switzerland: Springer, 2014, pp. 409–422.
- [17] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual explanations from deep networks via gradient-based localization,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.
- [18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Proc. 26th Annu. Conf. Neural Inf. Process. Syst.*, vol. 2, Dec. 2012, pp. 1097–1105.
- [19] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.
- [20] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

- [21] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.
- [22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent.*, 2015, pp. 1–14.
- [24] T. Kohonen and E. Oja, "Fast adaptive formation of orthogonalizing filters and associative memory in recurrent networks of neuron-like elements," *Biol. Cybern.*, vol. 21, no. 2, pp. 85–95, Jun. 1976.
- [25] C. C. Aggarwal, *Neural Networks and Deep Learning*. Cham, Switzerland: Springer, 2018.
- [26] F. Xie, M. Shi, Z. Shi, J. Yin, and D. Zhao, "Multilevel cloud detection in remote sensing images based on deep learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3631–3640, Aug. 2017.
- [27] A. Namozov and Y. I. Cho, "An efficient deep learning algorithm for fire and smoke detection with limited data," *Adv. Electr. Comput. Eng.*, vol. 18, no. 4, pp. 121–128, 2018.
- [28] R. O. Melo, C. F. F. C. Filho, and M. G. F. Costa, "Leak detection of natural gas with base on the components of color spaces RGB and HSI using novelty filter," *IEEE Latin Amer. Trans.*, vol. 12, no. 8, pp. 1560–1565, Dec. 2014.
- [29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
- [30] P. Molchanov, S. Gupta, K. Kim, and J. Kautz, "Hand gesture recognition with 3D convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2015, pp. 1–7.
- [31] J. Talukdar, A. Biswas, and S. Gupta, "Data augmentation on synthetic images for transfer learning using deep CNNs," in *Proc. 5th Int. Conf. Signal Process. Integr. New. (SPIN)*, Feb. 2018, pp. 215–219.
- [32] M. Miyagawa, M. G. F. Costa, M. A. Gutierrez, J. P. G. F. Costa, and C. F. F. C. Filho, "Detecting vascular bifurcation in IVOCCT images using convolutional neural networks with transfer learning," *IEEE Access*, vol. 7, pp. 66167–66175, 2019.
- [33] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn.*, vol. 1, 2015, pp. 448–456.
- [34] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [35] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.
- [36] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, CA, USA, 2014. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [37] T. Tieleman and G. Hinton, "Lecture 6.5-RMSPROP: Divide the gradient by a running average of its recent magnitude," *COURSERA, Neural New. Mach. Learn.*, vol. 4, no. 2, pp. 26–31, 2012.
- [38] G. E. Hinton, "A practical guide to training restricted Boltzmann machines," in *Neural Networks: Tricks of the Trade*, G. Montavon, G. B. Orr, and K.-R. Müller, Eds., 2nd ed. Berlin, Germany: Springer, 2012, pp. 599–619.
- [39] B. S. Rem, N. Käming, M. Tamowski, L. Asteria, N. Fläschner, C. Becker, K. Sengstock, and C. Weitenberg, "Identifying quantum phase transitions using artificial neural networks on experimental data," *Nature Phys.*, vol. 15, no. 9, pp. 917–920, Sep. 2019.
- [40] C. Kofler, R. Muhr, and G. Spöck, "Classifying image stacks of specular silicon wafer back surface regions: Performance comparison of CNNs and SVMs," *Sensors*, vol. 19, no. 9, p. 2056, May 2019.
- [41] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manage.*, vol. 45, no. 4, pp. 427–437, Jul. 2009.
- [42] D. S. Kermany, M. Golubau, W. Cai, C. C. S. Valentim, H. Liang, S. L. Baxter, A. McKeown, and G. Yang, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131.e9, Feb. 2018.
- [43] H. M. Song, J. Woo, and H. K. Kim, "In-vehicle network intrusion detection using deep convolutional neural network," *Veh. Commun.*, vol. 21, Jan. 2020, Art. no. 100198.
- [44] C. J. Van Rijsbergen, "Foundation of evaluation," *J. Documentation*, vol. 30, no. 4, pp. 365–373, Apr. 1974.
- [45] D. C. Howell, *Statistical Methods for Psychology*, 7th ed. Belmont, CA, USA: Cengage Wadsworth, 2010.
- [46] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 2, 2014, pp. 3320–3328.
- [47] A. Darwish, D. Ezzat, and A. E. Hassanien, "An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis," *Swarm Evol. Comput.*, vol. 52, Feb. 2020, Art. no. 100616.



**ROBERLÂNIO OLIVEIRA MELO** received the B.Sc. degree from Paulista University, in 2009, and the master's degree from the Electrical Engineering from the Federal University of Amazonas, in 2012, where he is currently pursuing the Ph.D. degree. His current research interests include neural networks and artificial intelligence, computer vision, and image processing.



**M. G. F. COSTA** (Member, IEEE) received the Electrical Engineering degree from the Federal University of Amazonas, Brazil, in 1983, and the M.Sc. and D.Sc. degrees in electrical engineering from the State University of Campinas, Brazil, in 1988 and 1996, respectively. She has been with the Federal University of Amazonas, since 1990, where she is currently a Full Professor with the Electrical Engineering Program, Federal University of Amazonas. Her research interests include biomedical image and artificial.



**CICERO F. F. COSTA FILHO** (Member, IEEE) received the Electrical Engineering degree from the Federal University of Pernambuco, Brazil, in 1982, and the M.Sc. and D.Sc. degrees in electrical engineering from the State University of Campinas, Brazil, in 1985 and 1996, respectively. He is currently a Full Professor with the Electrical Engineering Program, Federal University of Amazonas. His research interests include biomedical image and artificial intelligence.

\*\*\*

## APÊNDICE 2

ARTIGO CIENTÍFICO SUBMETIDO PARA PUBLICAÇÃO NO PERIÓDICO

COMPUTERS & ELECTRICAL ENGINEERING, ISSN: 0045-7906

<https://www.journals.elsevier.com/computers-and-electrical-engineering>

### Computers and Electrical Engineering

#### Natural Gas Leaks Detection in Oil Facilities Analyzing Closed-Circuit Television Frames Using Transfer Learning and Feature Fusion

--Manuscript Draft--

<b>Manuscript Number:</b>	COMPELECENG-D-21-03267
<b>Article Type:</b>	Research Paper
<b>Keywords:</b>	Gas Leak; Gas Detection; Convolutional Neural Network; Feature Fusion Gas and Oil Industry
<b>Corresponding Author:</b>	Roberlano Melo, Ph.D. Universidade Federal do Amazonas Manaus, AM BRAZIL
<b>First Author:</b>	Roberlano Oliveira Melo, Ph.D.
<b>Order of Authors:</b>	Roberlano Oliveira Melo, Ph.D. Marly Guimarães Fernandes Costa Cícero Ferreira Fernandes Costa Filho
<b>Abstract:</b>	<p>Large natural gas leaks lead to accidents that can be lethal to humans and highly destructive to oil installations, as they release methane into the atmosphere, negatively impacting our environment. Therefore, the detection of natural gas leaks is essential for the oil industry to prevent accidents. The literature provides different techniques for detecting natural gas leaks. However, there is still a gap in the literature on detecting clouds of natural gas leaks through digital images, without the need for specific sensors or cameras calibrated for the spectrum of methane molecules. In this paper, a method for detecting visible clouds of hydrocarbon vapors is presented, consisting of convolutional neural networks, applied to classify images (closed-circuit television frames) as classes "with" or "without" leakage of natural gas in onshore wellheads, improving the results previously obtained. This study proposes and presents the results of investigating the transfer learning technique with the pre-trained architectures DenseNet-201, GoogLeNet, MobileNetV2, and ResNet-18, and the feature fusion technique between these pre-trained models, to predict, through image analysis of an onshore oil wellhead, the presence or absence of natural gas leakage in this installation. Finally, the study provides a visual explanation of the region of the input image that was relevant for the classifier to decide that a leak occurred through the Gradient-weighted Class Activation Mapping Algorithm. The experimental results showed that the best-performing model presented an accuracy of 100.00% and a false negative rate of 0.0%, surpassing the performance of previous methods of detecting leakage of natural gas by closed-circuit television images.</p>